

https://dspace.org/

Algorithmic trading, market quality and information : a dual -process account

Gamzo, Rafael Alon

2017

https://hdl.handle.net/10539/26313

Gamzo, Rafael Alon, (2017) Algorithmic trading, market quality and information: a dual-process account, University of the Witwatersrand, Johannesburg, https://hdl.handle.net/10539/26313

Downloaded from a DSpace repository. https://dspace.org/

WITS BUSINESS SCHOOL

RAFAEL ALON GAMZO

ALGORITHMIC TRADING, MARKET QUALITY AND INFORMATION: A DUAL-PROCESS ACCOUNT

Student Number: 323979

SUPERVISOR: PROFESSOR ERIC SCHALING

PhD THESIS

UNIVERSITY OF THE WITWATERSRAND

Faculty of Commerce, Law and Management

Wits Business School

PhD Thesis

RAFAEL ALON GAMZO

ALGORITHMIC TRADING, MARKET QUALITY AND INFORMATION: A DUAL-PROCESS ACCOUNT

FEBRUARY 2017

© Wits Business School, 2017. All rights reserved. No part of this publication may be reproduced without the written permission of the copyright holder.

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Algorithmic trading, market quality and information: a dual-process account

Abstract

One of the primary challenges encountered when conducting theoretical research on the subject of algorithmic trading is the wide array of strategies employed by practitioners. Current theoretical models treat algorithmic traders as a homogenous trader group, resulting in a gap between theoretical discourse and empirical evidence on algorithmic trading practices. In order to address this, the current study introduces an organisational framework from which to conceptualise and synthesise the vast amount of algorithmic trading strategies. More precisely, using the principles of contemporary cognitive science, it is argued that the dual process paradigm - the most prevalent contemporary interpretation of the nature and function of human decision making - lends itself well to a novel *taxonomy* of algorithmic trading.

This taxonomy serves primarily as a heuristic to inform a theoretical market microstructure model of algorithmic trading. Accordingly, this thesis presents the first unified, all-inclusive theoretical model of algorithmic trading; the overall aim of which is to determine the evolving nature of financial market quality as a consequence of this practice. In accordance with the literature on both cognitive science and algorithmic trading, this thesis espouses that there exists two distinct types of algorithmic trader; one (System 1) having fast processing characteristics, and the other (System 2) having slower, more analytic or reflective processing characteristics.

Concomitantly, the current microstructure literature suggests that a trader can be superiorly informed as a result of either (1) their superior speed in accessing or exploiting information, or (2) their superior ability to more accurately forecast future variables. To date, microstructure models focus on either one aspect but not both. This common modelling assumption is also evident in theoretical models of algorithmic trading. Theoretical papers on the topic have coalesced around the idea that algorithmic traders possess a comparative advantage relative to their human counterparts. However, the literature is yet to reach consensus as to what this advantage entails, nor its subsequent effects on financial market quality. Notably, the key assumptions underlying the dual-process taxonomy of algorithmic trading suggest that two distinct informational advantages underlie algorithmic trading. The possibility then follows that System 1 algorithmic traders possess an inherent speed advantage and System 2 algorithmic traders, an inherent accuracy advantage. Inevitably, the various strategies associated with algorithmic trading correspond to their own respective system, and by implication, informational advantage. A model that incorporates both types of informational advantage is a challenging problem in the context of a microstructure model of trade. Models typically eschew this issue entirely by restricting themselves to the analysis of one type of information variable in isolation. This is done solely for the sake of tractability and simplicity (models can in theory include both variables). Thus, including both types of private information within a single microstructure model serves to enhance the novel contribution of this work.

To prepare for the final theoretical model of this thesis, the present study will first conjecture and verify a benchmark model with only one type/system of algorithmic trader. More formally,

a System 2 algorithmic trader will be introduced into Kyle's (1985) static Bayesian Nash Equilibrium (BNE) model. The behavioral and informational characteristics of this agent emanate from the key assumptions reflected in the taxonomy. The final dual-process microstructure model, presented in the concluding chapter of this thesis, extends the benchmark model (which builds on Kyle (1985)) by introducing the System 1 algorithmic trader; thereby, incorporating both algorithmic trader systems.

As said above: the benchmark model nests the Kyle (1985) model. In a limiting case of the benchmark model, where the System 2 algorithmic trader does not have access to this particular form of private information, the equilibrium reduces to the equilibrium of the static model of Kyle (1985). Likewise, in the final model, when the System 1 algorithmic trader's information is negligible, the model collapses to the benchmark model.

Interestingly, this thesis was able to determine how the strategic interplay between two differentially informed algorithmic traders impact market quality over time. The results indicate that a disparity exists between each distinctive algorithmic trading system and its relative impact on financial market quality. The unique findings of this thesis are addressed in the concluding chapter. Empirical implications of the final model will also be discussed.

DECLARATION

I, Rafael Gamzo declare that the research work reported in this thesis is my own, except where otherwise indicated and acknowledged. It is submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. This thesis has not, either in whole or in part, been submitted for a degree or diploma at any other university. Error and omission noted in this work are attributed to my imperfection.

SIGNATURE: 

ACKNOWLEDGEMENTS

This section is about acknowledging those who have enabled me to complete this work. I am steadfast in the knowledge that I owe it all to You, Hashem, my G-d. You are the ultimate judge and I hope that I have done right in Your eyes. I know that I am blessed because of You. Baruch Hashem.

To my family, words cannot express how thankful I am for your support. Aba, I dedicate this thesis to you, because without you, I would never have had the chance to write this. It's impossible to mention all the things that you've done, so all I can say is, Thank G-d I have a father like you. I love you, Dad! To my mother Lindsay, you've always had my back and it is your kindness that has ultimately allowed me to be where I am today. From the smallest to the biggest acts of kindness, I just want you to know that I appreciate everything you have done to support me. Liat, my sister, you are the ultimate role model and friend and you've believed in me every step of the way, even when I didn't believe in myself. Thank You.

I would also like to thank my supervisor, Professor Schaling. You're an academic giant in my eyes, and none of this would have been possible without you. I have gained so much from your knowledge, insight and unwavering support. I feel truly blessed to have had such a great teacher, who, more than a supervisor, is a real mentor to me. It has been an honour to work under you on this thesis.

Finally, I would like to thank The Wits Business School, and its staff, who have facilitated and assisted me along this journey. I hope that this thesis will reflect the great academic values that have been instilled in me.

Table of Contents

ABSTRACT iii
DECLARATION v
ACKNOWLEDGEMENTS vi
LIST OF TABLES xi
LIST OF FIGURES xii
CHAPTER 1 1
INTRODUCTION 1
1.1 FOREWORD 1
1.2 BACKGROUND OF THE STUDY 2
1.2.1. Algorithmic Trading 2
1.2.2 Judging the Impact of Algorithmic Trading 3
1.2.3. Theoretical Framework 8
1.2.4. Conceptual Framework and Theoretical Implications of the Taxonomy 15
1.3 PROBLEM STATEMENT 16
1.4 OBJECTIVE OF THE STUDY 19
1.5 RESEARCH QUESTIONS 20
1.6 RELEVANCE OF THE STUDY 21
1.7 LIMITATIONS OF THE STUDY 23
CHAPTER 2 25
LITERATURE REVIEW 25
2.1 INTRODUCTION 25
2.2. THE FOUNDATIONS OF COGNITIVE SCIENCE 25
2.2.1 Overview of Generic Dual Process Theories 29
2.3 THE CHANGING TRADING ENVIRONMENT 37
2.3.1 The Origins of Algorithmic Trading 38
2.4 AN OVERVIEW OF ALGORITHMIC TRADING STRATEGIES 40
2.5 ALGORITHMIC TRADING: EXTANT LITERATURE 45
2.5.1 Algorithmic Trading: Empirical Approach 46
2.5.2 Empirical Proxies for Algorithmic Trading 47
2.5.3 Empirical Evidence: Market Quality and Algorithmic Trading 48
2.5.4 Empirical Studies: Concluding Remarks. 58

2.6 THEORETICAL LITERATURE: ALGORITHMIC TRADING 59
2.7 LITERATURE REVIEW SUMMARY 65
CHAPTER 3 68
RESEARCH METHODOLOGY 68
3.1 INTRODUCTION 68
3.2 RESEARCH PARADIGM AND DESIGN 68
3.3 KEY ASSUMPTIONS 70
3.4 ALGORITHMIC TRADING: TAXONOMY DEVELOPMENT 71
3.4.1 The Functional Aspects of Algorithmic Trading: Technological Infrastructure 73
3.5 INSTANTIATION OF SYSTEM 1 AND SYSTEM 2 ALGORITHMIC TRADING 74
3.5.1 The Instantiation of System 1 Algorithmic Trader's: The Speed Advantage 74
3.5.2 The Instantiation of System 2 Algorithmic Trader's: The Accuracy Advantage 76
CHAPTER 4 79
MICROSTRUCTURE BACKGROUND 79
4.1 INTRODUCTION 79
4.2 HALLMARKS OF THE MICROSTRUCTURE APPROACH 80
4.2.1 Information Economics 81
4.3 DEFINING PRIVATE INFORMATION 82
4.3.1 Summary of Private Information 91
4.4 MARKET MICROSTRUCTURE MODELS: A FOCUSED OVERVIEW 93
4.5 THE RATIONAL EXPECTATIONS MODEL 94
4.5.1 The Model 95
4.5.2 Solving for Equilibrium 97
4.5.3 Expectations 98
4.5.4 Demand 101
4.5.5 Market-Clearing Price 102
4.5.6 Limitations of the Grossman-Stiglitz Model 104
4.6 THE KYLE AUCTION MODEL 105
4.6.1 Model Overview 108
4.6.2 The Model 113
4.7 CHAPTER SUMMARY 131
4.8 CONCLUSION 134

4.9 SOME ADDITIONAL THOUGHTS 137
4.9.1 Algorithmic Traders as Informed Traders: An Affirmation 137
CHAPTER 5 142
A THEORETICAL DUAL-PROCESS MODEL OF ALGORITHMIC TRADING: AN INTRODUTION 142
5.1 OVERVIEW 142
5.2 THE FINAL INTERGRATIVE DUAL-PROCESS *TAXONOMY* OF ALGORITHMIC TRADING 143
5.3 MODEL SPECIFICS 146
5.3.1 The Foundations of the Model 147
5.3.2 The Benchmark Concept 148
5.4 THE BENCHMARK MODEL: AN INTRODUCTION 149
5.4.1 Benchmark Model Setup 151
5.4.2 The Benchmark Model – Some Caveats 154
5.4.3 The Extended Processing Phase $(\boldsymbol{T}-)$ 161
5.4.4 The Trading Period $(\boldsymbol{T})$ 163
5.5 FORMAL BENCHMARK MODEL: PROOFS 176
5.5.1 Formal Equilibrium of the Benchmark Model 178
5.5.2 Computation of Equilibrium 179
5.6 EQUILIBRIUM ANALYSIS 183
5.6.1 The System 2 Algorithmic Trader 184
5.6.2 The Market Maker 185
5.6.3 Properties of the Solution 189
5.6.4 Interpretation of Equilibrium 192
5.6.5 Information Revelation 194
5.7 AN ILLUSTRATIVE EXAMPLE OF THE BENCHMARK MODEL 195
5.7.1 Equilibrium Illustration 197
5.8 CONCLUDING REMARKS 201
CHAPTER 6 202
ALGORITHMIC TRADING, MARKET QUALITY AND INFORMATION: A THEORETICAL DUAL-PROCESS MODEL 202
6.1 FOREWORD 202
6.2 THE FINAL MODEL: INTRODUCTION 203
6.2.2 System 1 Algorithmic Trading 213

6.3 SYSTEM 1 ALGORITHMIC TRADER: KEY ASSUMPTIONS 217
6.4 THE FINAL MODEL: INSIGHTS 217
6.5 A DUAL PROCESS MICROSTRUCTURE MODEL OF ALGORITHMIC TRADING 224
6.5.1 Model Setup 224
6.6 FORMAL EQUILIBRIUM OF THE FINAL MODEL 230
6.7 EQUILIBRIUM ANALYSIS 234
6.8 DUAL-PROCESS MODEL INSIGHTS 237
6.9 FINDINGS: MARKET QUALITY 240
6.9.1 Price Informativeness 242
6.9.2 Market Liquidity 243
6.9.3 Volatility 245
CHAPTER 7 246
ALGORITHMIC TRADING, MARKET QUALITY AND INFORMATION: A DUAL PROCESS ACCOUNT – SUMMARY AND CONCLUSIONS 246
7.1 FINAL SUMMARY AND CONCLUSIONS 246
7.2 AN OVERVIEW OF RESULTS 247
7.3 EMPIRICAL IMPLICATIONS 249
7.4 POLICY DISCUSSION 250
References 252
APPENDIX I 271
Empirical Support the Dual Process Proposition – Algorithmic Trading Infrastructure 280
A) Co-location Infrastructure 280
B) High-End Capability Computing Infrastructure 283
APPENDIX II 287
Private Information: The Dimensions of Private Information Accuracy 287
a) The Interpretation Mechanism of Private Information Accuracy 287
b) The Inference Mechanism of Private Information Accuracy 291
APPENDIX III 300
APPENDIX IV 304
APPENDIX V 307

LIST OF TABLES

Table 2: Synthesising the Diverse Literature on Algorithmic Trading…….....64

Table 4.1: Organising the Theoretical Literature on Algorithmic Trading and Market Microstructure.....133

Table 6.1: Comparative Overview of the Dual-Process Theory …….....202

Table I: A Comparative Summary of Dual-Process Concepts…….....286

Table V: Kyle, Benchmark and Final Model: A Comparison……...308

LIST OF FIGURES

Figure 4.1: Categories of Private Information ……..82

Figure 4.2: The Mechanisms of Private Information Accuracy ……...85

Figure 4.3: Visual Summary of Private Information in its Entirety ……92

Figure 5.4: Timeline of the Benchmark Model ……..174

Figure 6.1: A Typical Trading Round…….220

Figure 6.3: Timeline of the Final Model.....…...228

Figure 6.4: Equilibrium Strategies……..236

Figure 6.5: Equilibrium Parameters……..242

CHAPTER 1

INTRODUCTION

1.1 FOREWORD

This paper will propose a theoretical model of algorithmic trading that will draw on the dual-process theory of human cognition to inform an organisational framework in which to conceptualise and synthesize the vast amount of algorithmic trading strategies. The overall aim of this study is to determine the impact of algorithmic trading on market quality.[1] Building on previous knowledge in the area, which studies aspects of algorithmic trading strategies in isolation - a theoretical model is put forward that looks at the effects of mechanisms/processes individually and concurrently (that is, their constituent elements and dynamic interactions), within the same framework. It is hoped that this integrative model may shed light on contradictory findings and on previously unknown market variables i.e., the theoretical model may clarify existing empirical puzzles.

[1] Market quality refers to a market's ability to meet its central function of price discovery. Price discovery refers to the process by which a market incorporates new information about an asset's value into the asset's price (Brogaard, 2010; Hendershott & Riordan, 2009; O'Hara, 2014). These terms will be substantiated as we continue.

1.2 BACKGROUND OF THE STUDY

1.2.1. Algorithmic Trading

For hundreds of years, stock markets have operated as physical locations where market participants could meet to exchange their trading interests. These interactions were supported by floor-based mechanisms like the infamous open outcry system, whereby market participants gathered, and competed for transactions using verbal signals and hand gestures to indicate the quantity and price they were willing to buy or sell a specific financial instrument (Melamed, 2009).

The contemporary securities trading landscape is fundamentally different; the forces of technology, speed, and computer-based trading have facilitated an extraordinary evolution in the manner that current markets function (Jain, 2005).

A particularly intriguing aspect of this revolution remains the extent to which people are being removed from the direct decision making process and being replaced by automated trading systems. The emergence of algorithmic trading (AT) represents such a shift. Indeed, with over 78% of all U.S. equity traded 'volume'[2] originating from computer algorithms, algorithmic trading has asserted itself as the dominant force in financial markets (Johnson, Wang & Zhang, 2014).

Inevitably, there are myriad of questions on the topic of algorithmic trading attracting the attention of researchers (O'Hara, 2014). These issues run the gamut from the conceptual – how and what activities to regulate, to the more general – how to measure the activity of algorithmic trading? However, some of the more fundamental questions seem to concern the evolving nature of market quality (defined below) as a consequence of this activity.

[2] Here, algorithmic trading volume is measured as a percentage of total market volume. However, it is important to note that there are widely varying estimates of algorithmic trading volume in the academic literature (O'Hara, 2014).

The dimensions of market quality, as they appear in the literature, include: trading activity, prices, volume, liquidity, volatility and profits (Brogaard, 2010; Hendershott & Riordan, 2009; O'Hara, 2014). Accordingly, the aforementioned market quality metrics are all indicative of the markets primary role of price-discovery.

Concomitantly, price discovery can be defined as the process by which a market incorporates new information about an asset's value into the asset's price - a central function of financial markets.[3] Given that algorithmic trading has profoundly influenced the price-discovery process, it seems that elucidating the effects of algorithmic trading on market quality has become vital to our understanding of financial market performance (Gomber, Arndt, Lutat & Uhle, 2011).[4]

1.2.2 Judging the Impact of Algorithmic Trading

Algorithmic trading is broadly defined as the "use of computer algorithms to automatically make trading decisions, submit orders, and manage those orders after submission" (Hendershott & Riordan, 2009, p. 2).[5]

[3] If we consider that asset prices are determined by the outcome of supply and demand, then in a competitive market, traders are expected to rapidly assimilate any new information that is relevant to the determination of asset prices, and prices should adjusting accordingly. The manner in which markets "discover" this new information is aptly named the price "discovery" process. (Alagidede, 2008). Fama (1970) asserts that the price-discovery process is the central function of a financial market, whereby prices offer accurate signals for resource allocation and the efficient distribution of an economy's capital stock. Contextually, market quality is more general than price discovery.

[4] Financial market performance can be defined as how well (or poorly) a market is performing its vital economic function of price discovery (SEC, 2010). That is, a well-functioning securities market is a market in which firms can make production -investment decisions, and investors can choose among the securities that represent ownership of firms' activities under the assumption that security prices reflect *all available information* (Fama, 1970). Given that the price discovery process typifies the manner by which prices reflect available information, price discovery and financial market performance are inexorably linked.

[5] Defining algorithmic trading is difficult and there is no single agreed definition. Gomber et al, (2001) assert that algorithmic trading is in fact a misnomer, a seemingly precise term used to describe and diverse set of activities and behaviours. Appendix IV lists the relevant academic and regulatory definitions on algorithmic trading, noting the diversity.

Estimates of algorithmic trading typically exceeded 50% of total volume traded in U.S. listed equities (Goldstein, Kumar, & Graves, 2014; Kissell, 2013). A precise assessment of its impact on market quality is, however, challenging. This is partly a reflection of the complexity of defining the term itself. There has not, to date, been a consistent academic or regulatory definition of the term algorithmic trading – it is used in a variety of contexts and for various purposes. Invariably, *Automated Trading, Flash Trading, Program Trading, Low Latency Trading, Black Box Trading, Electronic Trading and High Frequency Trading* are just some of the labels ascribed to it in the literature (Kissell, 2013). An additional complexity in seeking to define algorithmic trading is that it encompasses many players, different infrastructural arrangements and, most importantly, a wide number of diverse strategies (Gomber et al., 2011).

There are two separate approaches to judging the impact of algorithmic trading. The first is empirical in that it draws on *data* from stock exchanges in order to infer impact. The second is theoretical in that it makes use of microstructure theory in order to model the behaviour of an algorithmic trader in a hypothetical, mathematically constructed market (Cvitanic & Kirilenko, 2010; Biais, Foucault & Moinas, 2011; Jovanovic & Menkveld, 2014; Menkveld 2016). Both approaches emphasise market quality as an important standard by which to judge the impact of algorithmic trading. Again, this is because market quality is indicative of the markets central function of price-discovery.[6]

[6] Typical market quality parameters include: trading activity, prices, volume, liquidity, volatility and profits (Brogaard, 2010; Hendershott & Riordan, 2009; O'Hara, 2014).

1.2.2.1. Empirical Approaches

The academic discourse on the topic of algorithmic trading has been largely empirical in nature. Yet this literature comes with well documented limitations (U.S. Securities and Exchange Commission (SEC), 2014). Obtaining useful data that can identify algorithmic trading activity is a formidable challenge. Much of the empirical evidence on the direct impact of algorithmic trading on the U.S. equity markets has relied heavily on either limited samples of *proprietary data* or *publicly available information*. Proprietary datasets are not publicly available, have been limited to particular products, are discretionary in nature and show only a small amount of algorithmic trading activity (Biais & Foucault, 2014).

Alternatively, publicly available data on orders and trades does not reveal the identity of buyers and sellers. As a result, at this time, it is *not possible* to identify orders and trades as originating from an algorithmic trader account when relying solely on publicly available information. Therefore, a variety of different metrics are used by researchers to estimate algorithmic trading. Analysing actual algorithmic trading is based on empirical proxies,[7] and results thus rely heavily on the quality of the proxy (SEC, 2014).

Empirical proxies for algorithmic trading are discretionary in nature and are often specific to a certain researcher i.e., research-specific. No standard proxy exists. As discussed further below, empirical proxies can greatly affect findings regarding the key factual characteristics of algorithmic trading activity. Additionally, the diversity of algorithmic trading strategies highlights the importance of exercising care when using metrics to define algorithmic trading (Conrad,

[7] A proxy can be defined as a figure that can be used to represent the value of something in a calculation.

Wahal, & Xiang, 2015). Ultimately, results in the academic literature are as diverse as the trading strategies themselves. (See the literature review for an in depth analysis)

1.2.2.2. Theoretical Approaches

If there are not any *a priori* ideas about the nature, the inner working, and the theoretical underpinnings of algorithmic trading, empirical studies may be limited. In turn, this may lead to the accumulation of a vast amount of data without any apparent hope of arriving at a succinct, precise, and meaningful understanding (LeBaron, 2006). By contrast, the theoretical literature attempts to provide a framework for understanding the key behavioural characteristics of an algorithmic trader. The theoretical approach embodies an explicit representation of an algorithmic trader in a mathematically generated hypothetical financial market. The foundations of these models are located in the market microstructure and behavioural finance literature (Madhavan, 2000; O'Hara, 1995).

Typically, microstructure models focus on the distributions of certain informational characteristics among agents. Given the agents' information set, the market's equilibrium condition can be determined analytically. Crucially, this information set depends solely on an agents underlying trading *strategy* (Francioni, Hazarika, Reck, & Schwartz, 2008).[8]

As discussed further below, algorithmic trading is not a single strategy phenomenon; rather it encompasses a diverse array of trading strategies each with its own distinct informational character. In terms of strategies, both numerous and diverse trading approaches currently exist. Amongst these are; Spread Capturing, Rebate Trading, Time Weighted Average Price (TWAP), Volume Weighted Average Price (VWAP), Implementation Shortfall, Adaptive Execution,

[8] Potential algorithmic trading strategies are discussed in Chapter 2 (Section 2.4).

Liquidity Detection, Data\Text Mining, Neural Network and Support Vector Machine Strategies. (See Section 2.4 for a review of the relevant strategies).

However, to date, theoretical models of algorithmic trading have been largely strategy specific, simplified, analytically tractable models with a single representative algorithmic trader. Indeed, current theoretical models of algorithmic trading tend to focus on a specific algorithmic trading strategy, often conflating that strategy with algorithmic trading in its entirety (Cvitanic & Kirilenko, 2010). Given that algorithmic trading is a dominant component of current market structure and likely to affect nearly all aspects of its performance, it seems that a model that accounts for the multifaceted nature of algorithmic trading has become vital to our understanding of financial market performance.

This research will put forward a theoretical model of algorithmic trading that will draw on the dual-process theory of human cognition to inform an organisational framework in which to conceptualise and synthesise the vast amount of algorithmic trading strategies. More precisely, using the principles of contemporary cognitive science, it will be argued that the dual-process paradigm – the most prevalent contemporary interpretation of the nature and function of human decision making – lends itself well to a novel *taxonomy* of algorithmic trading.

Thus, this taxonomy serves primarily as a heuristic to inform a theoretical model of algorithmic trading; with the view of explaining the evolving nature of market quality as a consequence of this practice. Arguably, a model that synthesises the dynamic aspects of algorithmic trading within a single overarching framework has the virtue of great verisimilitude, and may enable us to bridge a current research lacuna: that is, contention surrounding the impact of algorithmic trading on market quality. Moreover, this integrative model may shed light on contradictory findings and on previously unknown market variables.

1.2.3. Theoretical Framework

The dual-process paradigm has emerged as the most prevalent contemporary interpretation of the nature and function of human decision making (Sloman, 1996). Invariably, this construct is predicated on an assumption that cognitive decision making can be divided into two distinct processes or systems. More precisely, the dual-process paradigm maintains that all cognitive activity comprises two multi-purpose behavioral systems, for which Stanovich and West (2000, p.658) proposed the labels of System 1 and System 2.[9] One (System 1) having fast-process characteristics (impulsive, automatic, reflexive and fast, etc.), and the other (System 2) having relatively slower, more analytic or reflective processing characteristics (controlled, effortful and reflective etc.) (Evans & Over, 1996; Sloman, 1996; Stanovich, 1999).

As noted, this research presents the dual-process cognitive theory as a possible means by which to conceptualise and synthesise multiple algorithmic trading strategies under a single framework. More precisely, following the theoretical and empirical literature on both cognitive science and algorithmic trading, this research espouses that there exists two distinct types of algorithmic trading; one (System 1) having fast process characteristics, and the other (System 2) having relatively slower, more analytic or reflective processing characteristics. The intuition for this assertion will be developed in proceeding chapters.

As noted previously, this taxonomy serves to guide the construction of a single theoretical model; whereby the key characteristics of these traders reflected in the taxonomy will translate into key

[9] It should be noted that small terminological discrepancies do exist between the different theories. However, a critical evaluation of the various terms identified in the literature is beyond the scope of this thesis. Instead, we will emphasize the most neutral terms available, namely; the distinction between System 1 and System 2 processes (Kahneman & Frederick 2002; Stanovich 1999; Stanovich & West, 2000).

behavioural and informational assumptions of the agents in this model. Moreover, the study will attempt to support the above mentioned taxonomy with recent empirical evidence on algorithmic trading activities. Indeed, it will be asserted that a 'functional division'[10] of algorithmic trading strategies – predicated on infrastructural characteristics of algorithmic trading – provides robust empirical support for this unique taxonomy (See Appendix I).

1.2.3.1. Rationale for Focusing on the Dual Process Theory

The dual process theory was drawn on for two overarching reasons, namely a) The typical cognitive process in both human beings and digital computers rely on the same fundamental principles - the proof of this concept is provided below - and b) A functional division of actual algorithmic trading strategies by their requisite infrastructure reveals real non-trivial parallels between the dual process account of cognition and algorithmic trading.

a) Computation and Cognition

Naturally, one might have expected that in an environment dominated by computers, human cognition becomes irrelevant – the opposite is actually the case. In fact, a recent proliferation of empirical evidence suggests that the typical cognitive process in both human beings and computers rely on the same fundamental principles (Pylyshyn, 1980, 1984).[11]

Digital computers showed cognitive scientists that it was possible to explain the intelligent behaviour of a complex system without presupposing the intelligent behaviour of its components by employing the idea of computation - a set of rules to be followed in calculations or other

[10] By functional division we mean a difference in the functional *architecture* of algorithmic trading and not merely a difference in how it functions.

[11] Accordingly, computers and human organisms are both physical systems whose behaviour is correctly described as being governed by rules acting on symbolic representations.

problem-solving operations (Fodor, 1975, 1981, 1987, 1990, 1995; Pylyshyn, 1980, 1984). Computation has since emerged as a hypothesis about the literal nature of cognition; as opposed to a mere metaphor for its operation. Both the mind and the digital computer receive information from the environment, process it, and use it to perform an action.

Computation and cognition have become almost inseparable, and often indistinguishable, in much literature on the mind and brain (Pylyshyn, 1980). It can be said that humans and computers are just two species in the genus of information processing systems. *Overall, both cognition and computation operate on symbolic representations that can be physically instantiated in the form of internal rules and manipulated by some kind of processing element.*

In order to understand the statement above, it is important to understand the meaning behind symbols. On the one hand symbols are physical – like the ink on this page, the electrical impulses in a human brain or the magnetic records in a computer. On the other hand, symbols are representative of something else - for instance, the symbol $ is used to stand for money. That is not to say that the symbol $ is actual currency, rather, the symbol is a surrogate that refers to its referent, which is actual currency (Friedenberg & Silverman, 2011). In the case of representation, we say there is some symbolic entity 'in the head' or 'in the computer' that stands for real currency. In this view, both computers and minds are formal symbol manipulators.

A system is formal if it is rule governed. The rules of language and mathematics are formal systems since they stipulate which types of allowable changes can be made to symbols. Manipulations on the other hand, are actions which occur physically in some type of device or processing construct, e.g., a computer or the brain (Friedenberg & Silverman, 2011; Stillings, Weisler, Chase, Feinstein, Garfield & Rissland, 1995). Concurrently, this reinforces the aforementioned statement, namely,

that both cognition and computation operate on symbolic representations that can be physically instantiated in the form of internal rules and manipulated by some kind of processing element.

These nontrivial commonalities (between cognition and computation) imply that the mechanisms of human decision making are analogous with the decision making processes of an automatic economic decision making technology such as algorithmic trading – a fundamental doctrine of this thesis.

b) Infrastructure

The ideas behind this thesis are shaped by many dialogues in cognitive science regarding the architecture of the mind, particularly the computational processes underlying decision making. The present research is concerned with the dual emergent properties of types of cognitive (sub) systems, dubbed 'System l' (fast, superficial, reflexive, etc.) and 'System 2' (controlled, accurate, reflective, etc.).

This research considers the principle that deep functional differences (i.e., difference in the functional *architecture* of the system and not merely a difference in how it functions) exist between the two systems, and that this distinction can be expressed physically/anatomically. Lieberman (2007; 2009), provides the rationale for this position. Using the tools provided by neurological sciences, Lieberman (2007) infers, from the activation of a specific brain region, the phenomenological and representational characteristics of System 1 and System 2 information processing. Although of secondary importance to the information processing distinction emphasised in the cognitive literature, Lieberman (2007; 2009) highlights that the structural aspects of the processes' have clear descriptive relevance.

Such structural considerations are also relevant to our dual process decomposition of algorithmic trading. As mentioned previously, the dual process position on *human cognition* effectively synthesises the diverse computational process in the mind into functional sub-systems; with alternate brain regions responsible for each system. We argue that the diverse algorithmic trading computational processes are subject to a similar architectural distinction or functional distribution.

In light of the above, the current study will evaluate existing empirical research on the infrastructural aspects of actual algorithmic trading practices in order to determine the extent to which empirical research supports our dual process supposition of algorithmic trading (See Appendix I).

As evidenced in Appendix I, one can distinguish between two independent forms of infrastructure when it comes to algorithmic trading: that is, between co-location infrastructure and high-end capability computing infrastructure. In fact, empirical evidence seems to suggest that although ostensibly distinct, the array of different strategies associated with algorithmic trading require either co-location or high-end capability infrastructure to be performed (See, for example, Johnson, 2010; Gomber et al., 2011; Kissell, 2013; Frino, Mollica & Webb, 2014 and O'Hara, 2014).

These two disparate forms of infrastructure serve two distinct functions. Co-location allows firms to locate their 'servers'[12] next to the exchanges'. Placing ones server adjacent to the exchanges matching engine means that real-time market information can reach the algorithmic traders platform instantaneously. It therefore significantly reduces the time it takes to access the central

[12]. Strictly speaking, a server is a computational device which is dedicated to the running of a certain program. With regards to algorithmic trading, a server functions as an electronic communication device or electronic infrastructure that manages access to a centralized resource.

order book (where electronic information on quotes and prices are warehoused). It also decreases the time it takes to transmit trade instructions and execute matched trades. In exchange for a fee, those who subscribe to co-location services receive the infrastructure from the exchange itself. The package includes everything from the actual connection to the matching engine, to server cages, electricity, maintenance, and safety installations. The SEC (2010) notes that due to co-location facilities: "the speed of trading has increased to the point that the fastest traders now measure their latencies in microseconds"[13] (p.3605). In the past, those on the trading floor had faster access to the market than others; today, those co-located with the exchange market have faster access.

Co-location has emerged as a necessity and not merely a luxury or curiosity for those traders wishing to be fast. Given that latency or speed is the *raison d'être* for this type of infrastructure (co-location infrastructure) it seems reasonable to suggest that *speed* typifies those strategies that are conducted via co-location infrastructure.

High-end capability computing infrastructure on the other hand, serves as a lever to pry new insight from a mass of data or complicated mathematical formula. As such, capability computing is measured in terms of its complex, analytical, logical and deductive reasoning capabilities. Overall, those strategies that are performed on high-capability computer infrastructure exploit their superior ability to interpret public information in an attempt to make forecasts that are superior to the forecasts of other traders. In other words, these traders filter public information through an

[13] Where 1 microsecond is an IS (International System) unit of time equal to 1000000th of a second.

advanced platform (high-end capability computer infrastructure), in order to detect 'private'[14] patterns from public information – patterns that signal a firms future performance. Therefore, it seems natural to assume that several algorithmic trading strategies (those not reliant on colocation infrastructure) are simply predicated on an ability to make forecasts/analyses that are superior – in terms of *accuracy* – to those of other traders.

Overall, it is becoming increasingly clear that speed typifies one portion of algorithmic trading strategies – those reliant on co-location infrastructure – whilst accuracy epitomises the other portion of algorithmic trading strategies – those strategies reliant on high-capability infrastructure. This speed/accuracy distribution appears analogous with the dual system distinction of cognitive processing discussed above.[15]

In addition to providing robust empirical support for our dual system distinction of algorithmic trading, an infrastructural decomposition of algorithmic trading may also have descriptive relevance. As such, the structural decomposition of algorithmic trading will be used as a descriptive heuristic device that *contributes* to our taxonomy and modelling methodology.

Contextually, the dual process theory of human cognition will be used as the intellectual foundation for the unique theory of this thesis; this dual process theoretic precept, however, will be supported (and perhaps clarified) by actual phenomena discussed in the proceeding chapters.

---

[14] In financial market research, informed investors are seen to be those investors that have "private" information about the future states of the world (Grossman, 1976). This seems to imply that some investors are better than others when it comes to interpreting financial information and are, as a consequence, better at forecasting future market movements. Unlike public information which can be observed directly by all from the firms' accounting statements, private information relates to the more *abstruse* information about growth opportunities judged or interpreted privately by owners. This information reflects the more "decrypted" information - signalling the firm's future performance (see Chapter 4, Section 4.3, for a formal definition of private information).

[15] Thus, a functional division of algorithmic trading strategies by their requisite infrastructure reveals actual non-trivial parallels between the dual process account of cognition and algorithmic trading (this remains one of the overarching reasons why dual process theories were drawn on as a heuristic for our dual process taxonomy of algorithmic trading, see Section 1.2.3.1 – above).

In the following subsections we expand on the theoretical implications and conceptual foundations of our novel dual process taxonomy of algorithmic trading. See Appendix I for further information on the infrastructural characteristics of algorithmic trading.

1.2.4. Conceptual Framework and Theoretical Implications of the Taxonomy

Following key insights from cognitive science, this thesis advocates a taxonomy that classifies algorithmic trading with respect to the manner in which it processes information. The distribution and processing of information is a central consideration in market microstructure modelling, and thus appears to be a natural place to start in constructing the model. Two types of information processing are distinguished in the current set-up: (1) a type of processing that is fast, reflexive and superficial (2) a type of processing that is controlled, precise and reflective. From a theoretical microstructural modelling perspective, this supposition implies that two informational advantages underlie algorithmic trading.[16] That is, one portion of algorithmic traders possess' an inherent speed advantage, and another portion, an inherent accuracy advantage (relative to both each other and the rest of the market). This construct is largely consistent with the theoretical literature on algorithmic trading. Consider the following:

The view that algorithmic traders possess a comparative informational advantage relative to 'regular'[17] traders has emerged as somewhat a theoretical regularity in the algorithmic trading literature. However, evidence is anecdotal and suggests *either* a speed advantage or an accuracy advantage in isolation (Biais, Foucault & Moinis, 2011; Cartea & Penalva, 2010; Cvitanic &

[16] Arguably, some existing models are special cases of this general framework. Therefore, Table 4.1 in Chapter 4, provides some clarity on where exactly existing models fit in light of the abovementioned dual process regime. Concurrently, in Sections 2.5 and 4.9, it is suggested that this taxonomy is exclusive to algorithmic traders.

[17] The term 'regular' alludes to traditional (non-algorithmic) financial market participants vis-a-vis noise traders' market makers and any other human traders.

Kirilenko, 2010; Das, Hanson, Kaphart & Tesauro, 2001; Easley, Lopez De-Prado & O'Hara, 2012; Foucault, Hombert & Rosu, 2016; Gamzo, 2014; Johnson, 2010; Martinez & Rosu, 2011).

By advocating that dual advantages underlie algorithmic trading, our dual process supposition reconciles existing and competing notions of algorithmic trading. An infrastructural decomposition of algorithmic trading with respect to asymmetric/private information seems only to reinforce this assertion.

Moreover, given that algorithmic traders are almost certainly the fastest and the most accurately informed participants in the market, it seems reasonable to suggest that our taxonomy is exclusive to algorithmic traders (This statement is clearly justified in Chapter 4, Section 4.9). The prominence of this practice – accounting for a much larger proportion of trades relative to traditional human traders – underscores the relevance of our supposition.

1.3 PROBLEM STATEMENT

According to Jain (2005), 'algorithmic trading'[18] represents one of the most interesting and significant developments in the history of financial markets. With over half of all equity trading volume originating from automated computer algorithms, algorithmic trading has fundamentally altered the way stock markets function (Gomber et al., 2011). Understanding this practice has become vital to our understanding of financial markets.

The academic discourse on the topic of algorithmic trading has been largely empirical in nature. However, this literature comes with well documented limitations (Biais & Foucault, 2014).

[18] Appendix IV lists the variety of academic and regulatory definitions on algorithmic trading. For now, a broad definition of algorithmic trading should suffice i.e.: "Algorithmic trading refers to the use of computer algorithms to automatically make trading decisions, submit orders, and manage those orders after submission" (Hendershott & Riordan, 2009, p. 2). A more elaborate definition can also be found in Section 2.3.1 of the literature review.

Obtaining useful data that can identify algorithmic trading activity is a formidable challenge. At this time, it is *not possible* to identify orders and trades as originating from an algorithmic trader account when relying solely on publicly available information. Consequently, researchers have devised different methods to '*proxy'*[19] for algorithmic trading activities within traditionally-available datasets. Research-specific proxies for algorithmic trading can greatly affect findings about the key characteristics of algorithmic trading activity and its impact on financial market quality (Biais & Foucault, 2014).[20] In perhaps the greatest indictment of existing empirical studies to date, the SEC (2014) warns that interpreting the results of empirical studies should be done with caution; "an assessment of empirical papers must deal with the various metrics researchers used to define algorithmic trading, and how their definitions may affect their conclusions about algorithmic trading activity… Particularly, the different metrics (proxies) used to classify algorithmic trading can greatly affect findings about key characteristics of algorithmic trading and its impact on financial market quality" (p.5).[21] Ultimately, the lack of precise data is a sufficient enough reason to believe that the goal of understanding algorithmic trading strictly from empirical observations is unattainable.

[19] A proxy is simply a figure used to represent the value of something in a calculation. With regards to these proxies, examples are here barred. The introductory nature of this chapter (together with the theoretical focus of this thesis) precludes any detailed exposition of these proxies here. Readers who would prefer clarification at this stage are directed to Chapter 2, Section 2.5.2, where we delineate the variety of different proxies underlying empirical research on the topic of algorithmic trading, noting the diversity of conclusions reached.

[20] Proxies are discretionary in nature (highly subjective) and are often specific to a certain researcher i.e., proxies are research specific. Refer to the literature review in Chapter 2 for further insight.

[21] Other examples of such AT proxies derived from market-wide data include high message rates, bursts of order cancellations and modifications, high order-to-trade ratios, small trade sizes, and increases in trading speed. Note that a critical evaluation of the various proxies identified in the literature is consigned to Chapter 2. We emphasise here, only, that the diversity proxies for algorithmic trading can greatly affect findings about the key factual characteristics of algorithmic trading activity.

The theoretical literature offers little respite in the way of an accurate assessment of algorithmic trading. One of the primary challenges encountered when conducting theoretical research on the subject of algorithmic trading is the wide array of strategies employed by practitioners (Gomber et al., 2011). Disconcertingly, current theoretical literature (as exemplified by the work of e.g., Biais et al., 2014; Cartea & Penalva, 2010; Cvitanic & Kirilenko, 2010; Das et al., 2001; Easley et al., 2012; Foucault et al., 2016; Gamzo, 2014; Johnson, 2010; Martinez & Rosu, 2011) treat algorithmic traders as a homogenous trader group, forming a gap between academic discussions. Existing theoretical models on the topic (i.e., Cvitanic & Kirilenko, 2010; Foucault, Hombert & Rosu, 2016; Martinez and Rosu, 2011) have been largely strategy specific, simplified, analytically tractable models with a single representative algorithmic trader. Indeed, current theoretical models of algorithmic trading tend to focus on a *specific* algorithmic trading strategy, and often conflate that strategy with algorithmic trading in its entirety (Cvitanic & Kirilenko, 2010). Given that algorithmic trading is a dominant component of current market structure and likely to affect nearly all aspects of its performance, it seems that a model that accounts for the 'multifaceted'[22] nature of algorithmic trading strategies has become vital to our understanding of financial market performance (Securities and Exchange Commission (SEC), 2014).

In order to address this, the current study introduces an organisational framework from which to conceptualise and synthesise the vast amount of algorithmic trading strategies. More precisely, using the principles of contemporary cognitive science, it is argued that the dual process paradigm – the most prevalent contemporary interpretation of the nature and function of human decision making – lends itself well to a novel *taxonomy* of algorithmic trading.

[22] Algorithmic trading is multidimensional concept, encompassing a diverse array of trading strategies. Inevitably, these strategies differ with regards to their market impact and informational characteristics.

This taxonomy serves primarily as a heuristic to inform a theoretical market microstructure model of algorithmic trading. Therefore, this study proposes a model that integrates and synthesises the multitude of algorithmic trading strategies within a single modelling framework. It is hoped that this consolidated model will reconcile competing and existing notions – both theoretical and empirical – on the practice of algorithmic trading and its impact on market quality. This integrative theoretical model may also shed light on existing empirical puzzles.[23]

1.4 OBJECTIVE OF THE STUDY

The central aim of this thesis can be described concisely, namely; designing and executing a theoretical enquiry into algorithmic trading with the view of explaining the evolving nature of market quality as a consequence of this activity. To elaborate on the above, some reflection on an important tenet of this thesis is necessary. Specifically, the overarching *objective* of this thesis is to present the first unified, 'all-inclusive' theoretical model of algorithmic trading (all-inclusive in that it accounts for the multitude of algorithmic trading strategies within a single theoretical framework).

In accordance with the literature on both cognitive science and algorithmic trading, this thesis espouses that there exists two distinct types of algorithmic trader; one (System 1) having fast processing characteristics, and the other (System 2) having slower, more analytic or reflective

[23] The model has microstructural foundations. Thus, in order to understand the extent to which the model integrates existing models, an in-depth knowledge of market microstructure theory is required. Given that research within market microstructure is so extensive, some form of injustice in the form of neglecting otherwise important microstructure-theoretic issues would arise with further substantiation here. Therefore a precise assessment of the nexus between existing models and those of this thesis can only be accomplished following a detailed review of market microstructure theory – assigned to Chapter 4. Following the overview of theoretical market microstructure in Chapter 4, a comparative summary of the literature is presented in Table 4.1. Table 4.1 allows the reader to gauge - visually - how our model integrates existing theoretical models. It also highlights which existing models are special cases of our more general framework.

processing characteristics. Thus, from a theoretical level, this work explicates the interaction between two distinct systems of algorithmic trading, in contrast to the tendency of studying a single type in isolation. A theoretical model is put forward that looks at the effects of these mechanisms/processes individually and concurrently (that is, their constituent elements and dynamic interactions), within the same framework. The main motivation for developing this theoretical framework is that it may clarify the existence of a disparity between each distinctive algorithmic trading system and its relative impact on financial market quality.

Moreover, it is hoped that this integrative model may shed light on contradictory findings and on previously unknown market variables i.e., the theoretical model may clarify existing empirical puzzles.

1.5 RESEARCH QUESTIONS

Algorithmic trading has become a key aspect of modern financial systems and has been gaining market share for the past decade or so. Despite this, the academic literature is not yet fully developed and many important questions remain unanswered. Does algorithmic trading enhance or impede financial market quality? Does it lead to an increase in volatility and crowding out of human traders? Or, does the speed of computers help incorporate information quicker, thereby aiding in the price discovery process? Do algorithmic traders make markets more liquid, such that it is easier and less costly to trade, or do they in fact take liquidity out of the market?

Absent draconian regulations, algorithmic trading will likely remain an important feature of modern financial markets. Understanding its impact and behaviour is therefore crucial in forming an understanding of modern market functioning.

1.6 RELEVANCE OF THE STUDY

The SEC's (2014) Concept Release on Equity Market Structure recognized that algorithmic trading is one of the most significant equity market structure developments in recent years. It noted, for example, that estimates of algorithmic trading typically exceeded 50% of total volume in U.S. listed equities and concluded that, by any measure, algorithmic trading is a dominant component of the current market structure and likely to affect nearly all aspects of stock market performance.[24] Understanding the impact of algorithmic trading on market quality has been the subject of considerable interest in microstructure analysis.

Price-discovery is a central function of financial markets (Fama, 1970). On a micro level, algorithmic trading can be seen to have profoundly influenced the price-discovery process (Gomber et al., 2011). It follows that elucidating the effects of algorithmic trading on market quality – the typical standard by which to account for the price discovery process – is vital to our understanding of financial market performance. On a macro level, there are wider implications for the economy as a whole. The link between a well-functioning securities market and economic growth is well established in the literature (see e.g., Bencivenga & Smith, 1991; Obstfeld, 1992; Saint-Paul, 1992; Boyd & Smith, 1996; Rajan & Zingales, 1996). For these reasons, conducting microstructure analysis on the impact of algorithmic trading on financial market quality has important applications with respect to market regulation, market design and is of great importance to academics, exchanges and regulators.

The overall aim of this study is to determine the impact of algorithmic trading on market quality in a theoretical framework. As will be clarified in the proceeding literature review below, current

[24] Financial market performance can be defined as how well (or poorly) a market is performing its vital economic functions, including, price discovery (SEC, 2010). Readers are referred to footnote 4 for substantiation.

models of algorithmic trading are flawed because they fail to take into account the multidimensional aspects of its underlying strategies. For now, consider the following narrative that highlights, broadly, the limits of existing theoretical research:

Algorithmic trading is a multidimensional concept, encompassing a diverse array of trading strategies. Inevitably, these strategies differ with regards to their market impact and informational characteristics. The diverse nature of this practice means that research on the subject is often limited in scope. Existing theoretical models tend to focus on a specific algorithmic trading strategy and often conflate that specific strategy with algorithmic trading in its entirety. In fact, the SEC (2014) notes the many instances in which researchers attempt to determine the impact of algorithmic trading as a whole, but follow almost immediately by an emphasis on a specific isolated strategy – without an explicit indication or explanation for such a strategy specific approach. Therefore, an alternative method of inquiry that captures the diversity of algorithmic trading practices in a systematic and orderly fashion seems pertinent.

Accordingly, this thesis will present, what is arguably the first 'all-inclusive'[25] model of algorithmic trading; one that draws on the dual-process theory of human cognition in order to inform an organisational framework in which to conceptualise and synthesise the vast amount of algorithmic trading strategies in a systematic way. A model more representative of algorithmic trading in its entirety has the virtue of great verisimilitude, and may enable us to begin bridging a current research lacuna – i.e., questions' surrounding algorithmic trading's impact on financial market quality.[26]

[25] All-inclusive in that it accounts for the multitude of algorithmic trading strategies within a single theoretical framework.

[26] The model will allow us to explore, explicitly, financial market quality metrics (i.e. trading activity, prices, volume, liquidity, volatility and profits) as these are theoretical concepts in the model.

As reported in Goldstein et al., (2012) and Johnson (2010), algorithmic trading is also rapidly gaining popularity in Europe and Asia, accounting for approximately 45% of stock trading volume in the European Union, 40% in Japan, and 12% in the rest of Asia as of late 2012. In South Africa, algorithmic trading is estimated to account for as much as 76% of equity market volume on the Johannesburg Stock Exchange (JSE). Algorithmic trading is thus a worldwide phenomenon and as such, the results of this study may have broad implications, not just on U.S. equity markets, but also stock markets around the world.

1.7 LIMITATIONS OF THE STUDY

This thesis develops a theoretical model of algorithmic trading. While some empirical applications are discussed in the concluding chapter of this thesis, the contribution of this thesis remains purely theoretical. Arguably, the current study would have been enhanced had we been able to empirically verify the model. However, our exploration of the empirical applications of the developed model has been severely restricted. This is largely on account of the fact that the main immediate prerequisite for empirically examining algorithmic trading relies, primarily, on the access to high quality data.

With regards to algorithmic trading, currently available data sets make it difficult to carry out any methodologically consistent empirical analysis. Indeed, obtaining useful data that can identify algorithmic trading activity is a formidable challenge. Much of the empirical evidence on the direct impact of algorithmic trading has relied heavily on either limited samples of proprietary data, or publically available information. Furthermore, the strategies employed by algorithmic trading firms are marred by secrecy. This secrecy, coupled with the competitive nature of the industry, means that rich proprietary data is scarcely available.

Alternatively, analysing actual algorithmic trading activities using publically available data requires the use of empirical proxies. Unfortunately, these proxies are discretionary; and so may render the empirical literature meaningless for deriving useful conclusions.

In summation, the above mentioned data challenges have severely limited our ability to empirically verify the model, which could have brought further insights into the dynamic and evolving nexus between algorithmic trading and financial market quality. Nonetheless, we argue that the work here represents an important step in modelling algorithmic trading behaviour and opens up the door for empirical research as more data becomes available.

CHAPTER 2

LITERATURE REVIEW

CONTEMPORARY COGNITIVE SCIENCE, DUAL-PROCESS THEORIES AND ALGORITHMIC TRADING: EXTANT EVIDENCE

2.1 INTRODUCTION

The present chapter will begin by examining the available literature on contemporary cognitive science in order to provide a premise from which to discuss and evaluate the critical links between the modern cognitive approaches (dual process construct) and algorithmic trading.

2.2. THE FOUNDATIONS OF COGNITIVE SCIENCE

One of the most interesting intellectual developments of the past few decades has been the emergence of a novel interdisciplinary field named cognitive science. Researchers in psychology, philosophy, neurological science and computer science soon realised that they were attending to many of the same issues and had developed complementary and potentially synergistic methods of investigation (Boden, 2008).

Over the last six decades or so, cognitive science has overcome disciplinary boundaries, leading to the revolutionary reorientation in the science of the mind. Cognitive scientists seek to understand the mechanisms underlying important mental phenomena such as reasoning, learning and decision making (Boden, 2008). Their research is remarkably diverse. It includes, for example, analysing the nature of thought, studying the principals of neural circuitry in the brain, and programing computers to do complex problem solving.

Like all intellectual disciplines, cognitive science takes a definite perspective. Cognitive scientists view the human mind as a complex system that receives stores, retrieves, transforms, and transmits information (Stillings et al., 1995). The literature refers to these operations as *computations* (e.g., Pylyshyn 1980; Stillings et al., 1995).

This computational reference is particularly insightful and rests on certain academic doctrines surrounding the fundamental similarity between digital computation and human cognition. Human beings and digital computers are seen in the literature as two species in the genus of information processing systems (Stillings et al., 1995). In other words, both the mind and the digital computer receive information from the environment, process it, and use it to perform an action.

Computation and cognition have become almost inseparable, and often indistinguishable, in much literature on the mind and brain. In fact, a recent proliferation of empirical evidence suggests that the typical cognitive process in both human beings and computers rely on the same fundamental principles (Pylyshyn, 1980, 1984).[27]

Fodor (1975, 1981, 1987, 1990, 1995) and Pylyshyn (1980, 1984) advance a strong case for this empirical regularity, suggesting that the term computation exists as a *literal* translation of mental activity. To paraphrase Pylyshyn (1980): "There is no reason why computation ought to be treated as merely a *metaphor* for cognition, as opposed to a hypothesis about the *literal* nature of cognition"(p.114).

[27] Computers and human organisms are both physical systems whose behavior is correctly described as being governed by rules acting on symbolic representations (See Section 1.2.3.1 (a)).

In roughly equivalent terms, both cognition and computation operate on symbolic representations that can be physically instantiated in the form of internal codes and manipulated by some kind of processing element (Pylyshyn, 1984) (See Section 1.2.3.1 (a)).

Originally, this computational philosophy was predicated on the assumption that a single computational process motivates cognition. Yet, more recently, this entirely domain-general account has come under some scrutiny (Beaulac, 2010). Some of these more recent discussions, suggest that a growing number of cognitive scientists are adopting a more pragmatic outlook, one involving hybrid systems with different sorts of architectures for different sorts of sub-processes or tasks. For example, Stillings et al., (1995) posit that an intelligent system ought to "consist of a number of functional subsystems, or modules, that cooperate to achieve intelligent information processing and behavior" (p.16). Indeed, close inspection of recent research trends discloses a set of recurrent illustrations of *computational theories* in their conceptual, observational and instrumental applications that support a largely heterogeneous and functional computational system approach to cognition (Kahneman, 2003; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutch, 2004).[28]

More recently, attention has been directed to the dual-nature of computation and cognition.[29] These so called 'dual-process theories' (e.g., Epstein, 2003; Evans, 2008; Evans & Frankish, 2009; Kahneman, 2011; Kahneman & Frederick, 2002; Lieberman, 2007; Stanovich, 2011; Stanovich & West, 2000) are becoming increasingly relevant in explaining a wide variety of cognitive

[28] They share many common features both with each other and earlier theoretical accounts (e.g. Epstein, 1973; Tversky & Kahneman, 1983).

[29] See also, Evans, 1989, 2008; Frankish, 2007; Samuels, 2009; Sloman, 1996, 2002; Stanovich, 1999, 2009.

phenomena.[30] In fact, empirical evidence that is congruent with dual-processing accounts of cognition have become so vast in the literature that Evans (2008) considered the task of drawing up a synoptic overview of the topic to be too complex and demanding. Despite the multiplicity, all the various theoretical adaptions share a single fundamental link. That is to say, all dual-process theories are predicated on the assumption that cognitive activity can be divided into two distinct computational processes or systems (Barrouillet, 2011). More specifically, two types of processing exist in the mind, with different evolutionary histories, different functioning, and possibly distinct 'cerebral substrates'[31] (Epstein, 1994; Evans, 2006; Evans & Over, 1996; Sloman, 1996; Stanovich, 1999, Stanovich, 2009).

Although the aforementioned is a necessary simplification of dual-process theories; the basic assumptions of the models, as well as the domains in which they have been applied and tested, can be described concisely. Precisely, two distinct types of computational information processing exist within a single mind: (1) a type of processing that is fast, reflexive and superficial (2) a type of processing that is controlled, precise and reflective.

The discussions to follow are aligned on the assumption of a generic dual-process theory. It should be noted that slight terminological discrepancies do exist between these theories. As a critical evaluation of the various terms identified in the literature is beyond the scope of this research, we will instead emphasize the most objective terms available.[32] Such a construct implies that

[30] According to Evans and Frankish (2009), such theories can account for a wide range of phenomena in the reasoning, judgment, and decision-making literatures that have been the subject of several recent books (Evans, 2007; Kahneman, 2011; Stanovich, 2011).

[31] The term 'cerebral substrates' emanates from neurobiological nomenclature. The aforementioned cerebral reference suggests that dual process theories have foundations in neuroscience.

[32] Namely; the distinction between System 1 and System 2 processes (Kahneman & Frederick 2002; Stanovich 1999; Stanovich & West, 2000).

cognition is composed of two multi-purpose behavioral systems, for which Stanovich and West (2000) proposed the neutral labels of System 1 and System 2.[33]

The terms ‘process’ and ‘system’ are defined as follows: a 'system' refers to any interconnected set of components, from which interaction emerges a property or capacity (the systemic property or capacity). Systems have a specified structure (their parts and the way they are interconnected) and certain behaviour (way(s) of processing information) (Wimsatt, 1986).

The term 'process' is meant as a general term that refers to any casual sequence that produces a certain output. The processes of concern here will be *cognitive/computational* processes, vis-à-vis processes that manipulate, store and transform information in order to help agents behave adaptively or intelligently in their environment (Beaulac, 2010).

2.2.1 Overview of Generic Dual Process Theories

Dual-process theories are ubiquitous in the recent literature in cognitive science. A central principle of these theories is that cognitive behaviour is determined by the interplay of two distinct systems of information processing. In other words, the mind is composed of two groups of computational processes, or two distinct *systems* having predetermined characteristics. Dual-process theorists usually speak in terms of System 1 (S1) and System 2 (S2).

Accordingly, System 2 is claimed to be the source of our capacity for complex, analytical and deductive reasoning.

[33] Unfortunately, these all-encompassing theories can result in terminological inflation and confusion and are sometimes referred to as dual-*system* theories, as opposed to dual-*process* ones. However, in order to elucidate the distinction between the two, I will henceforth follow Nobel laureate, D. Kahneman (2011) and use the prototypical dual-process axiom, whilst still maintaining the System 1, System 2, distinction. I borrow this construct from Kahneman (2011), who has greatly influenced my thinking.

System 1, in contrast, is often thought to be evolutionary primitive and largely applicable to situations demanding a so called urgent 'fight or flight' default response. It is also thought to be responsible for heuristic processing, such as those described in the heuristic and biases' literature.[34] (See also, Kahneman, 2003; Kahneman & Frederick, 2002; Kahneman & Tversky, 1982, 1996; Tversky & Kahneman, 1974, 1983).

For further clarity, the term *heuristic* is used to describe a method for problem solving, whereby a result is not guaranteed to be optimal, but is sufficient given a set of objectives (Ballard, 1999). Heuristic techniques speed up the process of finding a satisfactory solution via mental shortcuts in order to ease the cognitive load of making a decision. Invariably, this terminology (heuristic) seems to be reserved specifically for System 1 processing.

(Given the obvious nexus between heuristics and System 1 processing, one should expect heuristic techniques to feature quite prominently in our taxonomy, in the context of System 1 algorithmic traders. However, at this early stage of the thesis, we eschew consigning explicit traits to the agents in our taxonomy here. More specifically, we will avoid making restrictive inferences before exhausting the literature on both algorithmic trading and cognitive science).[35]

It should also be noted that heuristics are simple, efficient processes that people follow when facing complex problems or incomplete information. However, the faster speeds associated with heuristics are a result of mental shortcuts; shortcuts that often lead to systematic errors or cognitive biases. A cognitive bias refers to a person's propensity to think in a certain way that can lead to

[34] As Kahneman (2000) notes: "Tversky and I always thought of the heuristics and biases approach as part of a two-process theory" (p.682).

[35] Notwithstanding the above caveat, we do note that the intuition for relating heuristics to System 1 algorithmic trading appears to be quite reasonable. Hence, there is a strong chance that heuristics will feature prominently in our System 1 formulation for algorithmic trading.

systematic deviations from a standard of rationality or good judgment. Prominent cognitive biases, amongst a host of many others include: the gamblers fallacy and the anchoring effect. The gamblers fallacy refers to a judgment bias, whereby a person believes that past events alter future probabilities.[36] Alternatively, the anchoring effect refers to the tendency to 'anchor' or depend too much on a single piece of information when making a decision. Of course there are many other related errors and biases that exist in the plethora of literature on the subject (e.g. Jordan, Miller & Dolvin, 2012; Kahneman, 2000; Carmo and Luis, 2005).

Some of these limits are studied by the heuristics and biases research program in psychology, and these heuristic psychologists offer a great deal of evidence in favour of dual process theories of cognition (Evans, 2003). Accordingly, the tasks in this literature were specifically designed to pit a heuristically triggered response (S1) against an analytic (S2) response (Stanovich et al., 2008, 254).

Kahneman and Tversky (1982), the initial proponents of this research program, discovered that cognitive biases are not random: they follow a pattern because the mind relies on specific 'innate rules of thumb' (S1 processes).

[36]Notably, people are subject to the gambler's fallacy when they assume that a departure from what occurs on average, or in the long run, will be corrected in the short run. Namely, people often believe that because an event has not happened recently, it has become 'overdue' and is more likely to occur. People sometimes refer (wrongly) to the 'laws of averages' in such cases. Consider the following example: Roulette is a random gambling game where gamblers can make various bets on the spin of the wheel. There are thirty-eight numbers on a typical roulette table, two green ones, eighteen red ones, and eighteen black ones. One possible bet is to bet whether the spin will result in a red number or in a black number. Suppose a red number has appeared five times in a row. Gamblers will often become confident that that the next spin will be black, when the true chance remains at about fifty percent. Of course, it is exactly eighteen in thirty-eight (Jordan, Miller & Dolvin, p.270).

However, we also have deliberate and slow System 2 processes (S2) allowing us to find the *right* answer. This is, according to Stanovich et al., (2008), an important source of evidence in favour of dual-process accounts of the mind. Kahneman and Frederick (2008) agree and state:

> The persistence of such systematic errors in the intuitions of experts implied that their intuitive judgments may be governed by fundamentally different processes than the slower, more deliberate computations they had been trained to execute (p. 267).

Overall, dual process theories – specifically, the distinction between System 1 and System 2 processing – have generated fruitful explanations for a broad array of cognitive phenomena (Samuels, 2009). For some authors, like Carruthers (2006), the S1 / S2 distinction is a good way to think about the processes of the mind, although it does not pick up any deep functional difference between the two systems.

On the other hand, Lieberman (2007; 2009) asserts that there are deep functional differences between S1 and S2. The author notes that S1 and S2 are well-defined sub-systems that correlate to specific neuronal activation patterns, and goes as far as to attribute some cognitive capacities to specific brain regions. In fact, Lieberman (2007; 2009) suggests that there are two computational systems at work in the mind, c1early divided between what he calls the *ref1exive* X-system, or System 1, and the *ref1ective* C-system, or System 2.

It should be noted that our approach is aligned with, and builds on, Lieberman (2007; 2009). Thus, in the interest of objectivity, we use Lieberman's terminology throughout, specifically, the labels 'System 1' and 'System 2'.

Although Lieberman's account is consistent with dual process theories in general, he provides additional insight into cognitive processes, mostly stemming from his field of inquiry - cognitive

neuroscience. His research shows that each system is associated with precise processing characteristics. Lieberman clarifies that a clear trade-off exists between fast and accurate information processing and this trade-off has neurological or structural relevance. Using the tools provided by neurological sciences, Lieberman can infer, from the activation of a specific brain region, the phenomenological and representational characteristics of System 1 and System 2 information processing. Stanovich and West (2008) espouse this view, confirming that domain general processes can account for much of cognition.[37]

In most of the literature, the two systems are seen to interact directly with one another. They are also *competitive* in nature.[38] Consider the following narrative that highlights the interactive conditions underlying the two systems: As suggested by Carmo and Luis (2005, p.136), human behaviour is not the product of a single process, but rather reflects the interaction of different specialised subsystems. These systems, the idea goes, usually interact seamlessly to determine behaviour, but at times they may compete. The end result is that the brain sometimes argues with itself, as these distinct systems come to different conclusions about what we should do (Kahneman, 2011).

As noted, the major distinction responsible for these internal disagreements is the one between System 1 and System 2 processes. System 1 is generally reflexive and heuristic-based, which means that it relies on mental shortcuts. It proposes fast heuristic answers to problems as they arise. System 2, which corresponds closely with controlled processes, is slow, effortful, precise and

[37] According to Beaulac (2010), Stanovich's framework (Stanovich and West, 2008) can also account for a shift in neuronal activation. In fact, as articulated by Beaulac (2010, p.46), "Stanovich's account is really close to Lieberman's, as he agrees there are domain general processes that can account for much of higher cognition".

[38] These interactive conditions often result in their own distinct cognitive phenomena (Evans & Frankish, 2009). Therefore, analyses'- that look at the effects of these processes both individually and concurrently - that is, their constituent elements and dynamic interactions - have become the focus of much of recent cognitive exploration.

reflective and can also be employed to monitor the quality of the answer provided by System 1. Notably, if System 2 is convinced that the heuristic answer is wrong, then it is capable of correcting or overriding the automatic judgment.

However there are instances when System 2 processes cannot (or do not) override System 1 processes (such is the case with many observed behavioural biases). A major cause of these observed idiosyncrasies of decision-making may be that controlled processing accounts for only part of our overall behavioural repertoire, and in some circumstances can face stiff competition from reflexive processes that are part of System 1. Alter, Oppenheimer, Epley, and Eyre (2007) provide a compelling demonstration of this phenomenon. Alter et al., (2007) examine how subtle changes in contextual cues, such as altering the legibility of a font, can facilitate switching between System 1 and System 2 processing.

In a series of experiments, the authors manipulated the 'perceptual fluency' of various sets of stimuli i.e., they made it harder for people to understand or decipher the scenarios they were asked to judge. In one experiment participants were asked a series of questions designed to assess the degree to which System 1 intuitive processes are engaged in decision-making. In this test the gut (System 1) reaction answer is fast, but invariably incorrect. (An example: if a bat and a ball together cost $1.10, and the bat costs $1 more than the ball, how much does the ball cost? If you answer '10 cents', then you are in the majority, but unfortunately also wrong).[39] Alter et al., (2007) found that by making the problem more difficult to read (by using greyed-out, reduced-size font),

[39] The correct answer is 5 cents. Think about it this way: the problem doesn't say that the bat costs *exactly* $1…just that it costs an *additional* $1. So if the ball costs 10 cents, you'd be paying $1.10 for the bat alone. Instead, the ball has to cost five cents so that you're only paying $1.05 for the bat (Alter et al., 2007).

participants seemed to shift to more considered, System 2 responses, and as a result answered more of the questions correctly.

In a similar vein, a recent movement in behavioural economics seeks to acknowledge the limitations of everyday decision-making (such as the apparent reluctance of workers to contribute to 401K plans) and therefore design institutions in such a way as to 'encourage' better choices (such as introducing default options for retirement savings). Benartzi and Thaler (2007) have demonstrated that, when people are asked to commit to saving money in the distant future (as opposed to right now), they end up making much more economically rational decisions. This is because System 2 seems to be in charge of making decisions that concern the future, while System 1 is more interested in the present moment.

Overall, dual process theories – specifically, the 'distinction'[40] between System 1 and System 2 processing – has provided a very rich and interesting organisational framework from which to conceptualise and synthesize the vast amount of computational processes underlying human decision making. *This research follows a similar approach by presenting a dual-process taxonomy (as pioneered by Lieberman (2007; 2009)) as a possible means by which to conceptualize and synthesize the variety of computational processes underlying algorithmic trading – another decision making construct*.

[40] To summarise, distinction between the two systems can be described concisely: System 2 is a high-level processor, abstracting information and expressing knowledge as production rules. The representations in this system are symbolic and unbounded, in that they are based on propositions that can be combined to form larger and more complex sets of propositions (Sloman, 1996). Conversely, the representations within System 1 are characterized as associative and are automatically generalizable, allowing for a fast inferential process. Thus, unlike System 2, the representations of System 1 cannot be combined in novel ways, so that reasoning in System 1 is limited to what has already been represented.

In conclusion, the ideas behind this thesis are shaped by the narrative in cognitive science concerning the architecture of mind, particularly the computational processes underlying decision making. Remarkably, the cognitive approach allows one to consider human decision making and digital computation in parallel. Using the principals of contemporary cognitive science, this thesis will argue that the dual-process paradigm – the most prevalent contemporary interpretation of the nature and function of human decision making – lends itself to a novel taxonomy of algorithmic trading. This taxonomy will be used to inform several theoretical models of algorithmic trading; with the view of explaining the evolving nature of market quality as a consequence of this practice.

Current theoretical models of algorithmic trading (such as Cvitanic & Kirilenko, 2010; Foucault, Hombert & Rosu, 2016; Martinez and Rosu, 2011) are flawed as they fail to take into account the multidimensional aspects of its underlying trading strategies (see Section 2.6). By integrating and synthesising the multitude of algorithmic trading strategies within a single framework, it is hoped that our models may shed light on contradictory findings and on previously unknown market variables, thus providing new empirical insight.

As mentioned previously, the dual process view of the mind rests on certain intuitions regarding the fundamental similarity between computation and cognition. Arguably, such an advanced interpretation could not have come to fruition without the recent developments in computer hardware and software technologies. Inevitably, information technology has played a vital role in the emergence of the cognitive studies of science (Friedenberg & Silverman, 2011). Concomitantly, financial market transactions have been radically transformed by the information technology revolution over the course of the past decade or so.

2.3 THE CHANGING TRADING ENVIRONMENT

Over the last few years, technology has revolutionized the way in which financial markets operate (Jain, 2005). From the way trader's trade, to the way markets are structured, to the process of price discovery – all are now different in today's technologically advanced financial markets. Indeed, in the last decade or so, the forces of technology, speed, and computer-based trading have facilitated an extraordinary evolution in capital market structure, as well profoundly influenced the price-discovery process – the course by which prices reflect new relevant information (Gomber et al., 2011).

In 2001 a major structural event in North America - known as the 'decimalization of the price quotes on US stocks' - saw both the New York Stock Exchange (NYSE) and the National Association of Securities Dealers Automated Quotations (NASDAQ) migrate from a system of quoting stocks in fractions (e.g., 1/16th of 1$) to a system of quoting stocks in decimals (e.g., cents per share). This 'decimalization' resulted in minimum quote increments decreasing from $0.0625 per share to $0.01 per share.

Another structural event in the USA that had dramatic effects on trading was Regulation National Market System. Regulation NMS (henceforth Reg NMS) consists of Exchange Act Rules 600-612. The key provisions of Reg NMS were promulgated in 2005, and include: Rule 611 (the Order Protection Rule), Rule 610 (the Access Rule), Rule 612 (the Sub-Penny Rule), and amendments to Rules 601 and 603 (the Market Data Rules).

The literature ubiquitously confirm that Regulation NMS is located at the centre of the constellation of factors leading to the advent of algorithmic trading in the US (e.g., McNamara, 2016; Morelli, 2016). However, Reg NMS is a highly advanced piece of legislative material and thus the introductory nature of the current section precludes a thorough discussion on the

aforementioned regulatory framework. A concrete elucidation of Regulation NMS is consigned to Appendix I, where the nexus between Reg NMS and algorithmic trading can be addressed more explicitly. [41]

2.3.1 The Origins of Algorithmic Trading

While Reg NMS was instrumental in the development of algorithmic trading, recent technological advances in computer technology have also acted as a catalyst in the proliferation (and now dominance of) algorithmic trading. In recent years, information technology (IT) has successively assuaged the use of physical trading floors, allowing automated trading systems to dominate. Indeed, algorithmic trading exemplifies the narrative that technology has conquered the trading process. While the origins of algorithmic trading remain a topic of debate, Kirilenko and Lo (2013) suggest that the use of algorithms in trading emerged out of the desire to eliminate the need for human judgment. Accordingly, algorithmic trading represents a technique to overcome the computational limitations of the human mind and our limited human rationality (Biais & Woolley, 2011). Since our cognitive abilities are relatively limited, computerised trading programs function as a sort of 'mental prosthetic' for the failings of human rationality.[42] An individual human operator can only incorporate a small amount of data into his or her decision-making process, and so investment strategies requiring the processing of hundreds of thousands of variables necessarily

[41] The current study evaluated existing empirical research on the infrastructural aspects of algorithmic trading in order to determine the extent to which empirical research supports our dual process supposition. Essentially, one can distinguish between two independent forms of infrastructure when it comes to algorithmic trading (co-location infrastructure and high-end capability computing infrastructure). As clarified in Appendix I, algorithmic trading infrastructure and Reg NMS are inexorably linked. Therefore, the concrete elucidation of Regulation NMS is consigned to Appendix I, where the nexus between Reg NMS and algorithmic trading can be addressed more explicitly vis-à-vis**,** co-location and high-end capability computer infrastructure.

[42] See e.g., Salmon and Stokes (2010), particularly the description of areas where algorithmic trading excels over human operators.

require computer technology. Likewise, human traders find it increasingly challenging to monitor prices on more than a very small number of financial markets simultaneously, so they require computer assistance (McNamara, 2016).

With that said, it is important to note that algorithmic trading is a relatively new phenomenon and has not yet been clearly defined. Thus, unsurprisingly, the definitions of algorithmic trading range from the very general - "the use of computer algorithms to automatically make trading decisions, submit orders, and manage those orders after submission" (Hendershott & Riordan, 2009, p. 2) to the specific:

> Algorithmic trading is computer-determined trading whereby super computers and complex algorithms directly interface with trading platforms at high speed, placing orders without immediate human intervention. It (algorithmic trading) employs cutting edge mathematical models, adept computational techniques and extraordinary processing power via advanced computer and communication systems and is capable of anticipating and interpreting market signals in order to implement profitable trading strategies (Gamzo, 2014, p.45).

While the above definitions capture most activities referred to as algorithmic trading, it should be emphasised that algorithmic trading is not one thing. Algorithmic trading, like many other financial market activities, is complex, and encompasses an array of specific practices. In terms of strategies, both numerous and diverse trading approaches currently exist. Both the International Organization of Securities Commissions (IOSCO, 2011) and the US Securities and Exchange Commission (SEC, 2014) have emphasised that distinguishing such strategies is pivotal in any research design:

"*Algorithmic trading is not a single strategy but it is rather a set of technological arrangements and tools employed in a wide number of strategies, each one having a different market impact"* (IOSCO, 2011, p.24).

Effectively, when attempting to assess algorithmic trading, it is necessary to study the individual strategies that make up this practice closely (Gomber et al., 2011). While the universe of algorithmic trading strategies is too diverse and opaque to name them all, we attempt to delineate some of the most prominent strategies identified in the literature below. This is based on the work of Almgren (2009) and includes information from Gomber et al., (2011), Gamzo (2014), as well as Johnson (2010).

In line with the above, the following subsections shed light on some of the best known and probably most prominent algorithmic trading based strategies. Appendix I concludes with a functional organisation of algorithmic trading strategies. The comparative summary in Appendix I assigns each individual algorithmic trading strategy to its respective system.

2.4 AN OVERVIEW OF ALGORITHMIC TRADING STRATEGIES

Spread Capturing Algorithms

These algorithms profit from the spread - between the bid and ask prices - by continuously buying and selling securities by computers (Gomber, Arndt, Lutat & Uhle, 2011). With each trade, they reap the spread between the (higher) price at which market participants can buy securities and the (lower) price at which they can sell securities.

As emphasised above, we do not categorise the individual strategies vis-a-vis their respective system here (before exhausting the necessary literature) - this task is assigned to Appendix I. The same holds for the discussions to follow.

Rebate Trading Algorithms

Rebate trading concerns incentive scheme revenue capturing. In order to attract liquidity providers, some trading venues have adopted asymmetric pricing: members removing liquidity from the market (taker; aggressive trading) are charged a higher fee while traders who submit liquidity to the market (maker; passive trading) are charged a lower fee or are even provided a rebate. An asymmetric fee structure is supposed to incentivize liquidity provision (Gomber, et al., 2011). Upon execution of bids, offers are immediately posted on the inside market with the intent of capturing rebates regardless of whether or not capital gains have been achieved.

The rationale for applying maker-taker pricing is as follows: traders supplying liquidity on both sides (buy and sell) of the order book earn their profits from the market spread. Fee reductions or even rebates for makers shall stimulate a market‘s liquidity by firstly attracting more traders to post passive order flow in form of limit orders. Secondly, those traders submitting limit orders shall be incentivized and enabled to quote more aggressively, thus narrowing the spread. The respective loss of profits from doing so is thus compensated by a rebate. If this holds true, those markets appear favourable over their rivals and market orders are attracted enhancing the probability for the makers to have their orders executed (Lutat 2010).

Time Weighted Average Price (TWAP) Algorithms

This strategy involves the realization of a specific pre-determined benchmark, or the Time Weighted Average Price (TWAP). TWAP algorithms divide a large order into slices that are sent to the market in equally distributed time intervals. Before the execution begins, the size of the slices as well as the execution period is defined. For example, the algorithm could be set to buy

12,000 shares within one hour in blocks of 2,000 shares, resulting in 6 orders for 2,000 shares which are sent to the market every 10 minutes. TWAP algorithms can vary their order sizes and time intervals to prevent detection by other market participants (Gomber et al., 2011).

Volume Weighted Average Price (VWAP) Algorithms

VWAP algorithms try to match or beat the volume weighted average price (their benchmark) over a specified period of time. VWAP can be calculated applying the following formula for n trades, each with an execution price *pn* and size *vn* (Johnson 2010):

$$\text{VWAP} = \frac{\text{Overall Turnover}}{\text{Total Volume}} = \frac{\sum_n v_n p_n}{\sum_n v_n}$$

Since trades are being weighted according to their size, large trades have a greater impact on the VWAP than smaller ones. VWAP algorithms are based on historical volume profiles of the respective equity in the relevant market to estimate the intraday/target period volume patterns.

Implementation Shortfall Algorithms

More multifarious than both TWAP and VWAP strategies, implementation shortfall algorithms seek to manage the trade-off between market impact and timing risk. In order to reduce the market impact of large orders, implementation shortfall algorithms consider timing risk - the possibility of adverse price reactions during the execution process. Consequently, these algorithms pre-determine an execution plan based on historical data, and split an order into as many as necessary but as few as possible sub-orders. (Gomber et al., 2011). In contrast to TWAP or VWAP, these orders will be scattered over a period which is just long enough to dampen the market impact of the overall order (Johnson 2010).

Adaptive Execution Algorithms

Adaptive execution algorithms follow a similar approach to TWAP and VWAP, except they are more adaptive in nature. Instead of following a pre-determined schedule, they are able to re-evaluate and change their execution schedule with changing market conditions (Johnson 2010; Gamzo 2014).

Liquidity Detection Algorithms

Liquidity detection algorithms attempt to decipher patterns made by other investors with the aim of making a profit. They generally concentrate on detecting large orders, so that an automated version of 'quote matching' can be employed. When the algorithm detects large orders, it immediately places its own buy order with a minimally higher limit, believing that the large orders will increase the price of the asset. This type of activity profits from a rise in share price. However should the price fall, the algorithm is able to use the large order as a hedge against which it can sell its shares.

Data\Text Mining Algorithms

Data mining can be defined as the practice of isolating legitimate, unidentified, coherent and actionable information from large databases and using it to make critical business decisions (Brusilovsky & Brusilovskiy, 2008). This type of information may be inferred from correlations between assets, both in the same market, across different markets or even different asset classes

(Johnson, 2010).[43] Data mining can assist in the process of discovering these relations, allowing one to generate forecasting models based on wide ranges of data. On the other hand, text mining concerns the automated classification of textual information by transforming unstructured information into machine readable format and using it to make astute transactional decisions (Gamzo 2014).

Neural Network Algorithms

Neural network algorithms have gained prominence recently including plotting input-output vectors for cases where traditional models fail to hold. Neural networks are information processing paradigms with a remarkable tolerance for noise, ambiguity and uncertainty. Suhas and Patil (2011) explain neural networks as "a collection of mathematical processing units that emulate some of the observed properties of biological nervous systems and draw on the analogies of adaptive biological learning" (p.2). Neural network algorithms are especially useful for recognizing relationships in convoluted and complicated data sets and are only limited by the power of their relative platform or infrastructure. Their remarkable ability to derive meaning from vast, complicated and imprecise data allows them to detect patterns and identify trends that are too intricate to be noticed by humans alone. In fact, research has documented their ability to accurately forecast future market movements for a variety of different instruments and markets.[44] Therefore by combining the search capabilities with the modelling power of the neural networks, a useful predictive tool can be created (Foster, 2002).

[43] For example, lead/lag relationships may be found between different assets across different markets and classes (see e.g. Johnson, 2010).

[44] For instance, Walczak (2001) observes that a neural network was able to forecast foreign exchange rates for a variety of currencies. Hutchinson, Lo and Poggio (1994) compose a neural network for predicting Standard & Poor's 500 futures options prices. Whereas, Castiglione (2001) construct neural network models to predict a variety of financial time series.

Support Vector Machine Algorithms

In many ways, support vector machine algorithms share many of the same characteristics as neural network algorithms, although their training is very different. Essentially, a support vector machine model is an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints. This is done instead of solving a non-convex, unconstrained minimization problem like those in standard neural network training.[45] Importantly, their ability to cope with problems spanning multiple dimensions makes them an excellent prediction tool for algorithmic traders in today's convoluted financial markets.[46]

2.5 ALGORITHMIC TRADING: EXTANT LITERATURE

Algorithmic trading remains one of the most controversial and actively discussed topics in the financial world. Inevitably, there are myriad of questions on the topic of algorithmic trading attracting the attention of researchers (O'Hara, 2014). Some of the more fundamental questions seem to concern the evolving nature of market quality as a consequence of algorithmic trading. Traditional literature considers market quality parameters (i.e., trading activity, prices, liquidity, volume, volatility and profits) as representative measures of the price discovery process – the process by which a market incorporates new fundamental information about an asset's value into the asset's price (e.g., Fama, 1970; Hayek, 1945). The price-discovery process is a central function

[45] For a more detailed view of support vector machines see Bennett and Campbell (2000).

[46] For example, Van Gestel et al., (2001) used a support vector machine to forecast time series and associated volatility for US short-term interest rates along with German DAX stock index. Regarding the sign of forthcoming returns, support vector machines proved to have around 5% greater predictive accuracy when compared to traditional methods. In addition, Mills (1991) found that support vector machines are a superior forecasting method when it came to predicting the weekly direction of NIKKEI 225 Index.

of a financial market, whereby prices offer accurate signals for resource allocation and the efficient distribution of an economy's capital stock (Fama, 1970). Accordingly, elucidating the effects of algorithmic trading on financial market quality is vital to our understanding of financial market performance (Gomber et al., 2011).

So what effect is algorithmic trading having on financial market quality?

Unfortunately, this question, however pertinent, remains the most controversial and open question in both theoretical as well as empirical financial market research. Indeed, a remarkable gap between the results of academic literature can be observed. This section of Chapter 2 will review the often-contradictory empirical literature on algorithmic trading. We follow with a discussion on the developing theoretical work in this field.

2.5.1 Algorithmic Trading: Empirical Approach

The academic discourse on the topic of algorithmic trading has been largely empirical in nature. However, this literature comes with well documented limitations (Biais & Foucault, 2014).

Obtaining useful data that can identify algorithmic trading activity is a formidable challenge. Much of the empirical evidence on the direct impact of algorithmic trading on the U.S. equity markets has relied heavily on either limited samples of *proprietary data* or *publicly available information* (SEC, 2014). By definition, proprietary datasets are not publicly available, have been limited to particular products, are discretionary in nature and show only a small amount of algorithmic trading activity (SEC, 2014).

Alternately, publicly available data on orders and trades does not reveal the identity of buyers and sellers. As a result, at this time, it is *not possible* to identify orders and trades as originating from an algorithmic trader when relying solely on publicly available information. Therefore, a variety

of different metrics are used by researchers to estimate algorithmic trading. Indeed, analysing actual algorithmic trading is based on empirical proxies,[47] and results thus rely heavily on the quality of the proxy (Chung & Lee, 2016).

2.5.2 Empirical Proxies for Algorithmic Trading

As noted, a precise empirical assessment of algorithmic trading and its impact on market quality is quite challenging. This is partly a reflection of the complexity of inferring algorithmic trading from the data (Chung & Lee, 2016; SEC, 2014). Moreover, since algorithmic trading is not clearly defined in the literature,[48] researchers are bound to use datasets with different definitions of algorithmic trading. Perhaps quite naturally, researchers have proposed different methods to *proxy* for algorithmic trading activities within traditionally-available datasets. For example, Zhang (2010) defines algorithmic trading broadly as all short-term trading activities of institutional investors that are not covered in Form 13f - the quarterly holdings report of large institutional investors in the United States.[49] Jones (2013, p.32) however, expresses concerns over this proxy because it is difficult to attribute the ensuing result purely to algorithmic trading activities.[50]

---

[47] A proxy can be defined as a figure that can be used to represent the value of something in a calculation.

[48] There has not, to date, been a consistent academic or regulatory definition of the term algorithmic trading. The term has been used in different ways and for various purposes. Invariably, *Automated Trading, Flash Trading, Program Trading, Low Latency Trading, Black Box Trading, Electronic Trading* and *High Frequency Trading* are just some of the labels ascribed to it in the literature (Kissell, 2013).

[49] Zhang's (2010) proposed proxy relies on elimination rather than identification. Zhang's (2010) intuition is as follows: In the United States, institutions with over $100 million in Assets under Management (AUM) are required to report their long term holdings in the 13f quarterly report of equity holdings. However, short positions are not required to be disclosed and are excluded from the report. Thus, by measuring trading volume relative to institutional portfolio changes in quarterly 13f filings, Zhang (2010) argues that his definition effectively proxies for algorithmic trading; that is because it captures trading frequencies greater than those of long term traditional/non algorithmic investors.

[50] Such a proxy demands two assumptions to facilitate an estimate of algorithmic trading. First, no algorithmic trading is assumed to have existed prior to the beginning of 1995. His second assumption is that algorithmic traders do not hold positions at the end of any quarter and thus do not file 13f reports with the SEC.

Kirilenko et al., (2011) define algorithmic traders as market participants with extremely high trading volume and well-balanced inventory. The SEC (2014) however, suggests that this proxy is extremely narrow and therefore fails to identify a large block of algorithmic trading activities.

Hasbrouck and Saar (2013) develop an algorithmic trading proxy called 'strategic run', which they describe as a "series of submissions, cancellations, and executions that are linked by direction, size, and timing, and which are likely to arise from a single algorithm" (p. 660). Conversely, Chung and Lee (2016) argue that broad inferences are difficult from Hasbrouck and Saar's (2013) limited sample.

In perhaps the greatest indictment of existing empirical studies to date, the SEC (2014) warns that interpreting the results of empirical studies should be done with caution; " an assessment of empirical papers must deal with the various metrics (proxies) researchers used to define algorithmic trading, and how their definitions may affect their conclusions about algorithmic trading activity… Particularly, the different metrics used to classify algorithmic trading can greatly affect findings about key characteristics of algorithmic trading and its impact on financial market quality" (p.5).[51]

2.5.3 Empirical Evidence: Market Quality and Algorithmic Trading

As qualified above, the bulk of the empirical academic literature regarding algorithmic trading thus far has emphasised market quality as an important standard by which to judge the impact of algorithmic trading (e.g., Zhang, 2010). Although the effects in those areas have dominated the

[51] Other examples of such AT proxies derived from market-wide data include high message rates, bursts of order cancellations and modifications, high order-to-trade ratios, small trade sizes, and increases in trading speed. Note that a critical evaluation of the various proxies identified in the literature is beyond the scope of this thesis. We simply note that the diversity proxies for algorithmic trading can greatly affect findings about the key factual characteristics of algorithmic trading activity.

literature, little consensus exists in the empirical literature regarding the overall impact of algorithmic trading on financial market quality.

One consistent refrain from supporters and practitioners of algorithmic trading is that algorithmic trading makes prices more informative.[52] Indeed, a number of empirical studies suggest price discovery gains from algorithmic trading. Hendershott and Riordan (2011) look at all NASDAQ trades in 2008 and 2009 and conclude that algorithmic traders reduce temporary pricing errors and quickly incorporate information into prices. According to Hendershott and Riordan (2011), algorithmic trades were positively correlated with permanent price changes and negatively correlated with transitory price changes, suggesting that algorithmic trading improves price discovery.

Concomitantly, Hendershott and Moulton (2011) report that the algorithmic trading reduced market noise; thereby aiding in the price discovery process. Algorithmic traders also place tighter quotes in the German stock market in a sign of increased price discovery due to price agreement (Hendershott & Riordan, 2013). Chung and Hrazdil (2012) show, using data for more than 2,000 firms in 2008, that algorithmic traders improve price discovery by increasing the speed of adjustments for all firms.[53] Humphery-Jenner (2011) demonstrates that algorithmic trading activity increases before important events such as takeovers. Accordingly algorithmic traders incorporate their prior knowledge into prices, even before events occur.

[52] The notion that prices in financial markets reflect and convey information is axiomatic within the realm of finance and is enshrined in the theory of market efficiency (Fama, 1970). Hayek (1945) wrote eloquently on this function of prices in sharing information among economic participants. He described prices as a form of telecommunication, enabling producers and consumers, in general all decision makers, to alter their behavior due to the essential information contained in one measure, price.

[53] The improvements for large firms are even more substantial, suggesting that algorithmic trading is useful across the market but that activity is greater for larger, more liquid stocks.

Zhang (2012), in contrast, suggests that algorithmic traders improve price discovery for only certain kinds of news; mostly numerical and short-run data. Contextually, Zhang (2012) divides news into 'hard' (numerical data or machine readable news) and 'soft' (text-based or informal sources). She finds that algorithmic traders are able to process hard data quickly and earn short-term profits, while non-algorithmic traders focus their efforts on soft data and take longer to trade so that prices reflect such information.

Though much of the empirical literature shows improvements in pricing discovery, there are detractors. Zhang (2010) examined the impact of algorithmic trading over the period 1985-2009. By using dividend surprises and analysts forecast revisions as proxies for firm fundamental information news, Zhang (2010) suggests that algorithmic trading is negatively associated with the market's ability to incorporate news about fundamentals into asset prices. Specifically, his paper showed that prices seemed to deviate systematically from their 'fundamental values'[54] when algorithmic trading was more evident. Finally, Zhang and Powell (2011) surmise that algorithmic traders have no intrinsic interest in a company's fundamentals and have a negative impact on the price discovery process.

It should be noted that the preceding discussion considers the empirical literature vis-à-vis the impact of algorithmic trading on the price discovery process. Market quality itself, however, is more inclusive than price discovery and consists of a variety of interrelated concepts or parameters.[55]

[54] The term fundamental value is formally defined in Chapter 4, Section 4.3.

[55] As noted throughout, market quality refers to a market's ability to meet its central function of price discovery. Market quality is thus more general than price discovery itself. Note however, that a variety of different variables are relevant here including, but not limited to trading activity, prices, volume, liquidity, volatility and profits.

The discussion to follow provides an overview of the empirical literature on algorithmic trading; focussing on the nexus between algorithmic trading and existing market quality metrics i.e., liquidity, volume, trading activity and volatility. Section 2.6 reviews the growing theoretical literature. (For easy reference, Table 2, at the conclusion of the current chapter will present a combined summary of empirical and theoretical literature in the field of algorithmic trading).

2.5.3.1. Liquidity and Algorithmic Trading

There is a contentious argument in the academic literature regarding the relationship between liquidity and algorithmic trading, and its resolution depends to a large extent on the chosen definition of liquidity. Liquidity is a complex construct and a precise and consistent definition remains elusive. Kyle (1985) suggests that liquidity encompasses three different components: tightness, depth, and resiliency. Since then, researchers have expanded on Kyle (1985) and utilised such varied measures as average dollar volume (Pastor & Stambaugh, 2003), quoted and effective spreads (Chalmers & Kadlec, 1998), share turnover (Chan & Faff, 2003), market depth (Chordia, Roll, & Subrahmanyam, 2001), number of zero volume days (Lesmond, Ogden, & Trzcinka, 1999; with an opposing view expressed by Mazza, 2015), and resiliency or price impact (Amihud, 2002).

Against the basic concerns reflected in different aspects of liquidity, researchers generally report improvements in market liquidity parameters (such as those discussed above) following the advent of algorithmic trading. For instance, Hendershott, Jones, and Menkveld (2011) document narrower spreads among U.S. stocks after the introduction of the NYSE's automated quote dissemination in 2003. They further claim this is evidence of improved liquidity due to algorithmic trading.[56] Similar findings are provided by Hasbrouck and Saar (2013), who measure algorithmic trading by

[56] Hendershott, Jones, and Menkveld (2011) measure algorithmic trading as electronic message traffic and report liquidity improvements for NYSE stocks from 2001 to 2005 based on the evidence of narrower spreads.

strategic patterns of order submission and cancellation. The authors report decreased spreads for NASDAQ stocks in 2007 and 2008 and suggest that this is a result of algorithmic trading. Chaboud, Hjalmarsson, Vega and Chiquoine (2009) study algorithmic trading in foreign exchange markets. They identify algorithmic trading as orders that are submitted using the electronic brokering systems (EBS) offered by Reuters. Consequently, these researchers posit that automated trading tends to slightly increase liquidity provisions in the foreign exchange markets after exogenous market events (such as news announcements).

However, not all studies report improvements in all aspects of liquidity. Brogaard (2010) uses trade and quote (TAQ) data for 2009 and 2010 to examine the liquidity effects of algorithmic trading. He documents an increase in volume and frequent quotes at the inside spread, though the provided depth is less than other traders. Jarnecic and Snape (2014) report similar findings of tighter spreads but less depth for the London Stock Exchange in 2009. Gai, Yao, and Ye (2012) also suggest algorithmic trading has a detrimental effect on market liquidity. Using data for NASDAQ stocks in 2009 and 2010, they observe a decrease in market depth due to algorithmic trading.

We conclude the present discussion by noting another important caveat. If liquidity and volume were identical, then the debate over the effect of algorithmic trading on liquidity would be moot, since the increase in volume is obvious even to casual observation (Chordia, Roll, & Subrahmanyam, 2008). Indeed, by nearly every empirical measure, algorithmic trading increases market volume (Hanson, 2014).

Concurrently, most studies report a reliable and direct relation between trading volume and price volatility.[57]

If we go by the above suggestion that algorithmic trading almost certainly increases trading volume, then researchers are justified in raising concerns about the broad effects of algorithmic trading and volatility. The next subsection considers studies that have sought to establish whether a relationship exists between algorithmic trading and volatility specifically.

2.5.3.2. Volatility Effects of Algorithmic Trading

The debates surrounding the nexus between algorithmic trading and price volatility were likely the result of a single event known as the Flash Crash. Given the Flash Crash is often-cited as the most direct evidence for the case that algorithmic trading has increased volatility we consider it in more detail below.

The 'Flash Crash' of May the 6th, 2010 resulted in the largest single-day point decline in the history of the Dow Jones Industrial Average (998.5). For about 5 minutes, approximately $1 trillion in market value had disappeared, only to bounce back just as quickly. The two most affected markets appeared to be the E-Mini futures market and the equities market. Although the exact

[57] Using daily individual security data, 1981-1983, Lamoureux and Lastrapes (1990) find a positive conditional volume-volatility relationship in models with Gaussian errors and GARCH-type volatility specifications. Using monthly measures, 1885-1987, Schwert (1989) finds a positive relationship in linear Koyck distributed lag regression of estimated volatility on current and lagged volume growth.Gallant, Rossi, and Tauchen (1990) investigate the relation between price and volume using a semi-nonparametric method. In their time-series analysis, they find that daily trading volume is positively related to the magnitude of daily price changes and that high volume follows large price changes. Crucially, *the relationship between trading volume and volatility* is found to hold in U.S. stocks in the time series from 1928 to 1987. Their finding of an unconditional volume-volatility relationship is consistent with many other studies (see Karpoff, 1987; Tauchen & Pitts, 1983).

degree to which algorithmic trading was responsible for the 'Crash' remains unknown, it seems clear from the evidence below that it was a major contributing factor (Gamzo, 2014).

The Flash Crash – 6th of May 2010

According to a report published by the SEC (2010), two separate incidents occurred on that day: (1) a liquidity crisis in broad index level in the E-Mini futures market and (2) a liquidity crisis in individual stocks. The SEC (2010) provides a detailed report of the events that led to the 'Flash Crash' of May, 2010 and should be consulted for further clarity. A summary of the SEC's (2010) findings are provided below:

On the 6th of May 2010 a large sell order (set to 9% of the trading volume calculated over one minute earlier) was initiated by a trader via an automated trading algorithm, programed to feed orders into the E-Mini futures market. This 'sell algorithm' was set to ignore market price and market timing and only focus on quantity. The sell pressure was *initially* absorbed by 3 distinct market participants:

1. Other algorithmic traders seeking to profit from the subsequent price increases.
2. Traditional buyers in the futures market.
3. Cross-market arbitragers.[58]

The above net buyers subsequently accumulated temporary long-term positions. However, about 60% of algorithmic traders net long positions were sold. The resulting growth in volume prompted the initial 'sell algorithm' to feed more orders into the market, even though the previous orders

[58] Who transferred this sell pressure to the equities markets by opportunistically buying E-Mini contracts and simultaneously selling individual equities in the S&P 500 Index.

had not yet been fully absorbed by the other participants. The result was a drop in E-Mini prices by about 3% in just 240 seconds.

By 2:45 pm the 'Stop Logic Function' saw a pause in trading for about 5 seconds, resulting in a decrease in sell side pressure and an increase in buy side pressure. This was followed, almost immediately, by price stabilization and recovery. Overall in just the four and a half minutes from 2:41pm prices on the E-Mini had sunk by more than 5%, only to recover moments later (Gamzo 2014).

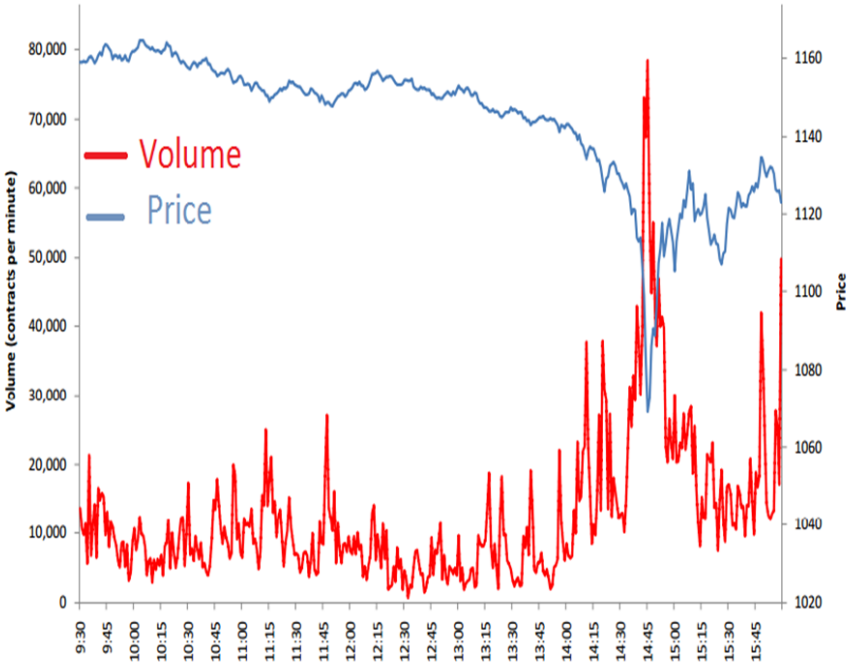


FIGURE 2- E-MINI VOLUME AND PRICE. SOURCE: SEC (2010). FINDINGS REGARDING THE MARKET EVENTS OF MAY 6, 2010. P. 19

The other liquidity crisis occurred in the equities market approximately 30 seconds before trading resumed in the E-Mini market at about 2:45 pm. Around 8000 individual stocks were traded, with the majority displaying similar price declines and reversals as those in the E-Mini futures market. "Over 20,000 trades across more than 300 securities were executed at prices more than 60% away from their values just moments before"(SEC , 2010, p.1) By the end of the day, major futures and equities indices recovered to close at losses of about 3% from the prior day.

Although an intra-day event, the 'Flash Crash' of 2010 cannot be considered inconsequential. On the contrary, it has been argued that the 'Flash Crash' of May the 6th 2010 represents the strongest evidence in support of the hypothesis that algorithmic trading has a destabilizing effect on the market (Gamzo 2014).

Perhaps the rarity of officially commissioned studies further emphasises the importance of understanding the relative impact of algorithmic trading. The SEC provided a short summary of the lessons to be learnt from the 'Flash Crash' of 2010:

> One key lesson is that under stressed market conditions, the automated execution of a large sell order can trigger extreme price movements, especially if the automated execution algorithm does not take prices into account. Moreover, the interaction between automated execution programs and algorithmic trading strategies can quickly erode liquidity and result in disorderly markets. As the events of May 6 demonstrate, especially in times of significant volatility, high trading volume is not necessarily a reliable indicator of market liquidity (SEC, 2010, p.6).

The 'Flash Crash' brought to light algorithmic trading's ability to exacerbate price movements in times of financial stress. Zhang (2010) examined the impact of algorithmic trading in a broader economic context using the period 1985-2009. Using data from the Centre for Research in Security

Prices (CRSP) and the Thomson Reuters Institutional Holdings databases in the U.S, Zhang (2010) found a positive correlation between algorithmic trading and stock price volatility. He revealed that even after controlling for fundamental firm-specific volatility, as well as other exogenous volatility variables, a single standard deviation increase in algorithmic trading activity is associated with a 5.6% rise in volatility.

In stark contrast to the above cited studies relating algorithmic trading to increased volatility; several other papers have reported no significant volatility effects of this practice. Hasbrouck and Saar (2013) analyse NASDAQ stocks from October 2007 and June 2008 and report that algorithmic trading does not lead to an increase in market volatility. On the contrary; the authors find that algorithmic trading activity lowers short-term volatility, reduces quoted spreads and the total price impact of trades, and increases depth in the limit order book. Moreover, Chaboud, et al., (2009) examine the foreign exchange market and find no relationship between algorithmic trading and volatility. A similar result is reported by Brogaard (2010) for 120 NASDAQ stocks in 2008 and 2009. Finally, Groth (2011) finds no relationship to volatility in the German DAX30 stocks in 2007, while Gsell (2008) takes a simulation approach and finds no significant increase in volatility.

To summarise, researchers have shown in various contexts that algorithmic trading both decreases (Hasbrouck & Saar, 2013) and increases measures of volatility (Zhang, 2010). Contextually, empirical literature on the topic of algorithmic trading has also not been able to identify

(concretely) the existence of a 'trade-off'[59] between liquidity and volatility – a recurring phenomenon in non-algorithmic related empirical research (Chordia et al., 2008; Hanson 2014).

2.5.4 Empirical Studies: Concluding Remarks.

The review above considered the extant empirical literature on algorithmic trading. Our discussion was prefaced by a broad summary of the inherent limitations associated with empirical studies to date.

Ultimately, the lack of precise data is sufficient enough reason to believe that the goal of better understanding algorithmic trading strictly from empirical observations is unattainable. Further difficulty is then encountered when one considers that algorithmic traders employ a diverse number of trading strategies. Indeed, we are not aware of any empirical studies that have been able to distinguish different algorithmic trading strategies.

Overall, given the data challenges facing researchers, and the variety of strategies and motivations for engaging in this practice, it is not surprising that little consensus exists in the empirical literature regarding the overall impact of algorithmic trading on financial market quality.

In the remaining subsections of this chapter we analyse the growing theoretical literature on algorithmic trading.

[59] Mike and Farmer (2008) argue that understanding liquidity is the first and principal step to understanding volatility. Previous work has shown that liquidity is typically the dominant determinant of volatility, at least for short time scales. Crucially, periods of high volatility correspond to low liquidity and vice versa (Farmer & Lillo, 2004; Gillemot, Farmer & Lillo, 2006; Weber & Rosenow, 2006).

2.6 THEORETICAL LITERATURE: ALGORITHMIC TRADING

The theoretical literature on the topic of algorithmic trading has its foundations in market microstructure theory (reviewed in Chapter 4). Given that the distribution of information is a central consideration in market microstructure modelling, the theoretical literature attempts to address related issues of information differentials between algorithmic traders and non-algorithmic traders and its subsequent market quality effects (Grossman, 1976).

The vast majority of papers have coalesced around the idea that algorithmic traders possess a comparative informational advantage relative to regular traders. However, research is divided as to what type of informational advantage algorithmic traders possess (e.g., Biais et al., 2011; Cartea & Penalva, 2012; Cvitanic & Kirilenko, 2010; Das et al., 2001; Easley et al., 2012; Foucault et al., 2016; Gamzo, 2014; Johnson, 2010; Martinez & Rosu, 2011).

Crucially, there remains a huge disconnect between available theoretical models of algorithmic trading and recent empirical research on the characteristics of algorithmic trading - purporting that algorithmic trading is a diverse phenomenon. Indeed, the developing empirical literature on the topic of algorithmic trading suggests that algorithmic trading constitutes several different trading strategies, each with its own distinct character and market impact. Unfortunately, theoretical models on the topic have been largely strategy specific, simplified, analytically tractable models with a *single* representative algorithmic trader. Researchers tend to focus on a specific algorithmic trading strategy and often conflate that specific strategy with algorithmic trading in its entirety. Indeed, the SEC (2014) notes the many instances in which researchers attempt to determine the impact of algorithmic trading as a whole, followed almost immediately by an emphasis on a specific isolated strategy - without an explicit indication or explanation for such a strategy specific approach.

Insofar as the literature is concerned, Cvitanic and Kirilenko (2010) provide the first ever theoretical model of algorithmic trading, in an attempt to address the theoretical impact of algorithmic trading on market quality. The authors construct a theoretical model that adds algorithmic traders (machines) into a market populated by non-algorithmic traders. However, they make an extremely restrictive assumption by assuming that algorithmic traders collectively follow a single strategy known as sniping.[60] This sweeping supposition is in stark contrast with the growing empirical evidence that supports the multiplicity of algorithmic trading strategies (Johnson 2010).

Unfortunately, by likening algorithmic trading to a single activity, Cvitanic and Kirilenko (2010) restrict themselves to the analysis of an isolated portion of algorithmic trading.[61] They do not provide an explanation for such a strategy specific approach. Inevitably, in their model, the only

[60] Sniping is a strategy designed to discover liquidity in the limit order book, or to 'pick-off' orders already in the book (Cvitanic & Kirilenko, 2010).

[61] Most importantly perhaps, Cvitanic and Kirilenko (2010) expose a commonly articulated flaw in existing theoretical models. As referenced throughout: Current theoretical models of algorithmic trading tend to focus on a *specific* algorithmic trading strategy, and often conflate that strategy with algorithmic trading in its entirety. In perhaps the most notable reflection of this (limiting) strategy specific approach, Cvitanic and Kirilenko (2010) *admit* that their model relates only to a specific strategy employed by algorithmic traders. Contextually, they state: "This is only one of the strategies used by actual algorithmic traders in real markets, and the only one we focus on" (p.2).

difference between normal and algorithmic traders – 'machines' – is that the latter have an inherent speed advantage relative to the rest of the market.[62]

Again, this common informational assumption is based exclusively on a speed advantage and is at odds with recent evidence on algorithmic trading that suggests that a number of different strategies underlie algorithmic trading (with different strategies having different informational characteristics and market impact). In addition to being less rich as a description of real markets, models based exclusively on one type of information are often empirically misspecified. Nevertheless, Cvitanic and Kirilenko (2010) find that the presence of the algorithmic trader is likely to affect trading volume and intertrade duration, i.e. the time span between trades. Their research implicates an increase in market liquidity measures based on trading volume and intertrade duration. Specifically, they find that the introduction of an algorithmic trader reduced the average trade value and resulted in lower volatility.

[62] In their paper Cvitanic and Kirilenko (2010) study the distribution of transaction prices generated in an electronic limit order market populated by orders from algorithmic traders (machines) and non-algorithmic traders (humans). They focus on the period between two human transactions – a very short period of time in a liquid market. The authors posit that during such a short horizon, the impact of changes in the fundamentals is negligible. Therefore, they model the incoming human buy order prices and sell order prices during the period as two *iid* sequences, arriving according to exogenous Poisson processes. For tractability, they also assume that the submitted orders are of unit size and at infinitely divisible prices. In the actual limit order book environment, traders submit orders of different quantities at discrete price intervals - ticks. At each tick, quantities get stacked up in accordance with a priority rule, e.g., time priority or order size and then time priority. Their idealized model, with the order prices coming from a continuous distribution and for one order only, can be thought of as taking the actual orders for multiple units stacked up at each tick and 'spreading' them between ticks. Thus, the algorithmic trader's speed advantage means that they can submit and cancel orders faster than human traders. Because of this advantage, machines dominate the trading within each period by undercutting slow humans at the front of the book. The timing protocol that exemplifies the algorithmic trader's speed advantage in the model is highly complex and mathematically intensive and replicating it goes beyond the scope of our discussion – we would be performing an injustice by trying to synthesise their methodology here. Instead we refer interested readers to Cvitanic and Kirilenko (2010) for a detailed analysis.

Cartea and Penalva (2012) provide an analogous example of Cvitanic and Kirilenko's (2010) (subsumed) speed advantage. These researchers, however, posit that algorithmic traders do not lower volatility; rather, they double trading volume and increase price volatility. Similarly, Foucault, Hombert and Rosu (2016) find that with a speed advantage, the algorithmic trader's order flow is much more volatile, accounts for a much larger proportion of trading volume (relative to other non-algorithmic traders), and forecasts very short-run price changes.

In contrast to the informational assumptions, where algorithmic traders are thought to be imbued with an inherent speed advantage (e.g., Cartea & Penalva, 2012; Cvitanic & Kirilenko 2010; Foucault et al., 2016), others argue that an algorithmic traders' relative informational advantage relates to their ability to more accurately forecast future market variables. Das, Hanson, Kaphart and Tesauro (2001) provide the intuition for this suggestion:

In order to determine whether computer traders can be considered superiorly informed relative to their human counterparts, Das et al., (2001) designed a simulated human versus machine experiment consisting of six challenges. By dividing the simulated population into human traders, fast computer agents and slow computer agents, these authors found that the computerized agents outperformed their human counterparts in all six challenges. Moreover, and perhaps more importantly, their result held for both fast and slow computerized agent populations. Crucially, this seemed to indicate that speed was *not* the sole factor accounting for the agents' edge in performance.

A plethora of subsequent studies show that advances in the capabilities and processes of computers has fundamentally influenced the accuracy of forecasting (Johnson, 2010). This is supported by Easley, Lopez De Prado and O'Hara (2012) who hypothesize that, contrary to popular perception, speed is not the defining characteristic that sets algorithmic trading apart. In their evaluation,

algorithmic trading is distinguished by an ability to make superior strategic decisions via the use of advanced computational techniques and processing power. Consistent with this assumption, Martinez and Rosu (2011) theorise that algorithmic traders have an *accuracy* advantage over their human counterparts.[63] They demonstrate this principle by postulating a theoretical model in which algorithmic traders' are superiorly informed about the fundamental value of the asset.[64] They show that algorithmic traders generate most of the trading activity and volatility in the market. Moreover, the model yields interesting patterns in volatility, volume and liquidity. In particular (according to their respective definitions for volatility, volume and liquidity), Martinez and Rosu (2011) show that a higher precision of information generates more price volatility and trading activity, but only marginally effects market liquidity.[65]

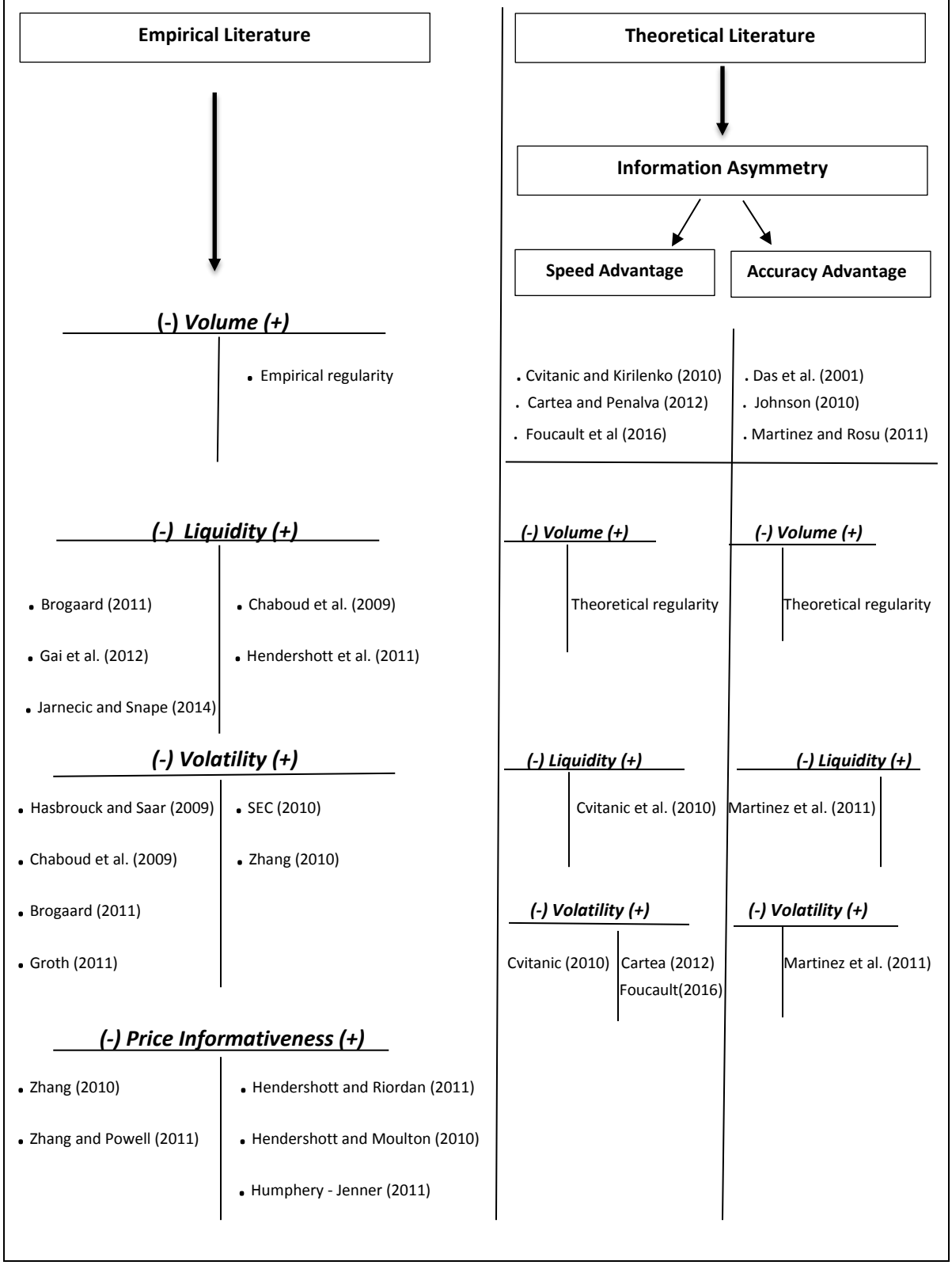
A combined visual summary of empirical and theoretical literature in this field (Table 2) is presented here. As emphasised in the introduction to this chapter, Table 2 applies to later discussions where we categorise and link existing results to those of our own study. Following this, a chapter conclusion and summary will be provided.

[63] This assumption is supported by previous empirical studies on algorithmic trading, which show that algorithmic traders are better informed than other market participants (see, e.g., Hendershott and Riordan, 2010; Brogaard, 2010; Kirilenko et al., 2011).

[64] Similar assumptions have been used in theoretical models by Foucault et al., (2012), and Gamzo (2014).

[65] For specificity, see e.g. Martinez and Rosu, 2011, High frequency traders, news and volatility, Discussion paper, HEC Paris. The model provides an explanation for why trading volume and volatility increase after news announcements.

Table 2: Synthesizing The Diverse Literature on Algorithmic Trading (Gamzo, 2017).



2.7 LITERATURE REVIEW SUMMARY

According to Jain (2005), we are in the midst of a trading revolution. People are being taken out of the direct decision-making process in financial markets around the world,[66] and being replaced by automated trading systems. According to ASIC (2013), an automated order processing program was responsible for at least 99.6% of all trading '*messages*'[67] on the New York Stock Exchange (NYSE) in 2013. Moreover, Johnson, Wang, and Zhang, (2014) estimate that, as of 2012, as much as 78% of all U.S. equity trading *volume* originated from computer algorithms. A precise assessment of its impact on market quality is, however, challenging - a remarkable gap between the results of academic literature can be observed.

Briefly, regarding the nexus between market quality and algorithmic trading, some *empirical* studies have linked the presence of algorithmic trading to increased trading activity, volume and/or improved market liquidity,[68] in both foreign exchange and equity markets (Hendershott, Jones & Menkveld, 2011; Chaboud, Hjalmarsson, Vega & Chiquoine, 2009; Brogaard, 2010, and Hendershott & Riordan, 2011). Other empirical studies assert that algorithmic trading creates an

[66] As reported in Popper (2012), algorithmic trading is also rapidly growing in Europe and Asia, accounting for approximately 45% of stock trading volume in the European Union, 40% in Japan, and 12% in the rest of Asia as of late 2012 (Goldstein, Kumar, & Graves, 2014).

[67] Trading messages are seen to constitute orders, quotes or cancellations.

[68] The bid-ask spread is often used as an indicator of liquidity; whereby, the narrower the spread, the greater the liquidity of a stock. However, as emphasized above, different samples, different markets, and different definitions of liquidity (abound in the literature) can change the findings significantly.

atmosphere of *instability* - increased risk of market crashes - and price volatility (Smith, 2010; Zhang, 2010).[69]

The theoretical literature offers little respite in the way of an accurate assessment of algorithmic trading. Disconcertingly, current theoretical literature treats algorithmic traders as a homogenous trader group, forming a gap between theoretical discourse and empirical evidence on algorithmic trading practices - supporting the multiplicity of algorithmic trading strategies. Therefore, one of the main contributions of this thesis is to bridge that gap by characterizing algorithmic trader subgroups and investigating their respective influence on market quality. This approach allows us to study the effects of mechanisms/processes individually and concurrently (that is, their constituent strategic elements and dynamic interactions), but within the same overarching framework.

A model that incorporates both types of informational advantage is a challenging problem in the context of a microstructure model of trade. Models typically eschew this issue entirely by restricting themselves to the analysis of one type of information variable in isolation. That being said, we must state unequivocally, that models can in theory possess both types of information variables within the same framework. It is hoped that this consolidated model will reconcile competing and existing notions on the practice of algorithmic trading and its impact on market quality.

To the previous caveats of this thesis, another must be added: naturally, some existing models are special cases of our general framework. We suggest however, that in order to understand the extent

[69] Contextually, empirical literature on the topic of algorithmic trading has also not been able to identify (concretely) the existence of a trade-off between liquidity and volatility – a recurring phenomenon in non-algorithmic related empirical research (Chordia et al., 2008; Hanson 2014).

to which the model(s) of this thesis integrate existing models, an in-depth knowledge of this market microstructure theory is required. Given that research within market microstructure is so extensive, a precise assessment of the nexus between existing models and those of this thesis can only be accomplished following a detailed review of market microstructure theory (Chapter 4).

Thus, following the overview of theoretical market microstructure in Chapter 4, a comparative summary of the literature is provided (Table 4.1). Table 4.1 allows the reader to gauge how our model integrates existing models. It also highlights which existing market microstructure models are special cases of our more general framework.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter presents the methodology proposed to investigate the subjects addressed by this thesis. The chapter highlights, and explains the rationale of an appropriate paradigm on which the methods of inquiry rests; provides an instantiation of the concept; and proposes an appropriate scientific model of algorithmic trading.

3.2 RESEARCH PARADIGM AND DESIGN

Algorithmic trading is one of the most controversial and actively discussed topics in the financial world. Understanding the relative impact of algorithmic trading has been the subject of considerable interest in microstructure analysis. On a micro level, algorithmic trading can be seen to have profoundly influenced the price-discovery process - the process by which a market incorporates new information about an asset's value into the asset's price. It follows that elucidating the effects of algorithmic trading on market quality – the typical standard by which to account for the price discovery process – has become vital to our understanding of financial market performance (Gomber et al., 2011).

However, a precise assessment of algorithmic trading is extremely challenging. This is partly a reflection of the complexity of defining the term itself. Invariably, automated trading, flash trading, program trading, low latency trading, black box trading, electronic trading and high frequency trading are just some of the labels ascribed to it in the literature (e.g., Frino, Mollica & Webb, 2014; Gomber et al., 2011; Johnson, 2010; Kissell, 2013; O'Hara, 2014).

Adding to the complexity of the issue it appears that, regardless of terminology, algorithmic trading is a highly diverse activity. Both numerous and diverse trading approaches currently exist. Amongst these are; spread capturing, rebate trading, time weighted average price, volume weighted average price, implementation shortfall, adaptive execution, liquidity detection, data\text mining, neural network and support vector machine strategies.

An alternative method of inquiry that captures the diversity of algorithmic trading practices in a systematic and orderly fashion is suggested here.

Using the principles of contemporary cognitive science, this thesis will argue that the dual-process paradigm – the most prevalent contemporary interpretation of the nature and function of human decision making – lends itself well to a novel taxonomy of algorithmic trading. This taxonomy serves primarily as a heuristic to inform a theoretical model of algorithmic trading, with the view of explaining the evolving nature of market quality as a consequence of this practice.[70] The foundations of this model are located in the market microstructure and behavioural finance literatures. We should emphasise however, that the aforementioned 'agent-based model' methodology will only be addressed in Chapter 4.

In this chapter we highlight only the important assumptions that lie at the heart of our dual process taxonomy of algorithmic trading. To maintain focus on essentials, Chapter 3 then proceeds to discuss the conceptual foundations of our taxonomy, provides an instantiation of the concept, and highlights an appropriate classification scheme on which our taxonomic system rests.

[70] Additionally the taxonomy may enable us to organise both the theoretical and empirical literature on the topic.

3.3 KEY ASSUMPTIONS

Traditionally,[71] in the *market microstructure* literature, a trader can be superiorly informed as a result of either (1) their superior speed in accessing or exploiting information, or (2) their ability to more accurately forecast future variables. To date microstructure models focus on either one aspect but not both (see Section 4.3). This common modelling assumption is also evident in theoretical models of algorithmic trading. Theoretical papers on the topic of algorithmic trading have coalesced around the idea that algorithmic traders possess a comparative informational advantage relative to regular traders. However, the literature is yet to reach consensus as to what this advantage entails, nor its subsequent effects on market quality.

Crucially, we assert that the key assumptions underlying our dual-process taxonomy of algorithmic trading implies that both of these informational advantages underlie algorithmic trading as a whole, with each advantage attributed to a different subgroup of algorithmic trader type. More formally, we suggest that algorithmic trading can be divided into two separate processes or systems; each with its own associated informational advantage.[72] Inherently, all the various strategies associated with algorithmic trading (as presented in Chapter 2) correspond to their own respective 'System' and thus, informational advantage. That is, System 1 algorithmic traders possess an inherent speed advantage and System 2 algorithmic traders, an inherent accuracy advantage. The functional

[71] Refer to Table 4.1, in Chapter 4, for an overview of traditional microstructure models. Table 4.1 organises the theoretical literature on both market microstructure and algorithmic trading vis-a-vis private information.

[72] A model that incorporates both types of informational advantage is a challenging problem in the context of a microstructure model of trade. Again, models typically eschew this issue entirely by restricting themselves to the analysis one type of information variable in isolation. That being said, we must state unequivocally that models can in theory possess both types of information variables within the same framework.

division of algorithmic trading strategies by their respective infrastructure, as seen in Appendix I, seems only to reinforce the above dual-process supposition.[73]

Our model will look at the effects of these two algorithmic trading systems individually and concurrently (that is, their constituent elements and dynamic interactions), but within the same framework. It is hoped that this integrative model – a model that integrates and synthesises the multitude of algorithmic trading strategies in a single workable framework – may shed some light on contradictory findings and on previously unknown market variables.

Next, we discuss the conceptual foundations of our taxonomy, provide an instantiation of the concept, and highlight an appropriate classification scheme on which the taxonomy rests. The presentation of the formal taxonomy is consigned to Chapter 5 (see e.g. Section 5.2).

3.4 ALGORITHMIC TRADING: TAXONOMY DEVELOPMENT

McKelvey (1978) gives a comprehensive coverage of the subject of taxonomies and classification for organizations under the heading of *systematics*. He sees this kind of science as "the science of diversity" (p. 12). The author points out the importance of systematics as a prerequisite to good scientific method, in providing clear delineation of the uniformities of classes of phenomena to be studied. The benefits of classifying objects of interest in a taxonomy is that like (as in similar) properties of a class of phenomena can be identified and a means is provided for comparing and contrasting classes. Following key insights from cognitive science, the taxonomy proposed classifies algorithmic trading with respect to the manner in which they process information.

[73] In Appendix I, it is argued that speed typifies one portion of algorithmic trading strategies – those reliant on co-location infrastructure - whilst accuracy epitomizes the other portion of algorithmic trading strategies –those strategies reliant on high capability infrastructure.

Two types of information processing are distinguished: (1) a type of processing that is fast, reflexive and superficial (2) a type of processing that is controlled, precise and reflective. The distribution of information is a central consideration in market microstructure modelling, and thus appears to be a natural place to start in constructing our model.

The only other characteristic identified in the literature as a possible candidate for distinguishing among classes are the functional attributes of the phenomena itself (by functional attributes we mean the functional architecture of a system and not merely how it functions). There are often deep functional differences between classes or systems. However, functional architecture is only indicative of a systems material/physical attributes and does not necessarily capture the behavioural aspects of that class or system (Gregor, 2006). Consequently, McKelvey, (1978) argues that the functional/architectural aspects of a system should be seen only as a *supportive* mechanism for the primary classifier – information processing in our case.

Thus, like Lieberman (2007, 2009), we use the functional attributes of the phenomena – algorithmic trading – as a secondary classifier for our dual process decomposition. Although of secondary importance to information processing, the architectural aspects of algorithmic trading will be used as a descriptive empirical device that contributes to our modelling methodology (this will become clearer as the research proceeds).We have identified technological infrastructure as a possible architectural/functional candidate for this task. See below.

3.4.1 The Functional Aspects of Algorithmic Trading: Technological Infrastructure

We sought to re-evaluate empirical research on algorithmic trading infrastructure in light of our cognitive dual process construct of information processing. This was done in order to determine the extent to which empirical research supports our proposed theoretical dual process decomposition of algorithmic trading. This empirical review will also have prescriptive relevance to our modelling methodology.

Using the dual process distinction of information processing as a guiding principle, we conducted our investigation into the infrastructural aspects of algorithmic trading (see Appendix I). Evidently, an infrastructural decomposition of algorithmic trading appeared to be analogous with our dual system distinction of algorithmic trading.

As evidenced in appendix I, trading infrastructure is seen as a key determinant in the trading process. In other words, financial markets rely on electronic information and communication technology (ICT) to support the exchange processes. Shapiro and Varian (2013) define ICT as: "the infrastructure that makes it possible to store, search, retrieve, copy, filter, manipulate, view, transmit and receive information" (p.8).

With regards to algorithmic trading, we can distinguish between two independent forms of infrastructure. That is, between co-location infrastructure and high-end capability computing infrastructure (Johnson, 2010; Gomber et al., 2011; Kissell, 2013; Frino et al., 2014 and O'Hara, 2014). Co-location is a latency specific mechanism – trading speed its ultimate condition. High-end capability computing on the other hand is an advanced trading platform that can be measured by its inherent complex, analytical, logical and deductive reasoning capabilities and intrinsic predictive abilities. Recall that from a theoretical perspective, our dual process categorisation of

algorithmic trading implies that one portion of algorithmic traders possesses and inherent speed advantage whiles the other, an inherent accuracy advantage. *It follows that the functional division of algorithmic trading by their respective infrastructure reflects the mechanisms behind these advantages*. In other words, System 1 algorithmic traders speed advantage relates to their access to latency specific mechanisms, e.g. colocation facilities: whilst high-end capability computing subsists as the mechanism by which System 2 algorithmic traders attain their accuracy advantage.

Below we provide an instantiation of System 1 and System 2 algorithmic trading. According to Gregor (2006), an instantiation refers to the creation of a real instance or particular realization of an abstraction or class of objects and processes. Again, this is intended to serve as a descriptive device that contributes to our modelling methodology.

3.5 INSTANTIATION OF SYSTEM 1 AND SYSTEM 2 ALGORITHMIC TRADING

3.5.1 The Instantiation of System 1 Algorithmic Trader's: The Speed Advantage

According to O'Hara (2014), all algorithmic trading is strategic because its goal is to maximize a particular strategy against a market's matching engine (notable examples are given in Chapter 6). Effectively, the matching engine determines how orders are processed, and thus controls the automation of trades and establishes a channel of communication between the market – the matching engine – and the trading algorithm. This is where System 1 algorithmic traders' speed advantage comes to the fore. In practical terms, speed or latency refers to the time it takes to access and respond to market information. It follows, that System 1's speed advantage relates to their faster access to the exchanges matching engine.

System 1 algorithmic traders leverage high speed connections to the market, locating their servers adjacent to exchanges matching engine. Placing one's server adjacent to the exchanges matching

engine means that real-time market information (such as market orders) can reach the algorithmic traders platform almost instantaneously. It therefore significantly reduces the time it takes to access the central order book (where electronic information on quotes to buy and sell as well as current market prices are warehoused). It also decreases the time it takes to transmit trade instructions and execute matched trades. As previously noted, in exchange for a fee, those who subscribe to co-location services get the infrastructure from the exchange itself. The package includes everything from the actual connection to the matching engine, to server cages, electricity, maintenance, and safety installations. The SEC (2010) notes that due to co-location facilities: "the speed of trading has increased to the point that the fastest traders now measure their latencies in microseconds" (p. 3605). 1 microsecond is an IS (International System) unit of time equal to 1000000th of a second. In the past, those on the trading floor had faster access to the market than others; today, those co-located with the exchange market have faster access.

Faster speeds also imply that these traders can trade more frequently, have smaller inventories and shorter holding horizons. The short holding horizon has important implications for our assumptions regarding the System 1 algorithmic trader's behaviour and choice of information. Consider the following: in the short-term, the resale value of a risky security is more likely driven by the order flow rather than the fundamental values. Hence, consistent with the insight of Froot, Scharfstein and Stein (1992), System 1 algorithmic traders in our taxonomy, will most likely focus on short-term order flow. It follows that, when establishing a position, the System 1 trader must have a plan to exit within a short time window. Therefore, the System 1 algorithmic trader's profit is not determined by the difference between his entry price and the fundamental value, but by the difference between his entry and exit prices.

To summarise, we will assume that System 1 algorithmic traders can effectively anticipate incoming market orders and trade rapidly to exploit normal-speed traders' latencies (details provided in Chapter 6). Their speed advantage is predicated on latency specific mechanisms e.g. co-location facilities and applies to information about incoming order flow and not about the fundamental value of the asset.

In order to fully appreciate our instantiation of System 1 algorithmic trading above, some terms require clarification (i.e., the terms public information, private information, information asymmetry and fundamental value). For the sake of brevity, we cannot define these terms effectively in the current section. Those who would prefer clarification at this stage are directed to Chapter 4 (Section 4.3), where we address these terms in detail. The instantiation above provides a sufficient introduction to the theoretic issues that arise later in this thesis.

3.5.2 The Instantiation of System 2 Algorithmic Trader's: The Accuracy Advantage

Crucially, System 1 algorithmic traders' speed advantage relates to information about incoming order flow. This is particularly pertinent given that order book information travels fast within the exchange. However, information on the macro-economy or firm fundamentals – which in general are larger and more complex than order book information – travels slower and is more complex to interpret than order book information. This is where System 2 algorithmic trader's accuracy advantage comes to the fore. In practical terms, accuracy relates to the precision of a long-run, firm specific, forecast. It is predicated on the extent to which an agent is informed about the intrinsic worth of a firm.

The following crucial assumption is thus made in our taxonomy: the System 2 algorithmic trader's accuracy advantage is linked to their ability to predict key, fundamental, value relevant, firm specific variables.

Therefore, unlike our assumptions for System 1 algorithmic trading, System 2 algorithmic trading is not characterized by a speed dimension; but rather an ability to make superior strategic decisions via the use of advanced computational infrastructure and techniques. Essentially, System 2 algorithmic trading would exploit information beyond the traditional order book data. This includes news, pre-news and other forms of textual, as well as numerical, information (Leinweber, 2009). Indeed, major news providers have started offering algorithmic traders access to electronically processable news feeds – providing these traders with valuable and actionable numerical and textual information.

By leveraging high-end capability computing infrastructure, these traders can apply advanced forecasting techniques – such as time series analysis, machine learning, neural networks, support vector machines tools, as well as text mining, in order to make accurate long term predictions of a firm's value. Thus, from a researcher's perspective, System 2 algorithmic trading can be measured by its inherent complex, analytical, logical and deductive reasoning capabilities and intrinsic predictive abilities.

Essentially, an assumption of System 2 algorithmic traders made in our taxonomy is that they exploit their superior ability to interpret public fundamental information in an attempt to make forecasts that are superior to the forecasts of other traders. In other words, these traders filter public fundamental information through an advanced platform (high-capability computer infrastructure), in order to detect private patterns from public information – patterns that signal a firms future performance.

This in turn rests on the principle of information filtration. Effectively, the concept of information filtration is aligned with the assumption that some traders are able to filter public information into private information signals. The terms public information and private information require some clarification. Again, given the inherent complexities involved in defining these terms, we devote an entire section to this task (refer to Chapter 4, Section 4.3). Only a broad overview is presented here.

According to Lyons (2001), at its most basic level, the term public information refers to any publicly available information flows. This includes, but is not limited to; earnings announcements, management and analysts' forecasts, and other summaries of detailed financial accounting statistics. Again, as the term suggests, this information is public knowledge and is available to all market participants. Conversely, private information can be defined as *any information that is not known by all trading parties* (Brunnermeier, 2001). Although this is the most natural definition, it is also extremely general. For example, under this definition, if a trader has any privileged information about a firms performance, and that information helps him or her in forecasting prices then this constitutes private information. However, a precise categorisation on what constitutes private information does not concern us at this stage. What matters here, is the assumption that some traders are able to utilise public information in a manner that reflects private valuations of a firms performance. This suggests that certain traders make judgments about a firm's performance that are superior to the judgments of other traders (Lyons, 2001).

The aforementioned serves to clarify the assumption made for System 2 algorithmic traders in our taxonomy *(i.e. that these traders filter public information through an advanced platform such as high-capability computer infrastructure, in order to detect private patterns from public information – patterns that signal a firm's future performance more accurately).*

CHAPTER 4

MICROSTRUCTURE BACKGROUND

4.1 INTRODUCTION

In Chapter 6 we present the first ‘all-inclusive’ theoretical model of algorithmic trading (all-inclusive in that it accounts for the multitude of algorithmic trading strategies within a single theoretical framework.)

As the model has microstructural foundations, market microstructure theory is reviewed in this chapter. It must, however, be pointed out from the outset that the work within market microstructure is so voluminous that in a single chapter, there is simply not enough room to present the main models in depth. Instead, our objective can be accomplished by presenting a coherent picture, or road map that readers can use for navigating further sections. For ease of reference, we present each model with its most valuable insights. Additionally, we present simple (less generalised) versions of each model in order to communicate the underlying economics as efficiently as possible. Clarity on the underlying economics is crucial for understanding the later sections applications.

The current chapter will begin by highlighting the important assumptions that underlie market microstructure theory. The terms private, public and fundamental information are then defined and discussed. For conceptual clarity, we present a few models/examples that speak directly to these definitions; however, briefly.

This chapter then proceeds to survey two influential models of market microstructure: 1) the rational expectations auction model and 2) the Kyle auction model.

4.2 HALLMARKS OF THE MICROSTRUCTURE APPROACH

The previous section introduces market microstructure in the context of trading but does not define the term clearly. Maureen O'Hara (1995, p1) defines market microstructure as "the process and outcomes of exchanging assets under explicit trading rules". We adopt a similar definition. However, because this definition is so broad, it may be helpful to clarify further.

Three important assumptions lie at the heart of the microstructure approach to financial markets. Indeed, Lyons (2001, p4) uses a three-point characterisation of market microstructure in order to differentiate it from traditional approaches (i.e., neoclassical asset pricing models). The focus is information, players and institutions:

1. Information: microstructure models recognise that some information relevant to securities is not publicly available.

2. Players: microstructure models recognise that market participants differ in ways that affect prices.

3. Institutions: microstructure models recognise that trading mechanisms differ in the ways that affect prices.

These three assumptions are the hallmarks of the microstructure approach. In particular, the richness of microstructure models come from (1) information structure: determining what market participants know; (2) heterogeneity: ascertaining what type of market participants are active in the market and identifying their motives for trading; and (3) institutions: examining the roles that each participant plays in the trading process and defining what trading information each trader has available.

The aforementioned hallmarks of the microstructure approach suggest that market microstructure rests resolutely on the dynamics of information i.e., the information economics of financial markets. In keeping with this focus, we move immediately to the economics of financial information; information heterogeneity, public, private and fundamental information are defined and discussed below. Section 4.4 of this chapter considers microstructure models in more detail.

4.2.1 Information Economics

By and large, the interplay between market participants and information has proven to be an important trend in market microstructure research. Apropos market information, a large amount of literature has begun to address issues of information differentials between traders and its subsequent market effects (Grossman, 1976). This information differential, commonly referred to as information asymmetry, arises when information is known to some, but not all market participants. The current approach emphasizes the distinction between informed and uninformed market participants.

Regarding this 'informed/uninformed' investor phenomenon, informed investors are seen to be those investors that have 'private' information about the future states of the world, while the uninformed investors are those that do not (Grossman, 1976). According to Brown, Richardson and Schwager (1987), this superiority or *advantage* is mostly driven by an informed investor's access to *either* more timely information *or* more accurate/precise information on which to base forecasts.[74]

[74] Information asymmetry is closely related to the problem of adverse selection. According to Grossman (1979): "The problem of adverse selection arises as a manifestation of asymmetrical information in any market in which buyers and sellers are not equally informed about the characteristics of the heterogeneous commodities they exchange"(p.336).

4.3 DEFINING PRIVATE INFORMATION

Although there are a host of definitions for private information available in the literature, a common and simplifying thread exists; that is, private information *is information that is not known by all trading parties*. This is a natural definition. We should say, though, that it is a bit broader than some people have in mind. For example, under this definition, if a trader has any privileged information about a firm's performance, and that information helps him or her in forecasting prices then this constitutes private information (Brunnermeier 2001; Lyons, 2001).

In order to add clarity and concreteness to the definition, let us discuss two sub-categories of private information as they appear in the microstructure literature (e.g. Foucault, Hombert, & Rosu, 2012, 2016). Accordingly, informed trading can be seen to can take one of two forms: (i) trading on more accurate information *or* (ii) trading on information faster than other investors. More precisely, private information consists of an accuracy dimension (advantage)[75] or a speed dimension (advantage). The following subsection extends this two category breakdown of private information by adding some granularity. Specifically, we begin by addressing the accuracy component of private information (i) before referencing the timing aspects of private information (ii).

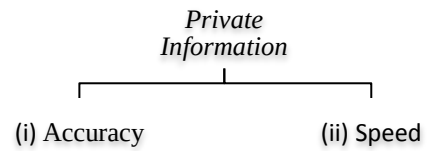


FIGURE 4.1: CATAGORIES OF PRIVATE INFORMATION.

[75] The term advantage is intended to communicate a superior ability relative to others.

(i) The Accuracy Advantage

What does it mean to have more precise/accurate information? We explore the components, and the formation process of information accuracy below.

Some caveats, however, are in order before we continue. Like microstructure theory in general, we view information accuracy as a fundamental component of private information. At its most basic level, the accuracy dimension of private information rests on the principal of information filtration. A single statement is suggestive of the narrative to follow: the accuracy component of private information is aligned on an assumption that some traders filter *public information* into private, and perhaps more accurate signals of a *firm's fundamental value*.

In order to understand this statement, the terms 'fundamental value' and 'public information' require clarification.

The term fundamental value is used to define the long-term or intrinsic value of a security. It is often presented as the discounted present value of a firms future cash flows (Vives, 2010).[76] A variety of other measures have emerged in recent works to account for a firm's fundamental value. We do not discuss these different measures here. We do note however, that *public information* appears to be a key variable in this process – the process of ascertaining a firm's fundamental value.

A comment on the term public information: it is well established that trading on financial markets is strongly influenced by public information (i.e., firm-specific, macroeconomic and other related information flows). The link between public information and a firm's fundamental value is

[76] This is not a proven theorem or validated theory, but a general assumption. Moreover, for different financial instruments, the methodology may differ.

somewhat of a theoretical regularity in the literature (e.g., Kim & Verrecchia, 1994). Markets react sensitively to this so called 'news' which is announced on a recurrent and intermittent basis.

A precise categorisation on what constitutes public information does not concern us at this stage.[77] What matters here, is the assumption that some traders are able to utilise public information in a manner that reflects private valuations of a firm's performance. Indeed, this is a central theme in the discourse on private information accuracy.

We are now in a position to clarify the statement used at the start of this subsection *(i.e., that the accuracy component of private information is aligned on an assumption that some traders filter public information into private, and perhaps more precise valuations of a firm's fundamental value).*

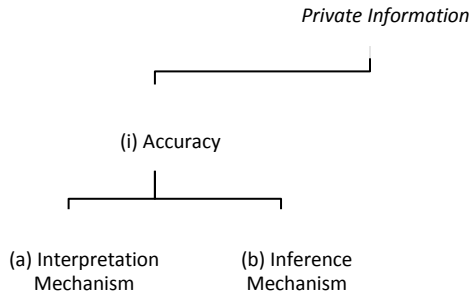
Accordingly, the accuracy advantage (a term used interchangeably with the accuracy dimension) suggests that certain traders make judgments about a firm's performance that are superior to the judgments of other traders. Next, we delineate two possible mechanisms behind this advantage.

Two related mechanisms are referenced in the literature to explain the accuracy phenomenon. The first position holds that, although agents observe the same public information, they *interpret* this public information differently. This idea was first introduced over a century ago by Bachelier (1900) and later developed by Holthausen and Verrechia (1990). The second proposition argues that agents use public announcements to *infer* new private information from the public information itself (e.g. Kim & Verrecchia, 1997). These mechanisms are defined as the interpretation mechanism (a) and the inference mechanism (b) of private information accuracy respectively.

[77] Our characterization of public information is sufficiently broad to include earnings announcements, management and analysts' forecasts, 10-K filings, and other summaries of detailed financial accounting statistics.

Consider a simple diagram (Figure 4.2) that illustrates the relationship between the mechanisms of private information accuracy.

FIGURE 4.2: THE MECHANISMS OF PRIVATE INFORMATION ACCURACY, GAMZO (2017).



a) The Interpretation Mechanism of Private Information Accuracy

The interpretation mechanism suggests that while all investors may observe the same public information, their interpretation of that information for the value of the firm need not be homogeneous. What is relevant is not that each investor sees the same public information for example, but that investors reach varying conclusions about the fundamental value of the firm following this information. This is akin to all agents observing the same reported earnings per share figure, and then determining the implication of the reported earnings for the value of the firm.

Holthausen and Verrechia (1990) validate this differential interpretation principle by postulating a theoretical model in which agents become differentially informed following a public information release. In order to facilitate the discussion to follow, we highlight only the key variables Holthausen and Verrechia (1990) use to model this phenomenon. The model is detailed extensively

in Appendix II (see (a) in Appendix II). For the purpose of continuity in the current section, a broad summary on this model should suffice:

Holthausen and Verrechia (1990) assess the extent to which heterogeneous interpretations of a public information release result in price and volume reactions. In their model, agents exchange a single risky asset and a single riskless asset over one period. The notation v represents the liquidating dividend of the risky asset and, thus, represents the risky asset's true, economic cash flow or *fundamental value*. During the trading period each agent receives a signal s and *interprets* what the signal implies about the liquidating value v, such that each investor's interpretation of the signal is given by:

$$A = v + e + \delta \tag{4.3.1}$$

The variable δ is an 'idiosyncratic'[78] noise term and the variable e is a common noise term. By modelling information with both common and idiosyncratic noise terms,[79] Holthausen and Verrecchia (1990) demonstrate that agents can be differentially informed following a commonly-observed signal. Here, each agent observes the same public signal (e.g. an earnings announcement) but each agent's interpretation of what the signal implies about the value of the liquidating dividend varies because of the agent-specific noise term (idiosyncratic noise term δ). A striking implication of the above is that some agents may have access to more accurate information than others (i.e., have less idiosyncratic noise in their signal). We refer to Appendix II (a) for a comprehensive review of the results of the model.

[78] The term 'idiosyncratic' is used to convey the message that something is specific to a certain individual alone.

[79] Indjejikian (1988) also models information with both a common and idiosyncratic component.

As alluded to above, we wish to avoid making restrictive inferences before exhausting the necessary literature. However, we do note that the intuition for relating the interpretation mechanism (above) to System 2 algorithmic trading appears to be quite reasonable. Hence, there is a strong chance that this informational dimension will feature prominently in our formulation for System 2 algorithmic trading. (Further information to follow in Chapters 4 and 5)

b) The Inference Mechanism of Private Information Accuracy

Holthausen and Verrecchia (1990) model private information by assuming that an earnings announcement provides each agent with a *single* signal ($A = v + e + \delta$). An alternative way to model private information with regards to information accuracy is discussed below. This alternative modelling methodology explicates the role of an agent's ability to infer *new* private information from public information itself. In other words, the inference mechanism of private information accuracy suggests that some agents are able to utilise information gathered in anticipation of public announcement in order to infer new private, and perhaps more accurate information from the information release.[80] Although closely related to the differential interpretation principle, the inference principle suggests that the process of gathering private information is a cumulative process. Kim and Verrecchia's (1994) rational expectations model provides some insight into this concept. However, given that Kim and Verrecchia's (1994) model is explicitly addressed in appendix II (b), the discussion to follow highlights only the informational aspects of the model itself.

[80] The characterization of information releases is sufficiently broad to include earnings announcements, management and analysts' forecasts, 10-K filings, and other summaries of detailed financial accounting statistics. For convenience, however, the focus is primarily on earnings announcements.

In Kim and Verrecchia's (1994) model, agents exchange a single risky stock and a riskless bond. Like Holthausen and Verrecchia (1990) an earnings announcement is set to occur during the trading period. Contextually, the earnings announcement communicates firm value with noise; that is:

$$A = v + \delta \tag{4.3.2}$$

Here, v denotes the liquidating value of the risky asset and thus, represents the risky asset's true, economic cash flow or *fundamental value*.

However, in contrast to Holthausen and Verrecchia's (1990) set-up (where the informed receives a *single* signal concerning the announcement); the informed agents in Kim and Verrecchia (1994) actively collect private information in anticipation of the announcement itself. Namely, before A is announced, the informed trader (denoted i) observes an 'anticipatory' signal in the form of:

$$K = \delta + e \tag{4.3.3}$$

Note that K alone is not an informative signal about the firm's liquidating value v , since both δ and e are independent of v. However, combined with the announced itself $(A = v + \delta)$, K generates a new signal $A - K = v - e$ and this provides more accurate information about the firm's performance.

Institutionally, K can be thought of as the information a trader gleans about a random error in financial reports by studying the firm. For example, in the case of an earnings announcement (characterized by $A = v + \delta$) , the random error δ represents the discrepancy between (v) and the forecast of (v) implicit in current accounting statement $(v + \delta)$.[81] Specifically, traders who

[81] This discrepancy arises, perhaps, from the failure of accounting profits to recognize unrealized gains, use consistent accounting procedures, or avoid questionable levels of capitalization. See Appendix II for a detailed discussion.

possess K are better prepared to translate current accounting figures into superior assessments of the firm. In the context of the model, when earnings are announced, the K can be used to partially correct for this error in the earnings report.

Intuitively, advanced agents would trade in the wake of an earnings announcement not just because of the information contained in the announcement itself, but also because their private anticipatory information (K) leads them to *infer* new private, and possibly more accurate, information from the announcement itself. In fact, this information is often defined as "uniquely privately *inferred* information about future earnings" (Barron, Harris & Stanford, 2005. p.404).We refer to Appendix II for a comprehensive review of Kim and Verrecchia's (1994) model.

Again it is too early to consign this type of information to System 1 or System 2 trading. Nevertheless, the nature of the inference mechanism seems to lend itself well to System 2 information processing.[82] Thus, it will likely feature in our formulation for System 2 algorithmic trading. The details will develop as we proceed to the models themselves.

Summary - the Accuracy Dimension of Private Information

Regarding the dimensions of private information *accuracy*, we have already mentioned that two notions exist to explain the mechanisms behind how some traders come to be more accurately informed. The first is to assume that agents interpret public information differently; and as a consequence some agents have more accurate information than others – this was demonstrated

[82] As noted above (e.g. Chapter 2, Section 2.1.2), with regards to human cognition, System 2 is a high-level processor. The representations in this system are symbolic and unbounded, in that they are based on propositions that can be 'combined' to form larger and more complex sets of propositions. (Sloman, 1996). This 'combination' reference bears remarkable similarity to the explicit features of the inference principle (above), which suggests that the process of gathering private information is a cumulative process (Kim & Verrecchia, 1994). Such insight seems only to reinforce the intuition that the inference mechanism will feature in our formulation for System 2 algorithmic trading.

above in (a). The inference mechanism of private information accuracy, (b), also above, provides an alternative explanation for this information superiority - in terms of accuracy. It explicates the role of an agent's ability to infer *new* private information from public announcement itself.

What follows here is a discussion on the second category of private information; the speed dimension (ii).

(ii) The Speed Advantage

Earlier we alluded to a two-category breakdown of private information. According to this scheme, informed trading can take one of two forms: (i) trading on more accurate information or (ii) trading on information faster than other investors. We have already discussed (i) above. Here we follow with a discussion that draws on (ii), the speed dimension of private information.

It is typical of market microstructure models to assume that all agents receive their information at the same time. While these models have provided many important insights, Hirshleifer, Subrahmanyam, and Titman (1994), argue that in reality some investors, either fortuitously or owing to superior skill, acquire pertinent information before others**.**

Motivated by the above, Hirshleifer et al., (1994) construct the first theoretical account of the speed dimension of private information. The model developed describes the investment choices of a set of investors who investigate the long-term prospects of firms. In order to assess the impact of speed differentials, some investors in the model uncover payoff-relevant information early, while other investors uncover this information later. Overall, the model demonstrates that the exact timing of when investors uncover relevant information may be *as* important, if not *more* important, than the accuracy of the information.

As Hirshleifer et al., (1994, p.1688) note: "since investors who receive information early, trade differently from investors who receive information late, the equilibrium in a securities market where investors receive their information before others can be fundamentally different from the equilibria in models of information acquisition in which investors receive their information simultaneously."

4.3.1 Summary of Private Information

All models are simplifications of reality, and market microstructure models are certainly no exception. One facet of economic models in which simplification has been common is their treatment of information. This simplification is frequently tacit and takes the form of an idealised 'informed/uninformed' investor paradigm. Accordingly, informed investors have 'private information' about future states of the world, while uninformed investors do not possess this information. A review on the literature on private information suggests that informed trading can take one of two forms: (i) trading on more accurate information, or (ii) trading on information faster than other investors. We discussed the various facets of each dimension.[83] However, this was done only to add some granularity to our discussion and ground it in theory. We emphasise here again that a simple appreciation of private information at its broadest level (i.e., the existence of a two-category breakdown of private information) provides sufficient introduction to the basic information-theoretic issues that arise in later sections.

A visual summary of private information is presented in Figure 4.3 below:

[83] For example, along with a description of private information accuracy, we also discussed two possible mechanisms for its formation (i.e., (a) the interpretation mechanism and (b) the inference mechanism of private information accuracy).

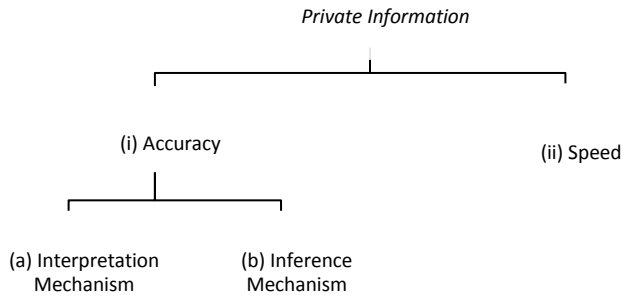


FIGURE 4.3: VISUAL SUMMARY OF PRIVATE INFORMATION IN ITS ENTIRETY, GAMZO (2017).

An important caveat needs to be noted here. In order to provide a coherent picture of private information, a condensation of the literature was necessary. Inevitably, the overview above may have suggested that research on the timing aspects of private information and research on the accuracy dimension of private information have evolved in unison. This is not actually the case. Although the literature does recognise the *existence* of a two category breakdown of private information, these dimensions are rarely, if ever discussed within the same theoretical framework. Indeed, characterising investors' use of both types of private information is a challenging problem in the context of a microstructure model of trade. Again, models typically eschew this issue entirely by restricting themselves to the analysis one type of information variable in isolation. As will become clearer as we proceed, the final model of this thesis reconciles this obvious disconnect by including both types of private information within a single overarching framework. Thus, a particularly novel feature of our model is that it lies at the confluence of these two seemingly disparate literatures.

Below we consider microstructure models in more detail. Following this overview, we close Chapter 4 with a combined visual summary of the theoretical literature on both algorithmic trading and market microstructure (Table 4.1). See Section 4.7.

4.4 MARKET MICROSTRUCTURE MODELS: A FOCUSED OVERVIEW

This subsection provides a focused overview of microstructure models. Given that research within market microstructure is so extensive, some form of literary injustice is inevitable when providing only an overview. However, the primary aim of this section will be accomplished by presenting a coherent framework that readers can use for navigating further sections.

Two distinct models are discussed in this section:

1. Rational expectations auction model

2. Kyle auction model

The rational expectations model of securities trading is a natural starting point for this section. Opening with this model clarifies how its shortcomings spurred the development of later microstructure models. Among these shortcomings is that the act of setting prices is not addressed in the model; there are no agents whose responsibility it is to set prices. When pressed about where these prices actually come from, researchers typically refer to a 'Walrasian' auctioneer, a hypothetical agent who collects orders, sets prices based on these orders and executes trades at the market-clearing price he or she sets (Grossman & Stiglitz, 1980). Interestingly, these hypothetical agents exist outside of the model itself.

The Kyle (1985c) model (henceforth the Kyle 1985 model) is a natural follow-up to the rational expectations model. Both models are equilibrium models in the sense that they derive relationships between the prices of traded assets based on some strong assumptions about investors behaviour

in the market (Cvitanić & Zapatero, 2004). The key difference is that the Kyle model addresses the act of setting price explicitly. This is achieved by introducing an actual auctioneer (henceforth, 'market maker')[84] to replace the hypothetical auctioneer of the rational expectations model. Kyle's market maker uses order flow information to determine the market clearing price. Also, because the protocol that governs trading is fully specified in the Kyle model (in contrast to the rational expectations model) his model produces an intimate link between trading protocol and price determination – a hallmark of microstructure modelling.

4.5 THE RATIONAL EXPECTATIONS MODEL

For contextual perspective, it may be helpful to distinguish between two types of rational expectations equilibria here.

These equilibrium types are

1. Fully revealing equilibrium

2. Partially revealing equilibrium

In a fully revealing equilibrium, all information, including private information, is embedded in price (so-called strong-form efficiency). More precisely, the price provides a sufficient statistic for the underlying fundamental, making private information redundant in the context of trading. Conversely, in a partially-revealing equilibrium, price reflects a combination of private information, as well as, extraneous noise.

[84] The Kyle model is not classified as dealership market because dealer quotes do not represent the best available price. The Kyle auctioneer (a term used interchangeably with market-maker) therefore does not share a characteristic that is true of dealers – that they first provide other individuals with a quote, and then orders are submitted conditional on that quote. Rather, orders are submitted to the Kyle auctioneer (market-maker) before price is determined, and the Kyle auctioneer then determines a market-clearing price based on those orders (Lyons, 2001, p.65).

Although fully revealing equilibria were the focus in early rational expectations models, the literature is beginning to favour the partially revealing approach. Partially revealing equilibria are also closer to the reality of trading. To produce a partially revealing equilibrium, we add some sources of noise to the trading process to make it difficult to disentangle the causes of price movements.[85]

The Grossman and Stiglitz (1980) model is widely referenced in the literature and is a natural place to begin our rational expectations review. Given that it includes private information and noise in asset supply, the model provides insight into the partially revealing equilibria discussed above. Model specifics are provided below.

4.5.1 The Model

In the particular one-period version of Grossman-Stiglitz (1980) rational expectations auction model, one risky asset is exchanged for a riskless asset among two kinds of traders: an informed trader and an uninformed trader. Both traders are risk averse and non-strategic (non-strategic meaning that that they act as perfect competitors and take market prices as given)[86]

The value of the risky assets end of period payoff is denoted as v. It is a normally distributed random variable with mean zero and variance σ_v^2. Before v is paid or observed, the risky asset is traded at price p.

Before trading at price p, one of the traders is 'informed', meaning that he receives private information about v in the form of signal s. Although the informed trader is the only one to receive

[85] For example, when noise relating to an assets supply is added to a model with private information, price rises do not indicate whether the cause is more positive private information or a smaller asset supply. This is because both would elevate the price.

[86] This definition of non-strategic behavior is Lyons (2001).

the signal s, both traders know that s is normally distributed, with mean v and variance σ_s^2. This signal is specified as

$$s = v + e,$$

e represents the noise in the signal.

The other trader is 'uninformed', meaning that he has not observed s and therefore does not have an information motive for trading. Specifically, Grossman and Stiglitz (1980) assume that uninformed traders are non-strategic and trade for motives other that information (such as hedging).[87]

Each trader initially receives a random endowment in units of the risky asset (in shares or currency) denoted X_I and X_U (I and U denote the informed and uninformed trader, respectively). Each endowment is normally distributed with mean zero and variance σ_X^2. We denote the aggregate supply of the risky asset by X, where $X = X_I + X_U$. (If $X_I > X_U$, we would expect the informed trader to be selling to the uninformed trader, ceteris paribus). X_I and X_U are distributed independently of one another and independently of the signal s and the payoff v.

The exponential utility function used here is common in microstructure theory (O'Hara, 1995). It is defined over end-of-period dollar wealth W:[88]

$$U(W) = -\exp(-W). \qquad (4.5.1)$$

[87] Uninformed traders are often recorded as noise traders in the literature. They are liquidity motivated, smoothing their inter-temporal consumption stream through portfolio adjustment. Regardless of how one chooses to define uninformed traders, one can safely assume that they have a non-informational motive for trading. Bagehot (1971) provides a detailed discussion on the topic.

[88] In this framework, random variables are normally distributed and utility exhibits constant absolute risk aversion (CARA). Equation 4.5.1 introduces the negative exponential utility function, a concept that is standard in many models of trade. See O'Hara (1995) for more details - in particular the appendix to Chapter 3.

The properties of the utility function are as follows: first, the reallocation of wealth in the trading process does not affect equilibrium. This is because the risky asset demands do not depend on wealth (Lyons, 2001). Second, when coupled with the normally distributed assumption underlying returns, the exponential utility function produces a demand function for the risky asset that takes a simple linear form.

A pricing rule rests at the centre of rational expectations models. This is because the pricing rule determines how the model's random variables determine equilibrium price. The functional form of the pricing rule is common knowledge to all traders. Knowledge of the pricing rule together with the market price implies that the uninformed trader is able to glean some of the information about the informed traders signal.

The pricing rule must meet two conditions: a rational expectations component, and an equilibrium component. The two conditions for rational expectations equilibrium are as follows:

1) Expectations of the payoff v are consistent with the equilibrium pricing rule

2) Markets always clear (excess demand equals zero for all random variable realisations.)

The solution to the equilibrium is given below. Here, we show that the proposed equilibrium conforms to the rational expectations conditions.

4.5.2 Solving for Equilibrium

Grossman and Stiglitz (1980) solve for this type of model 'by construction', that is, by conjecturing a pricing rule and then verifying that it meets the two conditions above.[89] The basis for proposing a linear pricing rule rests on the assumptions concerning utility and normal distributions. However,

[89] This is also known as the method of undetermined coefficients.

it is important to note that it is not *a priori* clear what the content of that linear rule should be (or that an equilibrium rule need even be linear). According to Lyons (2001), this is a matter of trial and error. Although not ostensibly obvious, Grossman and Stiglitz (1980) show that the following conjectured rule does in fact meet the two equilibrium conditions:

$$p = \alpha s - \beta X \tag{4.5.2}$$

We see that s, the realized signal and X, the realized risky asset supply, are the key components in the equation above. These are natural choices for the proposed rule: they are the random variables on which asset demands are based. The remaining random variable v is not a candidate because the payoff v is not observable at the time of trading. α and β are constants and their values are determined at the end of the solution process in a manner that makes them consistent with the optimising behaviour of both traders.

There are three more steps involved in solving for equilibrium. First we need expressions for each trader's expectation of the payoff v; these must be consistent with with the equilibrium pricing rule. Second, based on these expectations of of v from step one, we need expressions for each trader's risky-asset demand. Finally, we use those demands to find a market-clearing price that matches the proposed pricing rule in equation (4.5.2). Then, we have the rational expectations equilibrium, because it conforms to the second equilibrium condition (2) above. In that equilibrium, expectations are formed using the correct pricing rule, conforming to condition (1).

4.5.3 Expectations

It is relatively straightforward to demonstrate traders' expectations in this setting. This holds true especially in the informed trader's case given that he learns only from his own signal – he knows the other trader is uninformed. Grossman and Stiglitz (1980) show that we can write the informed

traders posterior beliefs about the payoff v conditional on his signal s as normally distributed with:[90]

$$E\,[v|s] = \left(\frac{\sigma_s^{-2}}{\sigma_s^{-2}+\sigma_v^{-2}}\right)s \quad \text{and} \quad Var\,[v|s] = \left(\frac{1}{\sigma_s^{-2}+\sigma_v^{-2}}\right).$$

These expressions make intuitive sense. As σ_s^2 – the variance of the signal s about v – goes to infinity (a weaker signal), $E\,[v|s]$ goes to the unconditional expectation of v, or $E[v] = 0$, and $Var(v|s)$ goes to the unconditional variance of v, or σ_v^2. These are the unconditional mean and variance of v. As σ_s^2 goes to zero (a stronger signal), $E\,[v|s]$ goes to s and $Var(v|s)$ goes to zero.

The uninformed trader does not have any private signal. Therefore, by definition, his expectation is not based on private information. Rather the uninformed trader gathers all his information from price, which, in equilibrium, has imbedded in it, information from the informed trader's trades. What the uninformed trader would like to know is the additional information afforded to the informed trader, the signal s. To utilize price to make inferences about the informed trader's signal s, the uninformed trader can use the proposed pricing rule – the parameters α and β – to transform

[90] Kyle's (1985) model introduces many of the modelling concepts utilized in our own study. Insofar as the Grossman and Stiglitz (1980) model is concerned the review here functions only to highlight key intuition. Naturally, this obviates explicit exposition of their results. However, for the sake of comprehensiveness, we illustrate broadly, the procedure used to derive the above results. Herewith, note the following: With normally distributed random variables, conditional expectations are normally distributed and take a convenient form. Specifically, define the following:

$$y = \bar{y} + \varepsilon_0$$
$$x_i = y + \varepsilon_i, \qquad i = 1, \dots, n,$$

Where the variable y is the variable of interest and the variable x_i are signals of y. Let each $\varepsilon_i,\ i = 0, \dots, n,$ be independently and normally distributued $N(0, \sigma_i^2)$. Then $E[y\,|x_1, \dots, x_n] = \frac{\bar{y}\sigma_0^{-2}+x_1\sigma_1^{-2}+\cdots+x_n\sigma_n^{-2}}{\sigma_0^{-2}+\sigma_1^{-2}+\cdots+\sigma_n^{-2}}$ and $V[y\,|x_1, \dots, x_n] = \frac{1}{\sigma_0^{-2}+\sigma_1^{-2}+\cdots+\sigma_n^{-2}}$, and these two moments fully charecterise the conditional expectation because the conditional distribution is normal. Consider the simplicity: The mean of the posterior distribution is the sum of the prior and signals, each weighted by its own precision (the inverse of its own variance), divided by the sum of the precisions. The conditional variance is just one over the sum of the precisions.

price p into information that is distributed about s. Specifically, using the pricing rule $p = \alpha s - \beta X$, the uninformed trader divides the price p that he observes by α, yielding:

$$\frac{p}{\alpha} = s - \left(\frac{\beta}{\alpha}\right) X.$$

This variable $\frac{p}{\alpha}$ is distributed around a mean of s, which is the information the uninformed trader wants to ascertain. For notational convenience, we will use K to denote this information:

$$K = \frac{p}{\alpha} = s - \left(\frac{\beta}{\alpha}\right) X.$$

Because $s \sim N(v, \sigma_s^2)$, $X \sim N(0, 2\sigma_X^2)$, and s and X are independent, we can summise that K is normally distributed about v with variance $\sigma_k^2 = \sigma_s^2 + 2\left(\frac{\beta}{\alpha}\right)^2 \sigma_X^2$. ("$\sim$" means distributed). With this value for σ_K^2, the uninformed trader's posterior distribution is normally distributed about v with

$$E\,[v|p, \alpha, \beta] = \left(\frac{\sigma_k^{-2}}{\sigma_k^{-2} + \sigma_v^{-2}}\right) K$$

and,

$$Var\,[v|p, \alpha, \beta] = \left(\frac{1}{\sigma_k^{-2} + \sigma_v^{-2}}\right).$$

Knowledge of α and β – the coeficients of the pricing rule – is crucial to the uninformed trader's inference. Because both traders are assumed to be non-strategic (i.e., take prices as given), neither trader conditions on the realization of his own endowment – X_I and X_U, respectively. We discuss the assumption of non-strategic behaviour in more detail below.

4.5.4 Demand

From the analysis above, we know that both traders' posterior distributions are normal. Thus, we can express each trader's demand in a simple fashion. Given that expected returns conditional on available information are still normally distributed, and our exponential utility specification, the demand functions for the informed trader D^I and the uninformed trader D^U – in units of the risky asset, e.g., shares or currency – take the following form:[91]

$$D^I = \frac{E\ [v|s]\ - p}{Var(v|s)}$$

$$D^U = \frac{E\ [v|p,\alpha,\beta]\ - p}{Var(v|p,\alpha,\beta)} \tag{4.5.3}$$

Note in equation (4.5.3) the information role that price plays in the demand of the uninformed trader (it enters in the conditional expectation and conditional variance). Inserting the values for $E\ [v|s]$ and $Var\ [v|s]$ into this expression for D^I and D^U results in equation (4.5.4):

$$D^I = (\sigma_s^{-2})s - (\sigma_s^{-2} + \sigma_v^{-2})p$$

$$D^U = (\sigma_K^{-2})K - (\sigma_K^{-2} + \sigma_v^{-2})p \tag{4.5.4}$$

[91] With respect to the Grossman and Stiglitz (1980), we consider only the key intuition underlying their model. Given that our own study is based on Kyle (1985), the explicit determination of equation (4.5.3) falls beyond the scope of this thesis.We refer interested readers to Lyons (2001), in particular the Appendix to Chapter 4, where he explicitly derives these demand functions.

4.5.5 Market-Clearing Price

The market-clearing price is determined by equating demand with supply so that no excess demand exists (i.e., excess demand equals zero):

$$D^I + D^U = X$$

Inserting our expression from equation (4.5.4) for D^I and D^U in this market-clearing condition yields a price of

$$p = \alpha s - \beta X, \tag{4.5.5}$$

where

$$\alpha = \left(\frac{\sigma_K^{-2} + \sigma_s^{-2}}{\sigma_K^{-2} + \sigma_s^{-2} + 2\sigma_v^{-2}} \right)$$

$$\beta = \left(\frac{1}{\sigma_K^{-2}(1 - \alpha^{-1}) + \sigma_s^{-2} + 2\sigma_v^{-2}} \right)$$

Recall that $\sigma_K^2 = \sigma_s^2 + 2\left(\frac{\beta}{\alpha}\right)^2 \sigma_X^2$. These values for α and β ensure that excess demand equals zero for all random variable realizations, which fulfils condition (2) above for rational expectations equilibrium. To fulfil equilibrium condition (1) above, we imposed in our derivation of these coefficient values, that the pricing rule used to form expectations is the actual rule used to determine price. Thus, using the method of undetermined coefficients, we have verified what we

set out to verify: that the conjectured pricing rule in equation (4.5.2) describes a rational expectations equilibrium.

This equilibrium is partially revealing, a fact taken out of the uninformed trader's expectation. Specifically, the uninformed trader does not know as much in equilibrium as the informed trader, as shown by the distribution of posterior expectations. Recall that the variance of the informed trader's posterior expectation is

$$Var\,[v|s] = \left(\frac{1}{\sigma_s^{-2} + \sigma_v^{-2}}\right),$$

and the variance of the uninformed trader's posterior expectation is

$$Var\,[v|p, \alpha, \beta] = \left(\frac{1}{\sigma_k^{-2} + \sigma_v^{-2}}\right)$$

The only difference is the replacement of σ_s^2 with σ_K^2, where σ_K^2 has a value of $\sigma_s^2 + 2\left(\frac{\beta}{\alpha}\right)^2 \sigma_X^2$. Because $2\left(\frac{\beta}{\alpha}\right)^2 \sigma_X^2$ must be positive, σ_K^2 must be larger than σ_s^2, so the variance of the uninformed trader's posterior expectation is larger.

4.5.6 Limitations of the Grossman-Stiglitz Model

We close this description of the model by discussing some drawbacks of Grossman and Stiglitz (1980) model – and rational expectations models in general.

4.5.6.1. The Implicit Auctioneer

The Grossman-Stiglitz (1980) model is not explicit about who sets prices. Like many other rational expectations models, the fiction of an implicit, Walrasian auctioneer is suggested. The Walrasian auctioneer is the traditional way to envision how prices are actually set in rational expectations models. The implicit auctioneer collects the 'preliminary orders' (Lyons, 2001), and uses them to find the model's market-clearing price. Without this narrative in the background, there is no way to understand how price is actually determined in real time. Strictly speaking, the model only requires that price clears the market and is consistent with expectations.[92]

4.5.6.2. Nonstrategic Behaviour

The demand function in equation (4.5.3) suggests that the informed trader takes the current price as given. The informed trader does not exploit the fact that his trade has a direct effect on that price (however, by definition, expectations are validated in equilibrium). The effect that individuals' trades have on price is not negligible, but these traders behave like perfect competitors (price takers) nonetheless.

[92] The Grossman and Stiglitz (1980) version of the model has a finite number of traders $i = i, \dots, n$, so the assumption of perfectly competitive behaviour is less problematic than in the two trader case presented here (but still a bit of a stretch). One commonly used technique to avoid this issue is to assume that the informed and uninformed represent separate continuums of traders, so that no single trader has a measurable impact on price. See Lyons (2001, p 288).

4.5.6.3. Knowledge of Pricing Rule

We can see from the model that the uninformed trader needs to condition his demand on the pricing rule – the variables α and β. However, practically speaking, one cannot determine how the uninformed trader acquires such knowledge as it remains unspecified in the model.

The drawbacks discussed above are addressed in the model that follows. Regarding the first count, Kyle's (1985) model addresses the act of setting price explicitly. The second and third drawbacks are also largely assuaged by the Kyle model. For example, it incorporates strategic behaviour. Moreover, the model does not rely on a pricing rule conjecture. Instead, an explicit market maker is introduced; who optimises his behaviour subject to the constraints imposed by the model. This will become evident as the chapter proceeds.

4.6 THE KYLE AUCTION MODEL

The Kyle (1985) model and the rational expectations models are closely related. Both have an auction market structure, and at the centre of both is a pricing rule that is consistent with expectations. The key conceptual difference is that the Kyle model includes an explicit auctioneer (market maker) rather than an implicit one. This changes the nature of the pricing rule because the act of price setting is now assigned to an agent within the model (a discussion on this will follow).

Notwithstanding the key conceptual contribution of an explicit auctioneer, the Kyle model generates many other novel insights – insights relative to the rational expectations model.

Before we formally introduce the model, some of these insights merit a discussion.

Insights

Three of the most important insights generated include:

1. Market makers earn zero profits. This assumption is consistent with free entry of competing market makers, a condition under which the single market maker cannot exercise monopoly power. (This zero profit condition is important to the model and is shared by many other models within microstructure.) Here, market makers are simply vote counters, not analysts of fundamentals; and the votes they count are the order flows.

2. Market makers are unable to differentiate between informed and uninformed orders. Informed traders can then use this to great advantage.

3. Market efficiency, market liquidity and price discovery are deeply related.

Though the model itself will enrich each of these insights, let us offer a few thoughts on insight (3).

The notion that prices in financial markets reflect and convey information is axiomatic within the realm of finance and is enshrined in the theory of market efficiency (Fama, 1970).[93] Hayek (1945) wrote eloquently on this function of prices in sharing information among economic participants. He described prices as a form of communication, enabling all decision makers to alter their behaviour due to the essential information contained in one measure, *price*.

However, when economists speak of a capital markets being efficient, they are usually referring to the extent to which prices reflect available information (Alagidede & Panagiotidis, 2009). If we consider that asset prices are determined by the outcome of supply and demand, then in a

[93] The concept of market efficiency emanates from Fama's (1970) academic doctrine of the Efficient Markets Hypothesis.

competitive market, traders are expected to rapidly assimilate any new information that is relevant to the determination of asset prices, and prices should adjusting accordingly (Alagidede, 2008).

The way in which prices adjust to this new information (or *discover* this new information) is aptly named the price *discovery* process.

The literature suggests that market liquidity is crucial to the price discovery process. Concomitantly, market liquidity can also be considered crucial for stock market efficiency (Lyons, 2001). Accordingly, a market is liquid in the sense that almost any amount of stock can be bought and sold immediately, and a market is efficient in the sense that small amounts of stock can always be bought and sold very near the current market price, and in the case of large amounts, can be bought or sold over long periods of time at prices that are (on average), very near the current market price (Black, 1971).

By showing the relation between liquidity, price discovery and market efficiency, Kyle's model has a fascinating message. The basic idea is that in efficient markets there are forces pushing to keep liquidity from moving predictably over time. More precisely, Kyle is able to determine that predictable variation in liquidity in an efficient market should not be 'large enough' to generate excessive risk-adjusted returns.

Although the above discussion serves well as an introductory note, we may have 'jumped the gun' somewhat by presenting preliminary results before presenting the model itself. Readers who find the above description more challenging should consider that this insight is thoroughly addressed at the end of the review below.

4.6.1 Model Overview

As clarified below, the Kyle (1985) model is a natural extension of the Grossman and Stiglitz (1980) model in that it addresses the act of setting price explicitly. This is achieved by introducing an actual market maker to replace the hypothetical auctioneer of the rational expectations model. Kyle's market maker uses order flow information to determine the market clearing price. Also, because the protocol that governs trading is fully specified in the Kyle model (in contrast to the rational expectations model) this model produces an intimate link between trading protocol and price determination – a hallmark of microstructure modelling.

The second key extension made by the Kyle model relative to its rational expectations counterpart, is that Kyle's model follows a more game-theoretic set-up.[94] In this respect, Kyle's equilibrium can be characterised as a Bayesian Nash Equilibrium (BNE).

The aforementioned 'Bayesian' reference emanates from game-theoretic nomenclature and has been used to indicate market maker's use of Bayes' rule to update their beliefs (Bayes' updating model is formally addressed in the overview of the model itself – below).

Before we define the constituent elements of a Bayesian Nash Equilibrium (BNE), consider the following: specifically, the Bayesian Nash Equilibrium concept is the second of two competing equilibrium concepts addressed in this theoretical overview – namely; the Grossman and Stiglitz (1980) Rational Expectations Equilibrium concept and Kyle's (1985) Bayesian Nash Equilibrium concept respectively.

[94] Essentially, Kyle's model extends the Grossman and Stiglitz (1980) model in two ways: (1) abandons the implicit Walrasian auctioneer and (2) introduces a game-theoretic set-up.

We expand briefly on the key conceptual differences between these two types of equilibrium below. This should provide the premise for the discussion to follow, where we explore Kyle's specific Bayesian Nash Equilibrium model in more detail.

Consider the following exert from Brunnermeier (2001), highlighting the conceptual difference between a Rational Expectations Equilibrium (REE) and a Bayesian Nash Equilibrium (BNE):

> In a REE, all traders behave competitively, that is, they are price takers. They take the price correspondence, a mapping from the information sets of all traders into the price space as given. In a BNE, agents take the strategies of all other players, and not the equilibrium price correspondence, as given (p.14).

Accordingly, the game-theoretic BNE concept provides a useful theoretical framework from which to analyse the strategic interactions that occur in a market in which traders take their price impact into consideration. Precisely, in the context of Kyle's BNE, the informed trader takes into account the effect his orders have on the resulting price. This contrasts the preceding REE model of Grossman and Stiglitz (1980), where the informed trader *does not* take his price impact into account (we can see from equation 4.5.3 of the previous model that the informed trader takes the market price as given, and thus does not consider the effect that his demand has on equilibrium price).

Kyle (1985) obtains a perfect Bayesian Nash Equilibrium by considering (1) that market-makers set prices based on their Bayesian interpretation of the information they obtain (expanded on below), and (2) that the informed trader chooses a demand function that maximises expected profits given his expectation of the impact of his order on the market price (this is also explicitly addressed with the respective notation below)

Kyle's BNE is probably best explained by illustrating the steps needed to derive the corresponding equilibrium. It should, however, be pointed out from the outset that the work within game theory in general (and BNE in particular) is so extensive that in a single section, there is simply not enough room present these steps in detail. Moreover, given that these precepts are adequately addressed in later chapters (namely, by the specific models of this thesis), a more comprehensive exposition of the topic cannot be justified here – no additional utility can be gained by providing a more detailed delineation of the steps involved in the derivation of a BNE. Instead, the primary goal of the proceeding discussion can be accomplished by presenting a coherent overview or roadmap that communicates the underlying economics of the model as effectively as possible. As noted, clarity on the underlying economics is crucial for understanding the later sections applications.

For the sake of descriptive convenience, the manual to follow (Box 4.6) presents a synoptic overview of the steps involved in the derivation of a BNE and is based on the work of Brunnermeier (2001), Fudenberg and Tirole (1991) and Osborne and Rubinstein (1994).

BOX 4.6: OVERVIEW OF THE STEPS INVOLVED IN THE DERIVATION OF A BAYESIAN NASH EQUILIBRIUM. SOURCE: BRUNNERMEIER (2001), FUDENBERG AND TIROLE (1991), OSBORNE AND RUBINSTEIN (1994).

Before extending our analysis to Kyle's (1985) seminal BNE model, we illustrates (broadly) the steps involved in the derivation of a BNE.

BOX 4.6: Overview of the steps involved in the derivation of a BNE

Introduction:

A Bayesian Nash Equilibrium is a term used in game theory to describe an equilibrium where each player's strategy is optimal given the strategies of all other players. This game-theoretic approach allows one to model the strategic interactions that emerge between traders when they take their price impact into account. Here, agents take the strategies of all other players, and not the equilibrium price correspondence, as given.[95] Concisely, a BNE is formed by a profile of strategies of all players from which no single player wants to deviate.

Steps in the derivation of equilibrium

Step 1: Specify the players' prior beliefs and conjecture a strategy profile for each player. More specifically, propose a whole set of profiles described either by a profile of general

[95] In a REE, all traders take the price correspondence as given. Brunnermeier (2001) asserts that the term price correspondence refers to "a mapping from the information sets of all traders into the price space" (p.14). Accordingly, a fundamental difference between a BNE and a REE is that in a BNE, agents take the strategies of all other players, and not the equilibrium price correspondence, as given. Herewith, by assuming that the strategies of all the other players are given, a player can choose his optimal strategy.

functions or by undetermined coefficients.[96]These profiles also determine the joint probability distributions between players' prior beliefs, their information, and other endogenous variables like other traders' actions, demand, and prices.

Step 2: Update all players' beliefs using Bayes' rule and the joint probability distribution, which depends on the proposed set of strategy profiles, for example the undetermined coefficients.

Step 3: Derive each individual player's optimal response given the conjectured strategies of all other players and the market clearing conditions.

Step 4: If the best responses of all players coincide with the conjectured strategy profile, nobody will want to deviate. Hence, the conjectured strategy profile is a BNE. In other words, the BNE is a fixed point in strategy profiles. If one focuses only on equilibria in linear strategies, the proposed set of strategy profiles can be best characterized by undetermined coefficients. Each player's best response depends on the coefficients in the conjectured strategy profile. The BNE is then derived by equating the conjectured coefficients with the ones from the best response.

Kyle's specific Bayesian Nash Equilibrium (BNE) is explored in more detail below.

[96] The "method of undetermined coefficients" (sometimes referred to as the method of *judicious guessing*) is a systematic way to determine the general form/type of the particular solution, say, $Y(t)$ based on the non-homogeneous term, say, $g(t)$ in a given equation. The basic idea is that many of the most familiar and commonly encountered functions have derivatives that vary little (in the form/type of function) from their parent functions. Consequently, when those functions appear in $g(t)$, we can predict the type of function that the solution Y would be. The steps are synoptically discussed below: Write down the (best guess) form of Y, leaving the coefficient(s) undetermined. Then compute Y' and Y'', put them into the equation, and solve for the unknown coefficient(s). Note that one can expect the model itself to clarify how this idea works in practice.

A Comment on Notation

Another caveat seems pertinent before we continue to Kyle's (1985) model. As emphasised above, the Grossman and Stiglitz (1980) model falls under the Rational Expectations Equilibrium (REE) framework. Conversely, Kyle (1985) utilises the more game-theoretic, Bayesian Nash Equilibrium (BNE) set-up. Relatively speaking, these two modelling approaches differ quite considerably in terms of notation.

Insofar as our review of the Grossman and Stiglitz (1980) model is concerned, we had to alter some of the models original notation in order to align it with Kyle's. This consideration emerged from the fact that the remaining model(s) of this thesis build on his framework.

For the sake of easy reference, Kyle's explicit notation will be used throughout the remaining chapters of this thesis. Again, this is because our own analysis is aligned with, and builds on Kyle (1985).

4.6.2 The Model

The following 'one-period version'[97] of the Kyle's (1985) auction model is a workhorse within the microstructure literature. The beauty of this model can be seen by its inherent simplicity. As a result, the model has become the benchmark of all microstructure models and has given rise to a whole strand of discourse and commentary that has tested and extended its predictions. Notably, the survey here introduces many of the modelling concepts utilized in our own study.

[97] The preceding Grossman and Stiglitz (1980) model also considered a single period model.

The model has three types of players. Particularly, an informed trader trades against a price setting competitive market maker, in the presence of uninformed traders.[98] Here, traders trade two assets: a risk-free asset (bond) with zero interest rate and a risky asset (stock).[99]

Kyle (1985) begins by assuming that the risky asset to be traded has an *ex-post* payoff of v where the value is drawn according to $v \sim N(p_0, \sigma_v^2)$. As evidenced below, nature selects a true value v from a prior normal distribution $N(p_0, \sigma_v^2)$ for the traded asset. Effectively, the *ex-post* liquidation value of the risky asset is its true/fundamental value and is a normally distributed random variable with initial mean p_0 and variance σ_v^2.

Note here that Kyle assumes that all players are risk neutral. Although not ostensibly intuitive, the assumption of risk neutrality means that p_0 also typifies the initial stock price (Chen, 2016). (We suggest that this insight can only be fully appreciated following a detailed exposition of the model itself – to follow.)

[98] Uninformed traders are often recorded as noise traders in the literature. Noise traders are liquidity motivated, smoothing their inter-temporal consumption stream through portfolio adjustment. Regardless of how one chooses to define noise traders, one can safely assume that they trade for reasons exogenous to the given model. Bagehot (1971) provides a detailed discussion on the topic.

[99] The majority of microstructure models include a risk-free asset (e.g. a riskless bond). Because risky asset demand is independent of wealth, in order for each individual's demand to be feasible, then each individual must be able to borrow and lend at the risk-free interest rate without constraint (Lyons, 2001). However, its economic role is trivial and some models omit it from exposition. We include it because Kyle includes it.

In Kyle's canonical version of the model, a single (risk neutral) information monopolist is the only one who has private information about the risky asset v, in the form of a signal $s = v + e$.[100] Strictly speaking, the informed trader has some uncertainty over v (in the form of noise e). Nevertheless, he remains an 'informed trader' because he has more information on v than the other traders.

Knowledge of v (in the form of a signal $s = v + e$) forms part of the informed traders private information, its prior distribution (mean and variance) is however, publically available (this will also be clarified as the model description continues).

Noise traders are assumed to trade for purely exogenous reasons. The quantity traded by noise traders, denoted u, is normally distributed with mean zero and variance σ_u^2: $u \sim N(0, \sigma_u^2)$

The random variables v and u are independently distributed. The quantity traded by the informed trader is denoted x and the price is denoted p.

[100] As noted in the introduction to this chapter, the primary aim of this review would have been accomplished by presenting a coherent framework that readers can use for navigating further sections. We stress that clarity on the underlying economics of each model is crucial for understanding the later sections applications. The following must then be noted: In Kyle's original (1985) model, the informed trader has private information about the risky asset, v, in the form of a signal $s = v + e$. However, given the formal objective of this overview, only the simple (less general) version of Kyle's model will be discussed. Case in point, in this chapters' review of the Kyles model, it is assumed that e the error in the informed traders signal is negligible (i.e., $e = 0$). This assumption is made solely for the sake of analytical convenience and has no conceptual implications. With regards to its RE counterpart/predecessor, the rationale for this assumption is simple, and a single statement is suggestive of the narrative to follow: the crucial extension of the Kyle model has nothing to do with the informed investors signal, rather, Kyle's model extends the rational expectations model by abandoning the implicit Walrasian auctioneer (replacing this hypothetical agent with an explicit market maker) and having a game-theoretic set-up. We should remark however, that in order to enhance the comprehensiveness of this thesis, a more general case of Kyles model i.e., where $e > 0$ is presented in Appendix III at the end of this thesis.

4.6.2.1. Conditions for Equilibrium – Formal Definitions

Formally, the Nash equilibrium studied in Kyle (1985) is defined by two functions: the informed traders trading strategy $X(\cdot)$, and the market maker's pricing function $P(\cdot)$. In equilibrium, these two functions have to satisfy two conditions; (1) a profit maximisation condition and (2) a market efficiency condition. In other words, an equilibrium in this model is defined as a pricing rule chosen by the market maker and a trading strategy chosen by the informed trader such that: the informed trader maximizes expected profits, given the market maker's pricing rule; the market maker sets the price to earn zero expected profits,[101] given the trading strategy of the informed trader.

For analytical consistency, Kyle's explicit notation is employed in the explanation below. Specifically, in the Kyle model, $X(\cdot)$ denotes the informed traders strategy, and x, the informed traders trading quantity. Likewise, $P(\cdot)$ denotes the market makers pricing function, and p, the price set by the market maker.

We should explain here that very specific functional form assumptions underlie Kyle's analysis. More precisely, by assuming that the relevant random variables are normally distributed Kyle's model acquires a tractable linear structure. Indeed, under these assumptions (that u and v are independent random variables) Kyle proves that there is a *linear* solution for his strategic game.

[101] It is assumed that market making is a perfectly competitive profession, so that the market maker sets the price p such that, given the total order submitted, his profit at the end of the period is expected to be zero. Again, market makers are simply vote counters, not analysts of fundamentals; and the votes they count are the order flows. The assumption that market makers earn zero profit is important to the model and is shared by many other models within microstructure (Lyons 2001). Crucially, this zero profit assumption satisfies the market efficiency condition described above – condition (2).

In order to provide some contextual perspective, we expand on the topic of linear equilibria below. Kyle's solution technique (the method used to solve for equilibrium) is addressed in what follows.

Linear Equilibrium

Regarding the existence of a linear equilibrium in the Kyle model, a single statement is suggestive of the narrative to follow: Kyle (1985) ostensibly permits X to be a mixed strategy,[102] but then suggests that mixed strategies are not optimal in the equilibrium he describes.[103] Since mixed strategies are not optimal in equilibrium, Kyle (1985) justifies a more intuitive interpretation of X as a measurable function of v, such that $x = X(v)$. Note that the technical aspects of this notation will be refined below - by the model itself.

At the risk of over simplifying matters, one can safely argue that Kyle's aforementioned interpretation of X,[104] effectively rules out nonlinear trading strategies. Indeed, in his proof of Theorem 1, Kyle (1985) writes: "The quadratic objective (implied by a linear pricing rule P) makes linear strategies optimal even when nonlinear strategies are allowed" (p. 1319).

Kyle's Solution Procedure

As noted, Kyle's proof has been premised on the existence of a linear equilibrium. In order to support Kyle's dictum of a linear solution, his model relies on a solution technique known in the

[102] In some situations a player may want to randomise between several actions. If a player chooses which action to play randomly, we say that the player is using a mixed strategy, as opposed to a pure strategy. In a pure strategy the player chooses an action for sure, whereas in a mixed strategy, he chooses a probability distribution over the set of actions available to him or her (Pekar 2008, p.16). Since mixed strategies do not feature in the model we have been succinct in describing the term. For a detailed discussion on the topic of mixed strategies in game theory consult Rasmusen (2001). See also, Brunnermeier (2001), Fudenberg and Tirole (1991), Osborne and Rubenstein (1994) and Vives (2010).

[103] With regards to the discussion above, Kyle only claims to find one equilibrium, which he does, and it does not involve a mixed strategy.

[104] Namely; $x = X(v)$.

literature as 'conjecture and verify' (Brunnermeier, 2001; Fudenberg & Tirole, 1991; Osborne & Rubensteim, 1994; Vives, 2001). Simply put, Kyle 'conjectures' *linear* strategies for both the informed trader and the market maker and only then 'verifies' that these conjectures are actually the best response to one another's strategies.[105]

Specifically, Kyle's standard solution method is based on conjecturing a price strategy for the market maker that is a linear function of aggregate order flow and a trading (demand) strategy for the informed trader that is a linear function of his private information and then verifying that such conjectures are consistent with equilibrium. Because all conjectures will be linear combinations of functions in which the coefficients are 'constants to be determined', this entire approach to finding particular solutions is formally called the method of undetermined coefficients.[106]

The question as to whether there are linear equilibria when traders employ nonlinear strategies is an interesting one. On this topic, Boulatov, Kyle and Livdan (2012) have suggested that the linear equilibrium in the single-period trading model of Kyle (1985) is unique;[107] implying that equilibria with a nonlinear structure cannot actually exist in the Kyle setting.

Boulatov et al., (2012) analyse Kyle's a one-period model where the risk-neutral informed trader can use arbitrary (linear or non-linear) strategies, and the market maker can use arbitrary pricing rules. These authors are able prove that the standard linear informed trader's trading strategy, and correspondingly, the linear pricing rule, lead to the unique equilibrium in the model.

---

[105] Kyle conjectures linear strategies for both the informed trader and the market maker and then verifies that such conjectures are consistent with equilibrium – i.e., the conjectured linear strategies form each player's best response to the other player's strategy. In Kyle's explicit formulation, the term optimal is referenced in order to convey the best response principal. See steps 3 and 4 in Box 4.6, where we highlight the meaning behind best response.

[106] Less formally, it is also called the method of (educated) guess.

[107] See Boulatov, Kyle and Livdan (2012) for a lucid description of what it means for an equilibrium to be 'unique'.

Such validation (of Kyle's unique equilibrium) has significantly simplified subsequent analyses' – analyses' that utilise the Kyle's framework. Consequently, all models based on the Kyle framework conjecture and verify only linear trading and pricing rules respectively.

Kyle's standard linear solution is mathematically rigorous and requires further analysis. Note however, that our model(s), to follow, conform to Kyle's linear equilibrium condition and should be seen as a substitute for explicit proof here (see Chapter 5).

4.6.2.2. Some Intuition for Kyle's Equilibrium

In Kyle's BNE nature moves first by selecting a 'true value'[108] (v) from a prior normal distribution $N(p_0, \sigma_v^2)$ for the traded asset, and by selecting a demand quantity u from a prior normal distribution $N(0, \sigma_u^2)$ for the noise traders.[109]

Since the noise traders' order is exogenous, we need only to consider the optimal actions of the market maker and the informed trader.

[108] In a simple model where the stock pays a one-time liquidation dividend, the true value corresponds to the (random) amount of the liquidation dividend. If the stock can be traded more than once, then *v* may represent the resale value at the next trading date.

[109] Again, the latter (noise traders) may be irrational traders acting on noise rather than information (Black, 1986) or liquidity traders who do not have time discretion i.e., *nondiscretionary liquidity traders* (Admati and Pfleiderer, 1988). In any case, one can assume that noise traders' trade for reasons exogenous to the model. Since *v and u* are independent, the noise traders' trading behaviour, *u*, contains no information regarding the true value (*v*) of the stock.

As discussed above, the single (risk neutral) informed trader is the only one who has private information about the risky asset v, in the form of a single signal $s = v + e$.[110]

In the context of the model the uninformed traders and the informed trader simultaneously submit market orders to the market maker to be executed at a single market-clearing price p. Thus, the submitted orders are of two kinds: the order from the informed trader, x, and orders from the uninformed traders, u. (If u is negative then the uninformed are, on balance, selling.) The informed trader does not observe u before submitting his order x. (Effectively, this precludes the informed trader from conditioning on the market-clearing price, a stark contrast from the rational expectations model, where all trades are conditioned on the market clearing price.)

For setting the price p, the market maker observes *only* the sum of two types of orders, $x + u$. The term y is used to denote the aggregate order flow $x + u$. Here $y = x + u$.

Contextually, conditional on y, the market maker determines a price, p, at which he will clear the entire order. That is,

$$p = P(y)\,, \qquad y = x + u$$

[110] Consider the following narrative: In Kyle's original (1985) model, the informed trader has private information about the risky asset, v, in the form of a signal $s = v + e$. However, given the formal objective of this overview, only a simple (less general) version of Kyle's model will be depicted here. More specifically, in this chapters review of the Kyle model, it is assumed that e, the error in the informed traders signal is negligible (i.e., $e = 0$). A more general case of Kyles model i.e., where $e > 0$ is presented in appendix III. That noted, the structure of the informed trader's private information (e.g., whether or not the informed trader observes v perfectly or with noise) is not crucial. As articulated by Rochet and Vila (1994): "Given that all traders are assumed to be risk-neutral it is only needed that the information structures be nested i.e., that the informed trader knows more than the market" (p.132).

The market maker's pricing rule (the terms rule and strategy are used interchangeably throughout) is pinned down by the assumption that he expects to earn a profit of zero.

This assumption (that market makers earn zero profits) is important to the model and is shared by many other models within microstructure.[111] In Kyle maker market makers are simply vote counters and not analysts of fundamentals (the order flow represents the votes counted by the market maker) Expected profit of zero implies that the market maker sets the equilibrium price p as a function of $x + u$ such that

$$p = E[v|x + u] \tag{4.6.1}$$

Price depends on this sum because the market maker does not observe x and u individually. As noted, the market maker observes *only* the total order flow $x + u$, but not x or u separately.[112] The u component of this sum is determined exogenously, simplifying inference. The complication comes from the x component, which depends on the strategy of the informed trader.

The informed trader in Kyle is both risk neutral and strategic. Strategic trading involves conditioning on the behavior of other agents, both the uninformed and market maker's (a standard game-theoretic definition of strategic behavior).

Specifically, the informed trader takes into account the effect of his orders on price. This contrasts the preceding rational expectations model (as evidenced by equation 4.5.3 above, the informed trader in the Grossman and Stiglitz model takes the market price as given, and thus does not consider the effect that his demand has on equilibrium price).

---

[111] According to Kyle the perfect competition among market makers ensures that no market maker can exercise monopoly power.

[112] The total order flow $x + u$ is denoted by the term y; whereby $y = x + u$. Given that the market maker observes only y (the aggregate order flow) equation (4.6.1) above can also be expressed as $p = E[v|y]$.

Because the informed trader in the Kyle model is risk neutral he will choose a trading strategy that maximises his expected end-of-period profits.[113] More precisely, under the assumption of risk neutrality, risk is not considered in the optimal behaviour of the agent as maximising expected profits and expected utility are equivalent (De-Jong & Rindi, 2009).[114] It follows that the informed trader's problem is to determine the optimal purchase (or sale) of quantity x.

His objective function is given by

$$\max_{x} E[\pi|v] = E[(v-p)x] \tag{4.6.2}$$

The interaction between the market maker's problem and the informed trader's problem is clear from the last two equations (4.6.1 and 4.6.2). The market makers pricing rule depends on the contribution of x to order flow, but the informed trader's choice of x depends on the impact orders have on the market maker's price p. This problem is known as circularity and is resolved in equilibrium.

4.6.2.3 Solving for Equilibrium

In Box (4.6) we provide a manual on how to derive a Bayesian Nash Equilibrium. Thus, to derive Kyle's Bayesian Nash Equilibrium, let us follow the steps highlighted above in Box (4.6). The first step is to propose linear equilibrium strategies for both the informed trader and the market maker. (As noted, Kyle focuses only on equilibria in linear strategies i.e., he *conjectures* linear

[113] Under the assumption of risk neutrality, the expected utility of end-of-period profits is simply equal to expected end-of-period profits.

[114] With an exponential utility function, we obtain: $Max\ E[U(\pi_i)] = Max\left[E(\pi_i) - \left(\frac{A}{2}\right)Var(\pi_i)\right]$. Adding risk neutrality ($A = 0$), maximisation of expected utility simplifies to maximisation of expected profits.

strategies for both the informed trader and the market maker and then *verifies* that these conjectures are actually the best response to one another's strategies).[115]

Let the proposed linear strategy for the informed trader be $X(v) = \beta(v - p_0)$ and for the market maker be $P(y) = p_0 + \lambda y$.

We follow convention here and use $X(v) = x = \beta(v - p_0)$ and $P(y) = p = p_0 + \lambda y$ to denote the informed trader and the market makers' strategy/rule respectively.[116] (Recall that x denotes the informed trader's traded quantity and X is his trading strategy, likewise p denotes the equilibrium price set by the market maker and P, the market maker's pricing rule – the notation is Kyle's).

An equilibrium is defined to be an $X(v)$ and $P(y)$ which simultaneously satisfy conditions/equations (4.6.1) and (4.6.2).

Because the coefficients β and λ will only be determined in equilibrium, we know that Kyle relies on the method of undetermined coefficients to solve for equilibrium.

The informed trader maximises his expected trading profit $E[\pi|v] = E[(v - p)x|v]$, where he takes into account the fact that according to his beliefs, $p = p_0 + \lambda(x + u)$. His optimal stock holding is then given by $x = (\frac{1}{2\lambda})(v - p_0)$

Recall, the market maker observes the aggregate net order flow $y = x + u$. Given his beliefs about the informed trader's trading strategy, $X(v) = x = \beta(v - p_0)$, the market maker tries to infer the value of the stock v from y. We should emphasise that very specific functional form assumptions underlie Kyle's analysis here.

[115] Because all conjectures will be linear combinations of functions in which the coefficients are 'constants to be determined', this whole approach to finding particular solutions is formally called the method of undetermined coefficients.

[116] Rule and strategy are used interchangeably throughout.

More precisely, the appealing simplicity of Kyle's equilibrium was seen to derive from the assumption that v and u are each normally and independently distributed. Provided that the informed trader's order x is linear in v, then, given v and u are normally distributed, the total market order, $y = x + u$, will be a linear combination of normals, implying that it will be normally distributed as well. That v and y are bivariate normal guarantees that the pricing rule – the expectation of the asset's value v conditional upon the market order y – will be a linear function of the market order. Consequently, y is said to provide a noisy signal of v. In particular, the net order flow provides a signal of the asset value v, so that the market maker can use the noisy signal y to form their expectation of v. The resulting conditional expectation $E[v|y]$ will generally differ from their unconditional expectation p_0. As the aggregate order size from noise traders' u is normally distributed and independent of v, the expected value of v conditional on y is provided by the projection theorem:[117]

[117] The projection theorem is very useful for deriving the conditional mean and variance. The proof of the projection theorem can be found in almost any statistics book (see for example, Goldberger, 1991). Here, the projection theorem follows quite directly from the ordinary least squares regression rule. Consider two joint normally distributed random variables, $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$, and denote their covariance σ_{xy}. A property of the bivariate normal distribution is that the conditional density of Y given $X = x$ is itself normal with conditional mean: $E[Y|x] = \mu_y + \frac{\sigma_{xy}}{\sigma_x^2}(x - \mu_x) = \left(\mu_y - \frac{\sigma_{xy}}{\sigma_x^2}\mu_x\right) + \frac{\sigma_{xy}}{\sigma_x^2}x$, which is the predicted value of Y from an ordinary least squares (OLS) regression of the equation $Y = a + bX$, upon setting the explanatory variable $X = x$. The slope coefficient $\frac{\sigma_{xy}}{\sigma_x^2}$ is precisely the OLS estimate of b. In Kyle's case the expected value of v conditional on y is provided by the projection theorem: $E[v|y] = E(v) + \frac{cov(v,y)}{var(y)}(y - E(y))$.

$$E[v|y] = E(v) + \frac{cov(v,y)}{var(y)}\big(y - E(y)\big) = p$$

$$= p_0 + \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}\, y$$

Kyle obtains his BNE by determining the coefficients β and λ. Given the best replies, these equilibrium coefficients were given by

$$\beta = \frac{1}{2\lambda} \qquad \text{and} \qquad \lambda = \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}$$

Moreover, in equilibrium $\lambda = 2\,(\,\sigma_u^2/\sigma_v^2)^{-1/2}$ and $\beta = (\,\sigma_u^2/\sigma_v^2)^{1/2}$.

Insights

Notice that the pricing and trading rules (P and X respectively) depend on the same two parameters – the variance of the uninformed order σ_u^2 and the variance in the payoff σ_v^2 – a natural consequence of being determined jointly and is analogous to the consistency criterion that governs the pricing rule in the rational expectations model. Notice also that the ratio of these two parameters is inverted in the two rules. This is quite intuitive: When λ is high, meaning that orders have a high price impact, then β is low, meaning that the informed trader trades less aggressively (to avoid the impact of his own trades). Here, β fully characterises the informed trader's trading intensity.

The constituent variance parameters are also easily interpreted. When σ_v^2 is high, the informed trader's information is more likely to be substantial, ceteris paribus, inducing the market maker to adjust price more aggressively. When σ_u^2 is high, the informed trader's order is a less conspicuous (better camouflaged) component of the total order flow, inducing him to trade more aggressively.

Thus, in equilibrium, the informed investor trades more aggressively the more noise trading camouflages his activity. Conversely, the market maker prices more sensitively the greater the information asymmetry (Lyons, 2001). To clarify, at the most basic level, the informed trader is seen to buy in proportion to the amount by which the risky asset is undervalued or sell in proportion to the amount by which the risky asset is overvalued. The factor of proportionality (β) – referred to as the trader's aggressiveness – is inversely related to the degree (λ)[118] to which orders affect the transaction price. Simultaneously, this price sensitivity (λ) responds positively to the degree of the information asymmetry (σ_v^2) and in some sense positively to trading aggressiveness.

In Kyle (1985), the market maker makes a loss on the trades with the informed trader, but recoups these losses on trades with the noise traders, making zero profit on average. Consequently, the informed traders expected profit is the liquidity traders' expected trading costs.[119]

Another interesting result of Kyle's (1985) model is the degree to which the informed trader's information is revealed by the equilibrium price. A common way to measure this is from the market maker's expectation of v. Specifically, after observing price, how much more precise is the markets expectation? (In the model, the 'market' is the uninformed traders because they have no information other than that conveyed by price.)

[118] Here, λ determines the price increase of an additional buy order. The reciprocal of λ can be viewed as, what Kyle (1985) refers to, as market depth. If λ is low, an additional order will not lead to a large price change and, thus, the market is very liquid. The small price impact of an additional order reflected by a low λ induces the informed trader to trade more aggressively.

[119] The expected profit for the informed trader in equilibrium is given by

$$E[(v-p)x] = \frac{1}{2}(\sigma_v^2, \sigma_u^2)^{\frac{1}{2}}$$

His expected profit is increasing in σ_v^2, since σ_v^2 measures the informed traders informational advantage.

Initially the market and the market maker have an expectation that is distributed about v with a variance we will denote as var_0. After seeing $x + u$, the market maker's expectation is distributed about v with a variance var_1; whereby,

$var_1 = \frac{1}{2} var_0$,

a fact that is easy to show using Bayes' rule for updating conditional variances.[120]

This is a striking result: Regardless of the realisations of v and σ_u^2, the updated variance is exactly one-half of the *prior* variance ($var_1 = \frac{1}{2} var_0$). The informed trader's strategy results in exactly one-half of his private information being revealed by the market price; that is, the new conditional variance is only half of the original unconditional variance.

Concomitantly, it can be shown in Kyle that the equilibrium pricing error is distributed as

$$(v - p)|y \sim N(0, \frac{\sigma_v^2}{2})$$

[120] The proof of the Bayes' theorem can be found in almost any statistics book. Therefore, we use it here without derivation. To see its formal derivation, we refer readers to the appropriate literature (i.e., Maurer & Ralston, 2005; McCarthy, 2000; Nanda, 2002; Platen & Heath, 2006). Broadly, Bayes' formulation for the determination of conditional variances of jointly distributed random variables can be expressed as:

$$var(Y|X) = var(Y) - \frac{[cov(X,Y)]^2}{var(X)}$$

For the purpose of the model, this formula can be restated as:

$$Var[v|p] = Var[v] - \frac{Cov[v,p]^2}{Var[p]}$$

Now, define $var_0 = \sigma_v^2$ and $var_1 = var[v|p]$. It is then simple to show using the formulation above that ($var_1 = \frac{1}{2} var_0$). Note that the unique models of this thesis provide analogous (albeit more detailed) examples of the above and should be seen as a substitute for explicit proof here - see e.g., Chapter 5, particularly Section 5.8.

Thus, congruent with the results obtained above, an expected square pricing error of $0.5\,\sigma_v^2$ implies that the market maker is able to infer half of the private information initially held by the informed investor.

In summary, the variance the market attributes to the fundamental value of the asset can be interpreted as how much information is incorporated into price. A variance of zero has to be interpreted as a perfect revelation of the information through prices, the closer this variance is to var_0 (with $var_0 = \sigma_v^2$) the less informative the price is. As by observing the price set by the market maker, the variance halves, it can be said that half of the information is incorporated into prices. Notably, the smaller the conditional variance, the more informative prices are.[121]

4.6.2.4. Kyle's Model: Discussion

There is extensive market microstructure literature on the topic of price formation in the presence of asymmetric information. Kyle's (1985) paper is related to this literature. As referenced above, asymmetric information is generally modelled in the form of informed traders – agents who have access to information that is unavailable to other market participants. This private information provides them with an advantage relative to the rest of the market. In the Kyle (1985) model, an informed trader trades against the price setting competitive market maker, in the presence of uninformed traders.

[121] That var_1 should settle at precisely $0.5\ var_0$ is not obstensibly clear, but it certainly is appealing on aesthetic grounds (Lyons, 2001, p.84).

In the static version of the Kyle's (1985) model, the market maker sets the price after observing the aggregate order flow.[122] The market maker sets the price equal to his best estimate, given his belief about the insiders trading strategy. Kyle derives a perfect Bayesian Nash equilibrium strategy where the informed trader's profit is increasing in his informational advantage and market depth. Intuitively, greater market 'depth'[123] implies that an additional order from the informed trader would not lead to a large change in prices, allowing the informed trader to trade more aggressively on his private information.

The original Kyle (1985) paper also extends this static model to a dynamic setting; focussing on the profit maximising temporal decision of the informed trader.[124] In the multi-period, dynamic version of the model, if the insider takes a larger position on the early periods, his early profits increase yet this comes at a cost of revealing his private information to the market.[125] Here (in the dynamic version of the model), the optimal strategy for the informed trader is to exploit the informational advantage over time. In other words, Kyle's continuous auction model suggests that the informed trader must consider the implications that current trades have on future opportunities. If the informed trader trades too much too soon, his private information will be revealed quickly, prices will adjust rapidly to his trades and his profit will be smaller. Instead, the informed trader

[122] In the context of the model, before and after do *not* refer to real time. Rather; these terms emanate from the application of conditional expectations/Bayes' rule/projection theorem, etc.

[123] Recall, λ determines the price increase of an additional buy order. The reciprocal of λ can be viewed as, what Kyle (1985) refers to, as market depth. If λ is low, an additional order will not lead to a large price change and, thus, the market is very liquid. The small price impact of an additional order reflected by a low λ induces the informed trader to trade more aggressively.

[124] Kyle (1985) solves the discrete-time and continuous-time versions of his model and proves convergence of the equilibria as the length of the time periods in the discrete-time model goes to zero. Thus, the discrete-time model with small time periods and the continuous-time model have equilibria that are approximately the same (Back, 2017, p.646). The question of how Kyle's dynamic model collapses to the static model is an interesting one. However, it is also a highly complex issue that requires more than just an overview. McCarthy (2000) devotes an entire thesis to this topic and should be consulted for further details.

[125] It is worth noting that it is assumed here that the future profits are not discounted to their present value.

chooses to reveal his information gradually by hiding trades among the uninformed trades as they arrive. Although not apparent at the outset, this generates the constant liquidity property referenced in insight (3). Briefly, the informed trader wants his trades to have minimal price impact. If price impact is less evenly dispersed, the informed trader could earn higher profit by trading more when impact is low and less when it is high. This incentive to *reallocate trading across time* results in constant liquidity in equilibrium.

Overall, Kyle's (1985) model has spawned a large literature of modifications, extensions, analysis, and applications to modelling other situations. Indeed, his model has been widely used as a basic framework in the market microstructure literature in finance. However, some features of the Kyle model limit its applicability. Following the literature (i.e., Lyons, 2001; McCarthy, 2000), we discuss some of these limiting features below.

4.6.2.5. Limitations of the Kyle Model

i) Private Information Acquisition is Not Specified

We can see from the model that the informed trader needs to condition his demand on his private information ($s = v + e$). However, practically speaking, one cannot determine how the informed trader acquires such knowledge as it is not specified in the model.

ii) A Single Informed Trader

Kyle (1985) develops a model in which a *single* informed trader with private information optimally exploits his monopoly power over time.[126] However, Kyle's (1985) assumption of a single informed trader is strong, in the sense that in actual financial markets, it is reasonable to expect that at least a few players will have access to private information, and will trade in the knowledge that they will face competition with other informed agents in the market. Thus, Kyle's model seems to neglect the crucial interactions that could occur between informed agents in more complex settings. In addition to being less descriptive of real markets, Kyle's supposition of a *single* informed trader may yield incomplete empirical implications involving price changes and agent behaviour relations (Holden & Subrahmanyam, 1992).

4.7 CHAPTER SUMMARY

Fama (1970) defines a 'strong form' efficient market as one in which security prices fully reflect all available information, including both publicly and privately held information. Private information was broadly defined in this chapter as *information that is not known by all trading parties*. However, to add some granularity to the definition, we also discussed the two subcategories of private information as they appear in the literature. A review of the literature suggested that informed trading takes one of two forms: (i) trading on more accurate information or (ii) trading on information faster than other investors. We followed with a discussion on the various facets of each dimension. For example, along with a description of (i) private information accuracy, we also discussed two possible mechanisms for its formation (i.e., (a) the interpretation mechanism and (b) the inference mechanism of private information accuracy).

[126] In contrast to Kyle's (1985) assumption of a single informed trader, Back, Cao and Willard (2000), Bernhardt and Miao (2004), Foster and Viswanathan (1996), Holden and Subrahmanyam (1992), and Li (2013), all develop market microstructure models in which an *aggregate of identical informed traders* strategically exploit their common private information. The respective models are analysed and discussed in detail in Chapter 6.

We also noted that whilst the literature recognises the *existence* of a two category breakdown of private information, these dimensions are rarely, if ever discussed within the same theoretical framework. We conclude our summary here with Table 4.1. As suggested in previous chapters, Table 4.1 provides an important comparative summary of the theoretical literature on both algorithmic trading and market microstructure theory.

TABLE 4.1: ORGANISING THE THEORETICAL LITERATURE ON ALGORITHMIC TRADING AND MARKET MICROSTRUCTURE: A COMPARATIVE SUMMARY.

This table organises the theoretical literature on both market microstructure and algorithmic trading according to the private information attributed to the informed agents in the respective models. Although numerous other examples have been referenced throughout this thesis, only the models that have been analysed in detail have been included herewith.

Market Microstructure Models

Author	Accuracy Advantage	Speed Advantage
Brunnermeier (2002)		✓
Grossman and Stiglitz (1980)	✓	
Kim and Verrechia (1994)	✓	
Kim and Verrechia (1997)	✓	
Kyle (1985)	✓	
Hirshleifer et al., (1994)		✓
Holthausen and Verrechia (1990)	✓	
Vo (2008)		✓

Algorithmic Trader Models

Author	Accuracy Advantage	Speed Advantage
Das et al., (2000)	✓	
Cartea and Penalva (2010)		✓
Cvitanic and Kirilenko (2010)		✓
Foucault et al., (2010)		✓
Johnson (2010)	✓	
Martinez and Rosu (2011)	✓	

4.8 CONCLUSION

One of the primary challenges encountered when conducting theoretical research on the subject of algorithmic trading is the wide array of strategies employed by practitioners. Current theoretical models treat algorithmic traders as a homogenous trader group, resulting in a gap between theoretical discourse and empirical evidence on algorithmic trading practices. In order to address this, the current study introduces an organisational framework from which to conceptualise and synthesise the vast amount of algorithmic trading strategies. More precisely, using the principles of contemporary cognitive science, it is argued that the dual process paradigm – the most prevalent contemporary interpretation of the nature and function of human decision making – lends itself well to a novel taxonomy of algorithmic trading.

This taxonomy serves primarily as a heuristic to inform a theoretical market microstructure model of algorithmic trading. In accordance with the literature on both cognitive science and algorithmic trading, this thesis espouses that there exist two distinct types of algorithmic trader; one (System 1) having fast processing characteristics, and the other (System 2) having slower, more analytic or reflective processing characteristics.

Concomitantly, the current microstructure literature suggests that a trader can be superiorly informed as a result of either (1) their superior speed in accessing or exploiting information, or (2) their superior ability to more accurately forecast future variables. To date, microstructure models focus on either one aspect but not both. This common modelling assumption is also evident in theoretical models of algorithmic trading. Theoretical papers on the topic have coalesced around the idea that algorithmic traders possess a comparative advantage relative to their human counterparts. However, the literature is yet to reach consensus as to what this advantage entails, nor its subsequent effects on financial market quality.

Notably, the key assumptions underlying the dual-process taxonomy of algorithmic trading suggest that two distinct informational advantages underlie algorithmic trading. It follows then, the possibility that System 1 algorithmic traders possess an inherent speed advantage and System 2 algorithmic traders, an inherent accuracy advantage.[127] Inevitably, the various strategies associated with algorithmic trading correspond to their own respective systems, and by implication, informational advantages. A model that incorporates both types of informational advantage is a challenging problem in the context of a microstructure model of trade. Models typically eschew this issue entirely by restricting themselves to the analysis of one type of information variable in isolation. This is done solely for the sake of tractability and simplicity (models can in principle include both variables). Thus, including both types of private information within a single microstructure model serves to enhance the novel contribution of this thesis.

We suggest here that Kyle's (1985) model provides the necessary microstructural foundations for our unique model of algorithmic trading. Indeed, specifying a model that captures the multidimensional aspects of algorithmic trader's within a contemporary modelling framework, such as Kyle's, remains a crucial consideration of this work. A fundamental shortcoming of his model however, is evidenced by the supposition of only a single informed trader. In contrast to Kyle (1985), we propose an integrative model featuring two informed traders (which we later claim to be algorithmic traders). Moreover, in an attempt to incorporate the key informational assumptions of our formal taxonomy (to follow in Section 5.3.2); we suggest that relative to a group of identical informed traders, the informed traders in our model have heterogeneous

With respect to each other and the rest of the market.

informational advantages. Hence, we will allow for heterogeneity in agents (algorithmic traders) signal timing as well as signal quality.

The *modus operandi* of the current study is detailed below:

To prepare for the final 'all inclusive' dual-process theoretical model of this thesis, the present study will first conjecture and verify a benchmark model with only one type/system of algorithmic trader. More formally, a System 2 algorithmic trader will be introduced into Kyle's (1985) static Bayesian Nash Equilibrium (BNE) model. The behavioural and informational characteristics of this agent emanate from the key assumptions reflected in the taxonomy. Effectively, the System 2 algorithmic trader replaces the informed trader in Kyle's (1985) static model.[128]

The final dual-process microstructure model, presented in the concluding chapter of this thesis, extends the benchmark model by introducing the System 1 algorithmic trader; thereby, incorporating both algorithmic trader systems. Note: the benchmark model nests the Kyle (1985) model. In a limiting case of the benchmark model, where the System 2 algorithmic trader does not have access to this particular form of private information, the equilibrium reduces to the equilibrium of the static model of Kyle (1985). Likewise, in the final model, when the System 1 algorithmic trader's information is negligible, the model collapses to the benchmark model.[129]

[128] The term 'replaces' is perhaps too strong a word, rather; one could simply view the System 2 algorithmic trader in our model as a modified version of Kyle's informed trader. The difference between replace and modify is largely semantic and has no noteworthy implications.

[129] The above dictum should be clarified by the models themselves. Those readers who require further clarification at this stage are referred to Table V (in Appendix V) at the end of this thesis. Table V provides a succinct comparison of the models of this thesis, noting the limiting cases. However, the technical aspects of the models themselves may inhibit Table V's applicability prior to a formal analysis of each model.

One might argue that Kyle's (1985) model could suffice as the benchmark model. However, we maintain that this thesis will be greatly enriched with the inclusion of our own research specific benchmark model.

4.9 SOME ADDITIONAL THOUGHTS

To conclude this chapter, some reflection on an important tenet of this thesis is necessary. Specifically, this study presents the first unified, 'all-inclusive'[130] theoretical model of algorithmic trading; the overall aim of which is to determine the evolving nature of financial market quality as a consequence of this practice.

Indeed, it can be stated that this final integrative model is *exclusive* to algorithmic traders; the basic principles of which emanate from the key assumptions made with regards to these traders in our taxonomy. The models parity with algorithmic trading is justified below.

4.9.1 Algorithmic Traders as Informed Traders: An Affirmation

Kirilenko and Lo (2013) suggest that the growth of algorithmic trading emerged out of the desire to eliminate the need for limited human judgment.[131] Since our cognitive abilities are relatively limited, algorithms have enabled markets to far exceed the cognitive bounds of humans in processing information (Biais and Woolley, 2011).[132]

[130] The final model is all inclusive in the sense that accounts for the multitude of algorithmic trading strategies within a single theoretical framework. Given the inherent complexities involved, such a model is yet to be explored in the academic literature. Again, models typically restrict themselves to the analysis one type of information variable in isolation. That being said, we can confirm the possibility of a model including both types of information variables within the same framework.

[131] See Chapter 2, Section 2.2.1, where we discuss some of the limitations in human judgment (refer also to the respective literature, i.e. Kahneman, 2003; Kahneman & Frederick, 2002; Kahneman & Tversky, 1982, 1996; Tversky & Kahneman, 1974, 1983)

[132] See e.g. Salmon and Stokes (2010), particularly the description of areas where algorithmic trading excels over human operators.

Specifically, an individual human operator can only incorporate a small amount of data into his or her decision-making process, and so investment strategies requiring the calculation of matrices with hundreds of thousands of variables necessarily require computer technology (McNamara, 2016).

Algorithms try to make sense of the mass of data generated by the market to generate a usable output. For example, an algorithm might analyse a listed company's past dividends, available data on its debt, and market conditions, and take a position on the current value of the shares. Once the algorithm has calculated the value of shares, it can decide whether to buy or sell shares based on a pre-set strategy. The algorithm might use established valuation modelling techniques from finance theory, crunching large amounts of data to arrive at a more exact valuation for the security using the model (Narang 2013). Algorithms can also harness complex computational techniques vis-à-vis advanced statistical analysis, artificial intelligence, machine learning, neutral networks, support vector machines tools, data mining as well as text mining and can transact at speeds measured increasingly in microseconds (1 microsecond is an international standard unit of time equal to 1000000th of a second)[133]

Because algorithms are tasked to perform complex trades using deep data and speed, they must possess some programmed 'decision-making' capacity. In other words, algorithms must be capable of evaluating the importance of data, attaching a value to its content, and then making a decision independently of human traders. Indeed, it is becoming increasingly evident that algorithmic traders represent a form of advanced automated 'decision-maker'.

[133] On average it takes approximately 300 milliseconds for a human being to blink (Wilkinson et al., 2013).

All things considered, it seems that humans are at a distinct disadvantage in a market dominated by computer algorithms. The axiom that algorithmic traders have powerful advantages over their human counterparts is further evidenced by the fact that computers have both a greater information processing capacity and a higher reaction speed relative to human beings. The literature on algorithmic trading (e.g. Biais et al., 2014; Cartea & Penalva, 2010; Cvitanic & Kirilenko, 2010; Das et al., 2001; Easley et al., 2012; Foucault et al., 2016; Gamzo, 2014; Johnson, 2010; Martinez & Rosu, 2011) indicate that algorithmic trader's consistently outperform their human counterparts in terms of trading speed as well as in terms of predictive accuracy.

A natural means of addressing the implications of the above is to orient it towards the following question*: do we need to modify the conventional view of informed traders in light of the emerging evidence surrounding algorithmic trading*? If we construe that algorithmic traders are indeed the most informed decision makers in the market then it is likely that the answer to this question would be in the affirmative.

Herein lies our rationale for suggesting that the unique models of this thesis are *exclusive* to algorithmic traders. Consider the following narrative that explicates this proposition:

Kyle's (1985) model has been widely used as a basic framework in the market microstructure literature in finance - the model's simplicity and flexibility makes it a useful point of departure for many microstructure researchers. It seems only natural then that we base our own analysis on his framework. However, it should be noted that Kyles (1985) model emerged prior to the advent of algorithmic trading. Accordingly, his analysis seems to have been constrained to suppositions

surrounding the cognitive abilities and information processing capacities of human traders; which relative to algorithmic traders are extremely limited.[134]

With this in mind, we felt it necessary to conjecture and verify a preparatory (benchmark) equilibrium model in order to reconcile Kyle's (1985) equilibrium characterisation with our own unique algorithmic trading considerations. More precisely, we begin by introducing a 'System 2' algorithmic trader into a static version of Kyle (1985). Essentially, the System 2 algorithmic trader will replace Kyle's (1985) informed trader. To justify that the System 2 trader is algorithmic, an attempt was made to align his information set with the key informational assumptions made in our taxonomy, delineated in the proceeding chapter. The basic tenets for differentiating between the System 2 algorithmic trader and the informed traders in Kyle's model can be described concisely.

The following discussion remains an overview and does not include notation. (Precise details follow in Chapter 5).

Consider the following succinct outline: In Kyle's original (1985) model there is some noise in the informed trader's private information. However, in line with the discussion above, we would like the System 2 algorithmic trader to be 'more'[135] informed than the informed (human) trader in Kyle. In light of this objective, information was modelled in such a way so as to suggest that the System 2 algorithmic trader could forecast the firm's value without error (note that the model itself will clarify and expand in this principle – with respective notation.) Specifically, our set-up allowed the System 2 algorithmic trader to filter out noise from his private information.

[134] See also, Salmon and Stokes (2010), particularly the description of areas where algorithmic trading excels over human operators.

[135] The term 'more' is in reference to the accuracy of the informed trader's private information; whereby, the respective trader is able to arrive at a more exact forecast of the security.

This was done intentionally to account for the presumed characteristics of System 2 algorithmic trader's in our taxonomy (see Section 5.2).

The final dual-process microstructure model, presented in the concluding chapter of this thesis, extends the benchmark model by introducing the System 1 algorithmic trader; thereby, incorporating both algorithmic trader systems. Like the case with the System 2 algorithmic trader, the key considerations for this algorithmic trader emanate from the key assumptions made for System 1 algorithmic trader's in our taxonomy.

Note that the formal taxonomy of this thesis is supported by actual empirical evidence on algorithmic trading practices. An instantiation of System 1 and System 2 algorithmic trading in Chapter 3 serves only to enhance the validity of this taxonomy.

To conclude, given (a) the suggestion in the literature that algorithmic traders have an informational edge relative to human traders and (b) that the key characteristics of the algorithmic traders in our model(s) emerge from a taxonomy *exclusive* to algorithmic traders; it is reasonable to suggest that the models of this thesis are exclusive to algorithmic traders.

In the next chapter (Chapter 5), the benchmark model is formally introduced. Chapter 5 also presents our definitive dual-process *taxonomy* of algorithmic trading. The final all-inclusive microstructure model is detailed in Chapter 6.

CHAPTER 5

A THEORETICAL DUAL-PROCESS MODEL OF ALGORITHMIC TRADING: AN INTRODUTION

5.1 OVERVIEW

There remains a huge disconnect between available theoretical models of algorithmic trading and recent empirical research on the practice of algorithmic trading, purporting that algorithmic trading is a heterogeneous phenomenon. Given that algorithmic trading has asserted itself as a dominant force in financial markets across the world, it seems that a model that accounts for the multifaceted nature of algorithmic trading has become vital to our understanding of financial market performance (Gomber et al., 2011). As earlier noted, this thesis will propose a theoretical model of algorithmic trading that will draw on the dual-process theory of cognition to inform an organisational taxonomy in which to conceptualise and synthesise the vast amount of algorithmic trading strategies. Following the literature on both cognitive science and algorithmic trading, this thesis espouses that there exists two distinct types of algorithmic trading; one (System 1) having fast, reflexive, process characteristics, and the other (System 2) having relatively slower, more analytic or reflective processing characteristics.[136]

[136] So far we have surveyed Kyle's (1985) Bayesian Nash Equilibrium model and the Grossman and Stiglitz (1980) Rational Expectations Model. The Grossman and Stiglitz (1980) model was natural starting point for our review in Chapter 4. Opening with this model also clarified how its shortcomings spurred the development of later microstructure models such a Kyle (1985). The Grossman and Stiglitz (1980) Rational Expectations model therefore served only to preface the Kyle's model - with Kyle's (1985) model providing the necessary microstructural foundations for the unique models of this thesis. The remaining focus of this thesis is thus aligned solely with Kyle (1985). With regards to Kyle's (1985) Bayesian Nash Equilibrium model and the Grossman and Stiglitz (1980) Rational Expectations Model one could say that if the informed traders in these models were algorithmic, then they would be System 2 algorithmic traders. This proposition is further clarified by the fact that there is no speed dimension in these models.

Using this integrative dual process taxonomy, we will attempt to construct an overarching theoretical model of algorithmic trading; whereby the key characteristics of the traders reflected in the taxonomy will translate into key behavioural and informational assumptions of the agents in this model.

From a theoretical modelling perspective, the key assumptions underlying our dual-process taxonomy of algorithmic trading suggests that two informational advantages underlie algorithmic trading. Naturally, all the various strategies associated with algorithmic trading (as seen in Section 2.4) correspond to their own respective 'System' and thus, informational advantage.

A review of Section 3.5 may prove useful for the interested reader. This serves to orientate the reader with the classification dimensions of our dual process construct of algorithmic trading. Below, we provide our final integrative taxonomy of algorithmic trading.

5.2 THE FINAL INTERGRATIVE DUAL-PROCESS *TAXONOMY* OF ALGORITHMIC TRADING

Following our analysis of *empirical*, as well as *theoretical* academic literature on both *algorithmic trading* and *human cognition* we define algorithmic trading as follows:

Algorithmic trading represents computer-determined trading whereby super computers and complex algorithms directly interface with trading platforms at high speed, placing orders without immediate human intervention. It (algorithmic trading) employs cutting edge mathematical models, adept computational techniques and extraordinary processing power via advanced computer and communication systems and is capable of anticipating and interpreting market signals in order to implement profitable trading strategies.

A key assumption made with regards to our taxonomy is thus; two distinct subsets of algorithmic trading currently exist – that is, System 1 algorithmic trading and System 2 algorithmic trading. It follows that System 1 algorithmic traders have fast information processing characteristics. System 2 algorithmic traders have relatively slower, more analytic or reflective information processing characteristics. Theoretically speaking, System 1 algorithmic traders possess an inherent speed advantage and System 2 algorithmic traders, an inherent accuracy advantage – relative to both each other and the rest of the market.

The following discussion considers the key assumptions made in the taxonomy; noting the theoretical implications for System 1 and System 2 algorithmic trader's in our model(s)

I) System 1 Algorithmic Trading

From a theoretical perspective, our dual process categorisation of algorithmic trading implies that System 1 algorithmic traders possesses and inherent speed advantage - relative to the rest of the market. In practical terms, speed or latency refers to the time it takes to access and respond to market information. Accordingly, following the instantiation of the concept in Section 3.5, we will assume that System 1 algorithmic traders can effectively anticipate incoming market *orders* and trade rapidly to exploit normal-speed traders' latencies. Their speed advantage is predicated on latency specific mechanisms e.g. colocation facilities and applies to information about incoming order flow and not about the fundamental value of the asset.

Faster speeds also imply that they can trade more frequently, have smaller inventories and shorter holding horizons.[137] Consistent with the insight of Froot, Scharfstein and Stein (1992), System 1 algorithmic traders in our model, will focus on short-term order flow, becoming less informed about long-term fundamentals. It follows that, when establishing a position, these traders must have a plan to exit within a short time window. Thus, System 1's profit is not determined by the difference between his entry price and the fundamental value, but by the difference between his entry and exit prices. (See Chapter 6 for proof of concept)

II) System 2 Algorithmic Trading

System 1 algorithmic trader's speed advantage relates to information about incoming order flow. This is particularly pertinent given that order book information travels fast within the exchange. However, information on the macro-economy or firm fundamentals, which in general are larger and more complex than order book information, 'travels'[138] slower and is more difficult to filter than order book information. This is where System 2 algorithmic trader's accuracy advantage comes to the fore. In practical terms, accuracy relates to the precision of a long-run, firm specific, forecast. It is predicated on the extent to which an agent is informed about the intrinsic worth of a firm. It follows, that System 2's accuracy advantage is linked to their ability to predict key, fundamental, value relevant, firm specific variables in equity markets.

[137] The short holding horizon has important implication on the fast trader's behavior and choice of information. In the short-term, the resale value of a risky security is more likely driven by the order flow rather than the fundamental values. We discuss the implications that these assumptions have on equilibrium below.

[138] The vast scale and complex nature of this information means that it takes *longer* for the market to *disseminate* this information.

In our evaluation, System 2 algorithmic trading represents an information processing paradigm with a remarkable tolerance for *noise*, ambiguity and uncertainty. Its ability to derive meaning from vast, complicated and imprecise information means that System 2 algorithmic traders can detect patterns in information and identify trends that are too intricate to be noticed by humans alone.

Accordingly, following the instantiation of the concept in Section 3.5, we will assume in our model that System 2 algorithmic traders exploit their superior ability to interpret public information in an attempt to make forecasts that are superior to the forecasts of other traders. In other words, these traders filter public fundamental information through an advanced platform (high-capability computer infrastructure), in order to detect private patterns from public information – patterns that signal a firms future performance.

5.3 MODEL SPECIFICS

Overall, specifying a model that harmoniously captures the multidimensional aspects of algorithmic traders within a contemporary microstructure modelling framework, such as Kyle's (1985), remains a crucial aspect of this work. As such, our equilibrium characterization will follow Kyle's (1985) Bayesian Nash Equilibrium (BNE).

A crucial shortcoming of his model however, is evidenced by the supposition of only a single informed trader. In contrast to Kyle (1985), we present a model featuring two informed traders (which we later claim to be algorithmic). Moreover, in an attempt to incorporate the key informational assumptions of our proposed taxonomy, we suggest that the informed traders have heterogeneous informational advantages. Hence, we will allow for heterogeneity in agents – algorithmic traders – signal timing as well as signal quality. (This supposition will be clarified by

the model itself).[139] Overall we seek to determine how the strategic interplay among differentially informed algorithmic traders affects market quality (i.e. trading activity, prices, volume, liquidity, volatility and profits) over time.

In our paper the System 1 algorithmic trader will be faster than the System 2 algorithmic trader but the System 2 algorithmic trader will have better information about the fundamentals. The difference between algorithmic traders (with respect to asymmetric/private information) is developed and addressed in the final model in Chapter 6.

5.3.1 The Foundations of the Model

Due to its position in the development of microstructure theories, Kyle's (1985) model has emerged as the standard to which all microstructure models are compared. It has spawned a large literature of modifications, extensions, analysis, and applications to modelling other situations. Given that this model has been widely used as a basic framework in the market microstructure literature in finance, it seems only natural that our model follows Kyle (1985). However, as earlier noted, Kyle's (1985) model was designed with agent specific considerations that are not ostensibly relevant to our specification of algorithmic trading. With this in mind, we begin by conjecturing and verifying a preparatory (benchmark) equilibrium model in an attempt to reconcile Kyle's (1985) equilibrium characterisation with our own research specific algorithmic trading considerations. See below.

[139] Technically speaking, because Kyle's (1985) model is aligned with the accuracy dimension of informed trading, one could suffice in saying that the informed trader in Kyle's model could be algorithmic (vis-à-vis a System 2 algorithmic trader). However, in the interest of verisimilitude, we remain resolute that the System 2 algorithmic trader should be better informed - have more accurate private information - than the (human) informed trader in Kyle (1985). Specifics provided in what proceeds.

5.3.2 The Benchmark Concept

As a preparatory step for our final model, we will begin by conjecturing and verifying a benchmark equilibrium market microstructure model. We will construct a benchmark model that is (1) analogous with Kyle's equilibrium specification and (2) includes a property of information processing that parsimoniously captures our specification of algorithmic trading.

The benchmark model provides the initial link between the dual process construct of algorithmic trading and market microstructure and allows us to formalize our final integrative model of algorithmic trading. Thus in actuality, two models are presented in this thesis; the benchmark model and the final integrative dual process microstructure model of algorithmic trading (*it could be argued that Kyle's (1985) model could function as the benchmark model. However, we maintain that this thesis will be greatly enriched with the inclusion of our own benchmark model.)*

To prepare for the additional assumptions of our final model, this section presents a benchmark model with only one algorithmic trading type/system. More formally, we introduce System 2 algorithmic trading into a static version of Kyle (1985). The System 2 algorithmic trader replaces the informed trader in Kyle's (1985) static framework. As clarified below, we make an assumption that the System 2 algorithmic trader has access to a unique form of private information (we define, and provide the rationale for this particular form of private information shortly).

It is important to note that in a limiting case of the benchmark model, where the System 2 algorithmic trader does not have access to this particular form of private information, the equilibrium reduces to the equilibrium of the static model of Kyle (1985). Inexorably, the baseline/benchmark model nests the Kyle (1985) model. The model itself will qualify this suggestion.

The intuition for including System 2 algorithmic traders in a model before System 1 algorithmic traders is as follows:

A key assumption made in our taxonomy is that System 1 algorithmic traders' private information relates to incoming order flow. This means that they have no independent source of information; their information is dependent on the activity (orders) of other traders –including, but not limited to the System 2 algorithmic trader. Thus, it seems inappropriate to include System 1 algorithmic traders without first modelling System 2 algorithmic trading.

CHAPTER 5 (Cont'd)

5.4 THE BENCHMARK MODEL: AN INTRODUCTION

In this section we present the first of our two microstructure models i.e., the benchmark model. The benchmark model includes System 2 algorithmic traders but not System 1 algorithmic traders. The final model, presented in Chapter 6, includes both types of algorithmic traders.

Technically speaking, Kyle's (1985) model is aligned with the accuracy dimension of informed trading. Therefore, one could argue that the informed trader in Kyle's original (1985) model could be algorithmic (vis-à-vis a System 2 algorithmic trader).

However, in the interest of verisimilitude, we remain resolute that the System 2 algorithmic trader in the benchmark model should be 'more'[140] informed than the informed (human) trader in Kyle (1985).

Our equilibrium characterization for the benchmark model will follow the lines of Kyle (1985). However we will augment his theoretical platform to allow for sequential information acquisition. We introduce an adjusted information structure into the static Kyle (1985) model and add a new type of trader who has access to unique private information. Again, this section presents the benchmark model with only one type/system of algorithmic trader – System 2 algorithmic trader in our case.

In the particular model we investigate, one risky asset is exchanged for a riskless asset among *three* kinds of traders: a single System 2 algorithmic trader who has private, sequential observations of the *ex-post* liquidation value of the risky asset; uninformed noise traders who trade randomly; and market makers who set prices efficiently (in the semi-strong sense) conditional on the information they have about the quantities traded by others.

[140] The term 'more' is in reference to the accuracy of the informed trader's private information; whereby, the respective trader is able to arrive at a more exact valuation of the security. Specifically, in Kyle's original (1985) model there is some noise in the informed trader's private information. However, in line with the insights discussed in Chapter 4 above, information in the benchmark model will be modelled in such a way so as to suggest that the System 2 algorithmic trader is able to forecast the firm's value without error (note that the model itself will clarify and expand in this principle – with respective notation.) Essentially, our set-up allows the System 2 algorithmic trader to filter out noise from his private information. This is done intentionally to account for the presumed characteristics of System 2 algorithmic trader's in our taxonomy. We remind readers that our taxonomy is supported by actual empirical evidence on algorithmic trading practices. An instantiation of System 1 and System 2 algorithmic trading in Chapter 3 serves only to enhance the validity of our taxonomy. Thus, given (a) the suggestion in the literature that algorithmic traders have an informational edge relative to human traders and (b) that the key characteristics of the algorithmic traders in our model(s) emerge from a taxonomy specific to algorithmic traders; it is reasonable to suggest that the benchmark model to follow, is exclusive to System 2 algorithmic traders.

5.4.1 Benchmark Model Setup

Assets and Traders The System 2 algorithmic trader and the uninformed noise traders' trade two assets: a risk free asset (bond)[141] with zero interest rate and a risky asset (stock). Like Kyle (1985), we begin by assuming that the risky asset to be traded has an *ex-post* payoff of v where the value is drawn according to,

$$v \sim N(p_0, \sigma_v^2)$$

Contextually, nature selects a true value (v) from a prior normal distribution $N(p_0, \sigma_v^2)$ for the traded asset. Effectively, the *ex-post* liquidation value of the risky asset is its true/fundamental value and is a normally distributed random variable with initial mean p_0 and variance σ_v^2. *The assets initial distribution is public information.*

As in Kyle (1985), all players are risk neutral. The assumption of risk neutrality means that p_0 also typifies the initial stock price - this insight can only be fully appreciated following a detailed exposition of the model itself (see below).

In line with conventional Kyle type microstructure models, two types of normal (non-algorithmic) traders exist: (1) uninformed noise traders randomly trade normally distributed $u \sim N(0, \sigma_u^2)$, shares for exogenous non-informational reasons;[142] (2) market makers set the pricing function,

[141] The majority of microstructure models include a risk-free asset (e.g., a riskless bond). Because risky asset demand is independent of wealth, in order for each individual's demand to be feasible, then each individual must be able to borrow and lend at the risk-free interest rate without constraint (Lyons, 2001). However, its economic role is trivial and some models omit it from exposition. We include it because Kyle (1985) includes it.

[142] Uninformed traders are often recorded as noise traders in the literature. They are liquidity motivated, smoothing their inter-temporal consumption stream through portfolio adjustment. Their motive for trade is often referenced as a hedging motive. Like Kyle (1985), u is exogenously determined (simplifying inference). See Bagehot (1971) for a detailed discussion on these topics.

absorb the residual order flow imbalances, and make zero expected profits.[143] (The key assumptions regarding these two agents emanate from Kyle (1985) and are detailed extensively in Section 4.6)

However, we introduce a new type of trader: System 2 algorithmic traders who receive multiple pieces of private information and are ultimately informed about the true value v of the risky asset. The properties and structure of their information set differentiates these traders from informed traders in existing models. Notably, in many existing microstructure models of trading, such as Back (1992), Holden and Subrahmanyam (1992) and Kyle (1985), informed traders receive a simplified, one-shot signal of firm fundamentals.[144]

As reflected in our taxonomy however, one of the defining features of a System 2 algorithmic trader lies in its complex, logical and deductive processing capabilities and inherent predictive abilities. More specifically, System 2 algorithmic traders exploit their superior ability to interpret 'public fundamental information'[145] in an attempt to make forecasts that are superior to the forecasts of other traders. They filter public fundamental information through an advanced

[143] Recall, the key conceptual difference between rational expectations models and Kyle's (1985) model is that Kyle includes an explicit auctioneer (market maker) rather than an implicit one. The market makers pricing rule is pinned down by the assumption that he expects to earn a profit of zero. Again, market makers are simply vote counters, not analysts of fundamentals; and the votes they count are the order flow. The assumption that market makers earn zero profit is important to the model and is shared by many other models within microstructure. This zero profit assumption is consistent with free entry of competing market makers, a condition under which the single market maker cannot exercise monopoly power. Essentially, the market makers determine the price p at which they trade the quantity necessary to clear the market.

[144] Typically, the informed trader has private information about the risky asset v, in the form of a *single* signal $s = v + e$. The term e is used to represent the noise in the signal. However, for future reference, consider the following principle: The structure of the informed trader's private information (e.g., whether or not the informed trader observes v perfectely or with noise) is not crucial. As articulated by Rochet and Vila (1994): "Given that all traders are assumed to be risk-neutral it is only needed that the information structures be nested i.e., that the informed trader knows more than the market" (p.132). This principle suffices as justification for our ability to modify Kyle's with our own agent specific informational considerations – detailed below.

[145] Although an observation of the risky assets final liquidating value v would be considered private information, its distribution (mean and variance) is publically available.

platform (high-end capability computer infrastructure), in order to detect private patterns from public information – patterns that signal a firms future performance. Essentially, we assume that System 2 algorithmic traders can derive meaning from vast, complicated and imprecise information.

Common sense suggests that this type of processing requires *multiple reflective and corrective observations*. Therefore, unlike most traditional approaches, the informed agent in our model receives multiple informative signals, which then can then be combined and filtered to produce a perfect forecast of an asset's true/fundamental value.[146] Most notably perhaps, the System 2 algorithmic trader can forecast without error. This is consistent with the reflective/effortful processing characteristics of System 2 algorithmic traders in our taxonomy. We model this type of information as a sequence of multiple signals in the spirit of what Kim and Verrecchia (1994, 1997) define as 'pre-announcement' and 'event-period' private information. However, given that the material on pre-announcement and event-period private information is highly advanced, we do not expand on Kim and Verrecchia's (1994, 1997) models here. The rationale for excluding these models in this section is provided below.

Evidently, market microstructure models are said to fall into three clearly defined levels of difficulty or 'modules': a basic module, an intermediate module, and an advanced module (Cukierman, 1992, p.9). We have covered the basic and intermediate modules in the previous sections. Given that Kim and Verrecchia's (1994, 1997) models can be characterized as advanced modules, they are treated in a separate appendix (Appendix II). This is done solely to enhance the fluidity of this thesis. Technical details of Kim and Verrecchia's (1994, 1997) models are

[146] When the trading period concludes, the risky and riskless assets are liquidated, and agents consume their holdings of each. Thus, the term fundamental value is often referenced as the *ex-post* liquidation value of the risky asset.

contained within Section (b) of Appendix II. That being said, the model itself (below) should serve to clarify this information-theoretic issue, namely, the structure of System 2's private information.

5.4.2 The Benchmark Model – Some Caveats

The benchmark model describes an economy with two periods. However, because only one of these periods can be classified as a *trading* period, the model is not a multi-period model. Evidently, to provide some clarity on this statement, some major caveats merit a discussion here before continuing.

It may appear, to some, that the models presented in this thesis are multi-period models. However, both models (i.e., the benchmark model and the final model) are in fact single-trading period models.

In order to understand this caveat, one needs to first appreciate the difference between multi-period and single trading period models. A brief discussion on Kyle (1985c) should clarify the conceptual difference between multi-period and single-period models. Consider the following:

As stated above, Kyle's model has been widely used as a basic framework in the market microstructure literature in finance. Invariably, when researchers (including the present study) refer to the Kyle's model, they are referencing his single-auction equilibrium model. Indeed, in keeping with convention, the present study has followed the literature. In actuality, there are three different versions of Kyle's model:

Kyle (1985a, 1985b, 1985c) – henceforth 'Kyle' - presents three models in his seminal work: 1) the 'single-auction equilibrium model'; 2) the 'sequential auction equilibrium model' and 3) the 'continuous auction equilibrium model'.

Kyle's 'single-auction equilibrium model' is a simple model of one-shot trading game. Here, a single risky asset is traded over the time period [0,1]. In other words, trading occurs at date 0, and the game ends at date 1.[147]

Kyle's second model, the 'sequential auction equilibrium model', is a discreet-time, *multi-period* version of the 'single-auction equilibrium model' – the first model. Here, there are multiple auctions, or rounds of trading, which take place sequentially.

The 'continuous auction equilibrium model'- the third model - is a continuous version of the the 'sequential auction equilibrium model'. In this third model trading takes place continuously rather than at discrete intervals.

We do not analyse the results of these models here, rather, for the purpose of the discussion to follow; we simply note that Kyle's models' variants highlight the conceptual difference between multi-period and single-period trading models. Evidently, when there are multiple auctions, or rounds of trading, the model can be classified as a multi-period model. Conversely, when there is only a *single auction*, or trading round, the model falls under the rubric of single-period model.

Again, the benchmark model describes a two-period economy. However, actual trading occurs within *one* of these periods. Therefore, we cannot classify the benchmark model as a multi-period trading model. To corroborate this suggestion, we expand on the benchmark models timeline and information structure below.

[147] At the end of the game all agents consume their holdings and receive their payoffs.

Timeline and Information Structure Like Kyle's single-auction equilibrium model, there is only a single *trading* period in the benchmark model. However, to prepare for the additional assumptions made for System 2 algorithmic traders in our taxonomy, we adjust his timeline somewhat and include, what can only be described as, an additional 'period'. This term period is used loosely here because what we include it is not a period, per se, but rather an extended phase of information processing. The intuition for this arrangement stems from the relevant theoretical microstructure literature (see for example, De long, Shleifer, Summers & Waldmann, 1990; Dugast & Foucault, 2016; and to some extent Kim & Verrecchia, 1997).

We include the extended processing phase/period (henceforth, the terms extended processing *period* and extended processing *phase* will be used interchangeably) in our model in an attempt to reflect an important assumption made in our taxonomy: that is the assumption that System 2 algorithmic trader's precise/effortful information processing characteristics translate into an informational advantage. In other words, we have assumed that these traders are able to filter *public information* (through an advanced platform)[148] into private, and perhaps more accurate signals of a *firm's fundamental value (v)*. Essentially, the System 2 algorithmic trader's ability to derive meaning from vast, complicated and imprecise information means they are able to filter out much of the *noise* associated with fundamental information analysis. In reality however, filtering out noise from a signal requires multiple reflective and corrective observations. Dugast and Foucault (2016) provide a concise articulation of this filtration principle and state that "information processing filters out noise in raw information but it *takes time*" (p.1). Similarly, Deb, Koo and Liu (2014) construct a model in which informed traders can either trade early on an unverified,

[148] This advanced platform is known as high-end capability computer infrastructure.

noisy signal or they can choose to wait and trade on verified information. Inexorably, Deb et al. (2014) imply that filtering out noise from a signal requires extended processing.

Thus, by including an extended processing period in the benchmark model, we are attempting to capture the aforementioned information filtration principle; the supposition that filtering out noise from a signal is an extended and complex process.

Note that in Kyle's original (1985) formulation the informed trader receives a simplified, one-shot signal of firm fundamentals (i.e. $s = v + e$). However, as evidenced by the error term (e) there is some residual noise in his signal. Naturally, Kyles (1985) model seems to ignore the information filtration principle.

Below we present an overview the benchmark model's information structure, as it pertains to the System 2 algorithmic trader.

5.4.2.1 Overview of signal structure

Before we formally introduce the benchmark model, we provide a basic overview of the model's information structure (the structure and formation process of System 2 trader's private information). Note the introductory nature of this overview precludes an in-depth discussion here. Precise details are provided in the next sections.

Consider the following broad outline of the benchmark model's information environment.

Notably, there is only *one* trading round in the benchmark model. We denote this single trading round with the upper case letter $\boldsymbol{T}$. In $\boldsymbol{T}$, a single auction takes place in the interval [0.1]. More formally, a security is traded in a single auction in a time interval which begins at time $t = 0$ and ends at time $t = 1$.

The securities value at the end of trading – its fundamental value – is denoted by v which is assumed to be normally distributed with an *initial* mean p_0 and variance σ_v^2; in short $v \sim N(p_0, \sigma_v^2)$. In line with conventional microstructure models of trading (e.g., Kyle, 1985; Foster & Viswanathan, 1996; Holden & Subrahmanyam, 1992; Vayanos, 1999) the informed trader in our model (the System 2 algorithmic trader) is set to receive private information in the trading period ($\boldsymbol{T}$) in the form of a noisy signal s about the securities end of trading value v. This signal is set to satisfy:

$$s = v + e_s \qquad (e_s \text{ represents the noise in the signal}).$$

(The subscript s in e_s was used to affirm that the error relates specifically to the signal s)

This is where the extended processing period comes to the fore.

To prepare for the assumptions made for System 2 algorithmic traders in our taxonomy, we extend Kyle's (1985) timeline somewhat and introduce an extended phase of information processing. The extended processing phase is an information gathering period for System 2 algorithmic traders and there is no trading in this period. We have already alluded to this no-trade condition in the preceding paragraphs. That noted, we should emphasize that the rationale for this no-trade condition has been motivated in part by De long et al., (1990).

In lieu of a lengthy analysis of their model, we simply note that De long et al., (1990) adopt a similar approach to ours by including an extended period (Period 0)[149] in their model. Herewith, period 0 occurs immediately preceding the trading round. De long et al., (1990) provide a neat summary of the function of period 0. They write:

[149] Theoretical microstructure models typically begin with period 1.

> Period 0 is a reference period. The price is set at its initial fundamental value of zero, and there is no trading. Period 0 provides a benchmark against which traders can measure the appreciation or depreciation of stock from period 0 to period 1…Since there is no trading in period 0, market clearing conditions are automatically satisfied (p. 387).

Dugast and Foucault (2016) provide a similar example of this no-trade condition. Specifically, these authors presume that in period 0, speculators can only acquire information about the asset and do not trade until period 1.

With regards to the benchmark model, the System 2 algorithmic trader is granted an additional phase of information processing in order to account for the inherent complexities associated with filtering out noise from a signal (s) – further details will be provided below.

Consistent with the relevant literature (i.e., De long et al., 1990; Dugast & Foucault, 2016), the extended processing phase, denoted $\boldsymbol{T^{-}}$, occurs in the period immediately preceding the trading round.[150] Here, through a cumulative three-step filtration process, the System 2 algorithmic trader produces uniquely private information about the noise (e_s) in the forthcoming value relevant signal $s = v + e_s$ (recall, the System 2 algorithmic trader only obtains this value relevant signal in the trading period i.e., in period $\boldsymbol{T}$).

We refer to the extended periods ($\boldsymbol{T^{-}}$) private information as advanced K information and denote the sum of its parts K_0 (see Box 5.4 for a visual summary). We utilize parts of Kim and Verrecchia's (1994, 1997) notation here. Refer to appendix II (b) for further details. A review of

[150] The more traditional term 'period 0' could easily be misconstrued to imply that the model is a multi-period model. To avoid this, we use the term $\boldsymbol{T^{-}}$ to denote the extended processing phase.

appendix II (b) may have added benefit in that may orientate the reader with some of the informational dimensions of our analysis to follow.

In its totality, advanced K information (K_0) communicates $s - e_s$, in short

$$K_0 = s - e_s.$$

Essentially, as evidenced by the variables s and e_s, this private information concerns a future signal error/noise and says nothing about the fundamental value v itself – this is also implied by the assumption that the System 2 trader only obtains the actual v relevant signal ($s = v + e_s$) in the next period i.e., the trading round ($\boldsymbol{T}$).

Crucially, advanced K information is *only* useful in the trading period ($\boldsymbol{T}$), in conjunction with the fundamental signal itself $s = v + e_s$. Consequently, we do not view advanced K information as actionable information until the trading period ($\boldsymbol{T}$), when it can be combined with the value relevant signal. We reiterate that advanced K information (gathered in the extended processing phase $\boldsymbol{T^-}$), only becomes actionable information when the trading round starts ($\boldsymbol{T}$) and the fundamental value relevant signal ($s = v + e_s$) is received. In other words, K information i.e., $K_0 = \boldsymbol{s - e_s}$ is only used in the actual trading period, in conjunction with the signal $s = \boldsymbol{v + e_s}$.

Once the fundamental signal ($s = \boldsymbol{v + e_s}$) is obtained, advanced K information ($K_0 = \boldsymbol{s - e_s}$), gathered in the extended processing phase (T^-) can be used to correct for the error ($\boldsymbol{e_s}$) in the value relevant signal ($s = v + \boldsymbol{e_s}$).

(The aforementioned discussion remains an overview of the signal structure in the benchmark model. This will be expanded on and explored further in the relevant subsections that follow)

We discuss the extended processing phase $\boldsymbol{T^-}$ (specifically the structure of advanced K information) in more detail below. Section 5.4.3 considers $\boldsymbol{T}$, the actual trading period in more detail.

5.4.3 The Extended Processing Phase ($\boldsymbol{T^-}$)

As emphasized above, advanced K information alone is not informative about the firms liquidating value v in the extended processing phase;[151] there is no trading here. Since there is no trading in the extended processing phase, the market clearing conditions are automatically satisfied.

5.4.3.1. Advanced K information

Recall, advanced K information is acquired through a three-step cumulative information processing mechanism. Before we outline these steps, consider the following important principle: Namely, the structure of the informed trader's private information is not crucial. As articulated by Rochet and Vila (1994): "Given that all traders are assumed to be risk-neutral it is only needed that the information structures be nested i.e., that the informed trader knows more than the market" (p.132). This principle suffices as justification for our ability to modify Kyle's model with advanced K information without inducing irreconcilable dynamics.

[151] Recall that, in its totality, advanced K information (K_0) communicates $s - e$. As evidenced by the variables $\boldsymbol{s}$ and $\boldsymbol{e}$, this private information concerns a future signal error/noise and says nothing about the fundamental value $\boldsymbol{v}$ itself – this is also implied by the assumption that the System 2 trader only obtains the actual v relevant signal ($s = v + e$) in the next period i.e., the trading round ($\boldsymbol{T}$).

The steps involved in the derivation of advanced K information are outlined below.[152]

1. In step one, the System 2 algorithmic trader receives a *piece* of advanced K information in the form $K_2 = s + \delta$. The term, δ represents an idiosyncratic noise term, the type described by Kim & Verrecchia (1994) in Appendix II (by adding additional noise we intend to capture the principal that filtering information is a cumulative process that requires multiple reflective and corrective observations).

2. In step two, the System 2 algorithmic trader acquires an additional piece of K information in the form $K_1 = \delta + e_s$.

3. Finally, in step three, the algorithmic trader combines K_2 and K_1; generating K_0, uniquely private information about an error/noise in a forthcoming signal of firm value. Here K_0 communicates $s - e_s$. In other words, collectively the K_s generate private advanced extended processing period information in the form of K_0, where,

$$K_0 \equiv K_2 - K_1 = \boldsymbol{s - e_s}$$

The structure of advanced K information is presented below in Box 5.4.[153]

---

[152] Our decision to structure advance *K* information as a cumulative information gathering process follows, quite directly, from a key assumption made in cognitive science regarding System 2 cognitive processing. Particularly, in cognitive science, System 2 is seen as a high-level processor, abstracting information and expressing knowledge as production rules. The representations in this system are symbolic and unbounded, in that they are based on propositions that can be combined to form larger and more complex sets of propositions. (Sloman, 1996; Smolensky, 1988). Herein lies our rationale: Just as System 2 processes are able to 'accumulate' and 'combine' bits of information to produce novel outputs, so too, System 2 algorithmic traders in our set-up acquire advance *K* information through a cumulative process which can be combined to produce a novel output.

[153] Kim and Verrecchia (1994, 1997) follow a similar approach when it comes to modelling additional information. Consequently, we utilize parts of Kim and Verrecchia's (1994, 1997) notation here. Refer to Appendix II (b) for further details. A review of appendix II (b) may have added benefit in that may orientate the reader with some of the informational dimensions of advance *K* information.

BOX 5.4: THE STRUCTURE OF ADVANCED K INFORMATION: OUTLINING THE STEPS INVOLVED IN ITS DERIVATION.

Box 5.4 provides a visual representation of the structure and steps involved in the derivation of advanced K information., We should emphasize again that the structure of the informed trader's private information is not crucial. Under the assumption that all traders are risk neutral, it is only needed that the information structures be nested i.e., that the informed trader knows more than the market (Rochet & Vila, 1994).

Box 5.4. Structure of Advanced K Information

(1) $K_2 = s + \delta$

(2) $K_1 = \delta + e_s$

(3) $K_0 \equiv K_2 - K_1 = s - e_s$

5.4.4 The Trading Period (T)

Recall, there is only *one* trading round in the benchmark model. T denotes the trading round. In this single trading round, trading takes place after T^-, in the interval [0.1]. More formally, a security is traded in a *single* auction in in a time interval which begins at $t = 0$ and ends at $t = 1$. Consistent with the literature (e.g. Kyle, 1985), we begin by assuming that the risky asset to be traded has an *ex-post* payoff of v where the value is drawn according to $v \sim N(p_0, \sigma_v^2)$. As evidenced below, nature selects a true value v from a prior normal distribution $N(p_0, \sigma_v^2)$ for the traded asset.

Effectively, the *ex-post* liquidation value of the risky asset is its true/fundamental value and is a normally distributed random variable with initial mean (p_0)[154] and variance (σ_v^2).

Like Kyle (1985) the informed agent - the System 2 algorithmic trader in our case – receives private information in the form of a noisy signal about *v* at time $t = 0$ that satisfies $s = v + e_s$. Again, this signal communicates *v* but with a substantial amount of noise e_s. (Because of e_s this signal is only partially revealing of firm value - note again that the subscript s in e_s is used to affirm that the error relates specifically to the signal s.)

This is where advanced K information, gathered in the previous period comes to the fore. The advanced *K* information becomes actionable information now - in the trading period. Recall, in its totality, advanced K information communicates $\boldsymbol{s - e_s}$**.** Here, the System 2 algorithmic trader combines the fundamental value relevant signal s with the advanced K information (gathered in the previous period, $\boldsymbol{T^-}$) to correct for the noise in the signal; thereby forming a perfect forecast of v:

(1) $s = v + e_s$

(2) $K_0 = s - e_s$

(3) $s_2 = v$

Effectively, s_2 reflects the System 2 algorithmic trader's private information in the trading round, in the form of a revised and perfect estimate of fundamental value *v*. Most notably, our assumption that the informed trader's private information in equilibrium is perfect (i.e., can forecast without

[154] Note here that that all players are risk neutral. Chen (2016) suggests that under the assumption of risk neutrality, p_0 also typifies the initial stock price (this will become clearer shortly).

noise) is a stronger assumption than that made in Kyle (1985), where the informed trader has a weaker/less perfect signal of v - we compare our results with Kyle's at the end of this chapter.

Some Intuition for the Equilibrium:

Recall that in the trading period (time $t = 0$), the System 2 algorithmic trader combines $s = v + e_s$ with $K_0 = s - e_s$, to form a perfect forecast of the firm's end of period fundamental value *v*.

Once the algorithmic trader is perfectly informed about *v* (i.e., observes *v* without any noise), he and uninformed noise traders submit market orders to the market maker to be executed at a single market-clearing price p. These submitted orders are of two kinds: the order from the informed trader, x, and orders from the uninformed traders, u. The u component is a normally distributed random variable that is independent of v, with mean zero and variance σ_u^2. We reiterate that u is determined exogenously, simplifying inference.

The market maker sets a price p and trades the quantity necessary to make the markets clear. The prices determined by the market maker are assumed to equal the expectation of the liquidation value of the commodity, conditional on the market maker's information sets at the dates the prices are determined. Their information consists of observations of the current and past quantities traded by the informed and uninformed traders combined. Like Kyle (1985) we call these aggregate quantities the 'order flow'. The aggregate order flow is denoted by y, where $y = x + u$.

Contextually, conditional on y, the market maker determines a price, p, at which he will clear the entire order. That is,

$$p = P(y)\,, \qquad y = x + u$$

As in Kyle (1985), the market makers pricing rule is pinned down by the assumption that he expects to earn a profit of zero.

This assumption is consistent with free entry of competing market makers, a condition under which the single market maker cannot exercise monopoly power. Again, market makers are simply vote counters, not analysts of fundamentals; and the votes they count are the order flows. The assumption that market makers earn zero profit is important to the model and is shared by many other models within microstructure. Essentially, when market makers determine the price p at which they trade the quantity necessary to clear the market they observe $x + u$, but not x or u (or even v) separately.[155] Expected profit of zero implies that the market maker sets price p as a function of the sum $x + u$ such that

$$p = E[v|x + u] \tag{5.4.1}$$

We reiterate that based on the generalizations relating to System 2 algorithmic traders (which we have modelled explicitly above); we have suggested that the System 2 algorithmic trader observes v noiselessly in equilibrium – this assumption follows directly from the assumptions made in our final dual-process taxonomy of algorithmic trading – a taxonomy that was supported by an instantiation of the concept and consolidated with empirical evidence on algorithmic trading activities. It follows, that in the trading period, the System 2 algorithmic trader's problem is to determine the optimal purchase (or sale) of quantity x.

His objective function is given by

$$\max_{x} E[\pi|v] = E[(v - p)x] \tag{5.4.2}$$

[155]Here x simply denotes the System 2 algorithmic traders order, his strategy however is denoted X.

Our equilibrium characterisation follows Kyle (1985). Like Kyle (1985) we focus on a Bayesian Nash Equilibrium (BNE), where all strategies are linear in equilibrium (namely, we conjecture linear strategies for both the System 2 algorithmic trader and the market maker and then verify that these conjectures are actually the best response to one another's strategies).[156] Section 4.6.2.1 highlights the rationale for focusing on a linear equilibrium.

We obtain a Bayesian Nash Equilibrium by considering (1) that the informed System 2 algorithmic trader chooses a demand function that maximises expected profits,[157] given his expectation of the impact of his order on the market price and (2) that market-makers set prices based on their Bayesian interpretation of the information they possess. Precisely, we show that there is an equilibrium in which the System 2 algorithmic trader's quantity x is a linear function of his private information on v, and the equilibrium asset price p chosen by the market maker is a linear function of $x + u$.

Below we provide an overview of the equilibrium concept. We highlight some key assumptions and define the equilibrium conditions. While the overview does not provide detailed proofs (proofs are covered in Section 5.5), it does attempt to outline the important steps and highlight the key intuition. By clarifying the underlying economics as effectively as possible, the overview serves as a premise from which we can comprehensively analyse the benchmark model. The benchmark model is formally introduced in Section 5.5.

---

[156] Because all conjectures will be linear combinations of functions in which the coefficients are 'constants to be determined', this whole approach to finding particular solutions is formally called the method of undetermined coefficients.

[157] The System 2 algorithmic trader is risk neutral. Thus, risk is not considered in the optimal behaviour of the agent as maximising expected profits and expected utility are equivalent (De-Jong & Rindi, 2009). Specifically, under the assumption of risk neutrality, the expected *utility* of end-of-period profits is simply equal to expected end-of-period profits. Consider the following proof: With an exponential utility function, we obtain: $Max\ E[U(\pi_i)] = Max\left[E(\pi_i) - \left(\frac{A}{2}\right)Var(\pi_i)\right]$. Adding risk neutrality ($A = 0$), maximisation of expected utility simplifies to maximisation of expected profits.

Definition 5.1 (Equilibrium conditions). The Nash equilibrium studied in the benchmark model from time $t = 0$ onwards is formally defined as two functions, a trading strategy $X(\cdot)$ and a pricing rule $P(\cdot)$, satisfying a profit maximisation condition and a market efficiency condition. (Note that x denotes the System 2 algorithmic trader's traded *quantity* and $X(\cdot)$ is his trading *strategy*. Likewise, p denotes the *price* set by the market maker and $P(\cdot)$ the market maker's pricing '*rule*'[158] – the notation is Kyle's).

The profit maximisation condition states that the quantity traded by the System 2 algorithmic trader $x = X(\cdot)$, maximises the System 2 algorithmic trader's expected profits, taking the pricing rule $P(\cdot)$ as given, i.e.,

$$x = X(v) = \max_{x} E[x(v - P(y))|v].$$

The market efficiency condition states that the market maker's expected profits equal to zero, conditional on observing the order flow, and taking the System 2 algorithmic traders strategy as given, i.e.,

$$p = P(y) = E[v|X(v) + u = y].$$

Contextually, an equilibrium is defined to be an $X(v)$ and $P(y)$ which simultaneously satisfy conditions/equations (5.4.1) and (5.4.2).

Like Kyle, we rely on the method of undetermined coefficients to solve for equilibrium. Namely, the coefficients β and λ will only be determined in equilibrium.

[158] The terms 'rule' and 'strategy' are used interchangeably throughout.

In equilibrium these coefficients are given by

$$\lambda^* = \frac{1}{2}\sqrt{\sigma_v^2/\sigma_u^2} \tag{5.4.3}$$

and,

$$\beta^* = \sqrt{\sigma_u^2/\sigma_v^2} \tag{5.4.4}$$

Outline of Proof:

Our proof starts by showing that a linear trading strategy (chosen by the System 2 algorithmic trader) and the linear pricing rule (chosen by the market maker) are mutual best responses. The problem reduces to two best response equations with two unknown parameters (β, λ) which yield (5.4.3) and (5.4.4) respectively. Following an outline of the proof, properties of the equilibrium are discussed.

Begin with the assertion that the market maker prices linearly. That is,

$$P(y) = p_0 + \lambda(x + u). \tag{5.4.5}$$

Given his revised and perfect information on v, the System 2 trader's problem becomes

$$\max_x E[\pi|v] = E[(v - p)x] \tag{5.4.6}$$

$$s.t. \quad p = p_0 + \lambda(x + u).$$

Inserting the constraint into the objective function yields the following quadratic objective:

$$\max_x E[\pi|v] = E[(v - p_0)x - \lambda x^2 - \lambda xu|v]$$

$$= (v - p_0)x - \lambda x^2 \tag{5.4.7}$$

where the assumption that $E(u) = 0$ eliminates the last term. Profit is maximised when

$$X(v) = \frac{1}{2\lambda}(v - p_0) = \beta(v - p_0) \quad (5.4.8)$$

where,

$$\beta = \frac{1}{2\lambda} \quad (5.4.9)$$

Plugging $X(v)$ back into the objective function yields an expected profit function for the System 2 algorithmic trader of

$$\pi^{S_2} = \frac{(v-p_0)^2}{4\lambda} \quad (5.4.10)$$

Evidently, $E[\pi|v] = \frac{(v-p_0)^2}{4\lambda}$ characterises the System 2 algorithmic trader's conditional profits.[159] Thus, we have shown what we have set out to show: given a pricing rule which is a linear function of the market order, the informed System 2 trader optimally plays a strategy which is linear in $v - p_0$.

Conversely, suppose that the informed System 2 trader employs the proposed linear strategy. Then the market maker's pricing rule is given by

$$p = E[v|y] = E[v|\beta(v - p_0) + u] \quad (5.4.11)$$

As per standard game theoretic nomenclature, the symbol " $|$ " is used to denote "conditional upon" (Rasmusen, 2001).

[159] The System 2 trader's *unconditional* profits are given by $E(\pi) = \frac{\sigma_v^2}{4\lambda}$. The difference between his unconditional profits $E(\pi) = \frac{\sigma_v^2}{4\lambda}$ and his conditional profits $E[\pi|v] = \frac{(v-p_0)^2}{4\lambda}$ is as follows: unconditional profits increase with asset value variance, but conditional profits decrease with it. This is because, unconditionally, a high asset value variance makes it more likely that the difference between v and p_0 will be large. Conditionally, this variance does not affect the observed difference, but still affects the market maker's reaction to the observed order volume.

Like Kyle, we assume that *v* and *u* are uncorrelated and normally and independently distributed. (Such assumptions are typical in models based on Kyle's framework and are made with little or no theoretical justification – a distribution is chosen because it 'fits' and /or it is tractable (McCarthy, 2000)).

Nevertheless, being a linear combination of normals, *y* is also normal and its distribution given by

$$y \sim N(0, \beta^2, \sigma_v^2 + \sigma_u^2) \tag{5.4.12}$$

The value (v) and the total market order (y) are therefore bivariate normal and are positively correlated provided that β is positive. Specifically,

$$cov(v, y) = cov(v, \beta(v - p_0) + u) = \beta\sigma_v^2 \tag{5.4.13}$$

A convenient property of the bivariate normal distribution is that the expectation of v conditional upon y is linear. In particular, the net order flow provides a signal of the asset value v, so that the market maker can use the noisy signal y to form their expectation of v. The resulting conditional expectation $E[v|y]$ will generally differ from their unconditional expectation p_0. As the aggregate order size from noise traders' u is normally distributed and independent of v, the expected value of v conditional on y is provided by the projection theorem:[160]

[160] The projection theorem is very useful for deriving the conditional mean and variance. The proof of the projection theorem can be found in almost any statistics book (see for example, Goldberger, 1991). Here, the projection theorem follows quite directly from the ordinary least squares regression rule. Consider two joint normally distributed random variables, $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$, and denote their covariance σ_{xy}. A property of the bivariate normal distribution is that the conditional density of Y given $X = x$ is itself normal with conditional mean: $E[Y|x] = \mu_y + \frac{\sigma_{xy}}{\sigma_x^2}(x - \mu_x) = \left(\mu_y - \frac{\sigma_{xy}}{\sigma_x^2}\mu_x\right) + \frac{\sigma_{xy}}{\sigma_x^2}x$, which is the predicted value of Y from an ordinary least squares (OLS) regression of the equation $Y = a + bX$, upon setting the explanatory variable $X = x$. The slope coefficient $\frac{\sigma_{xy}}{\sigma_x^2}$ is precisely the OLS estimate of b. In our case the expected value of v conditional on y is provided by the projection theorem: $E[v|y] = E(v) + \frac{cov(v,y)}{var(y)}(y - E(y))$.

$$p = E[v|y] = E(v) + \frac{cov(v,y)}{var(y)}(y - E(y))$$

$$= p_0 + \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}\, y \qquad (5.4.14)$$

$$= p_0 + \lambda y$$

where,

$$\lambda = \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2} \qquad (5.4.15)$$

Hence, the market maker's best response is a linear strategy.

The appealing simplicity of the model is seen to derive from the assumption that v and u are each normally distributed – such assumptions have become standard in Kyle type models. We follow convention here. Contextually, provided that the System 2 trader's order x is linear in v, then, given v and u are normally and independently distributed, the total market order, $y = x + u$, will be a linear combination of normals, implying that it will be normally distributed as well. That v and y are bivariate normal guarantees that the pricing rule - the expectation of the asset's value v conditional upon the market order y - will be a linear function of the market order. Concurrently, a linear pricing rule implies that the System 2 algorithmic trader's profit was a quadratic function of his order, which in turn, guarantees that his optimal order, $X(v)$, would be linear in v.

The equilibrium solutions for λ^* and β^* are obtained by solving the following system of two equations and two unknowns:

$$\begin{cases} \lambda = \dfrac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2} \\ \beta = \dfrac{1}{2\lambda} \end{cases}$$

which gives:

$$\begin{cases} \lambda^* = \dfrac{1}{2}\sqrt{\sigma_v^2/\sigma_u^2} \\ \beta^* = \sqrt{\sigma_u^2/\sigma_v^2} \end{cases}$$

This verifies that the equilibrium values of λ and β are those shown in equations (5.4.3) and (5.4.4) respectively. This concludes the overview of the benchmark model. For easy reference, we close our overview with a visual representation of the benchmark models timeline – Figure 5.4.

Formal analysis of the benchmark model follows in Section 5.5. Section 5.5 also demonstrates exactly how the benchmark model nests the Kyle (1985) model.

FIGURE 5.4: TIMELINE OF THE BENCHMARK MODEL

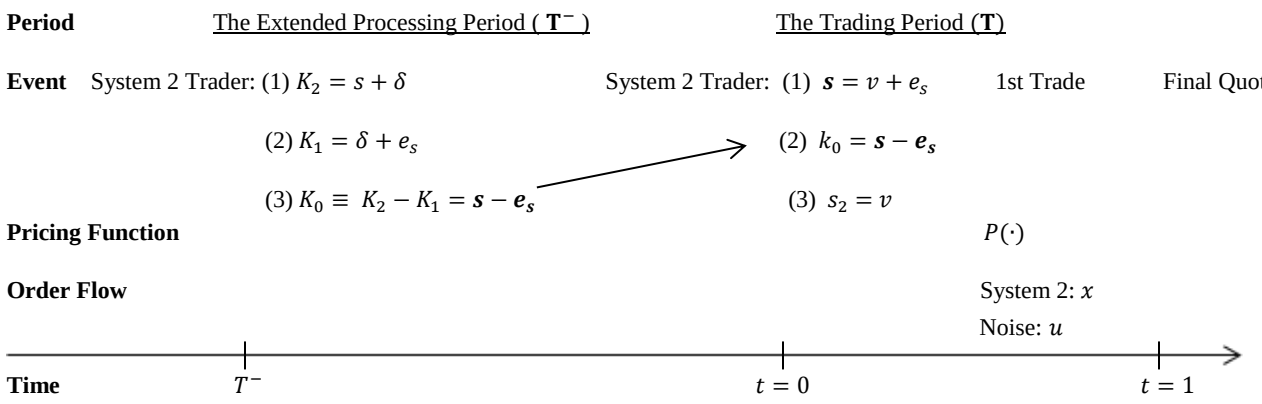


Figure 5.4. Timeline of the benchmark model with a single System 2 algorithmic trader. There are two periods in the benchmark model: (a) the extended processing period/phase and (b) the trading period. T^- characterises the extended processing period. Time T represents the trading period.

To account for the reflective/effortful processing characteristics of System 2 algorithmic traders in our taxonomy, we include an additional phase of information processing in our benchmark model. The extended processing phase, denoted T^- , occurs in the period immediately preceding the trading round. Here, through a cumulative three-step filtration process, the System 2 algorithmic trader produces uniquely private information in the form of advanced *K* information (Box 5.4 in Section 5.4.2, provides a visual representation of the structure and steps involved in the derivation of advanced *K* information – notation included). In its totality, advanced *K* information communicates $k_0 = \boldsymbol{s} - \boldsymbol{e_s}$ and concerns the noise in a forthcoming value relevant signal However, advanced *K* information alone is not informative about the firms liquidating value v at T^-. It concerns a forthcoming signal error/noise and is only useful in conjunction with the forthcoming signal itself. Thus, there is no trading in T^- and market clearing conditions are automatically satisfied.

T denotes the trading round. Recall, there is only one trading round in the benchmark model. In this single trading round, trading takes place in the interval [0.1]. More formally, a security is traded in a single auction in a time interval which begins at $\mathrm{t} = 0$ and ends at $\mathrm{t} = 1$. The securities value at the end of trading is denoted by v which is assumed to be normally distributed with an initial mean p_0 and variance σ_v^2. Like Kyle (1985), the informed agent - the System 2 algorithmic trader in our case - receives private information in the form of a noisy signal about *v* at time $\mathrm{t} = 0$ that satisfies $s = v + e_s$. This signal communicates v but with a substantial amount of noise e_s (because of e_s this signal only partially reveals firm value.) This is where advanced *K* information, gathered in the previous period/phase comes to the fore. Advanced *K* information becomes actionable information now - in the trading period $t = 0$. Recall, in its totality, advanced K information communicates $\mathbf{s} - \boldsymbol{e_s}$.

In the trading period (T), the algorithmic trader combines the fundamental value relevant signal s with the advanced K information to correct for the noise in the signal; thereby forming a perfect forecast of v:

(1) $s = v + e_s$
$\searrow$
(2) $K_0 = s - e_s$

(3) $s_2 = v$

Effectively, s_2 reflects the System 2 algorithmic trader's private information in the trading round, in the form of a revised and perfect estimate of fundamental value v. After observing v, the algorithmic trader chooses to trade the quantity x. At the same time, noise traders choose to trade quantity u. Market makers set a price p and trade the quantity that makes markets clear. The game ends at time 1. Although the benchmark model conforms to Kyle's standard set-up, our assumption that the informed traders signal is perfect is a stronger assumption than that made in Kyle (1985). In Kyle, the informed trader observes the fundamental value v, but with some residual noise: (e). Because the System 2 algorithmic trader in our model has a perfect forecast of v (we modelled explicitly the process by which he acquires this information above) the benchmark model exemplifies type 2 information asymmetry. An elegant property of the benchmark model is thus; if we nullify advanced K information (rendering it completely useless), then the benchmark model collapses to Kyle's original (1985) model. Inevitably, the benchmark model nests the Kyle (1985) model. The assumption that the System 2 algorithmic trader observes advanced K information is the main difference with Kyle (1985). In our benchmark model, advanced K information enables the System 2 algorithmic trader to form a perfect forecast of v. However, the assumption that the algorithmic trader observes v (complete information) is not crucial. According to Rochet and Vila (1994, p.132), "under the assumption that all agents are risk neutral it is only needed that the information structures be nested i.e. that the informed trader knows *more* than the market maker". This principle suffices as justification for our ability to modify Kyle's model with advanced K information without inducing irreconcilable dynamics.

5.5 FORMAL BENCHMARK MODEL: PROOFS

We consider a benchmark model with (a) two periods, (b) two assets and (c) three types of traders.

a) There are two periods in the benchmark model: (1) the extended processing period and (2) the trading period/phase. T^- characterises the extended processing period. Time T represents the trading period. In the actual trading round T, time is indexed $t = 0$ and $t = 1$. Actual trading begins at time $t = 0$ as does our derivation of equilibrium.

b) Traders trade two assets: a risk-free asset with zero interest rate and a risky asset. The risky asset has an *ex-post* payoff of v where the value is drawn according to $v \sim N(p_0, \sigma_v^2)$.

c) Three kinds of traders exist: (1) uninformed noise traders randomly trade normally distributed $u \sim N(0, \sigma_u^2)$, shares for exogenous non-informational reasons; (2) a market maker who sets the pricing function, absorbs the residual order flow imbalances, and makes zero expected profits and (3) a System 2 algorithmic trader with private, sequential observations of the *ex-post* liquidation value of the risky asset. Evidently, the System 2 algorithmic trader replaces the informed trader in Kyle (1985). Our results are contrasted with Kyle's original model at the end of this chapter

At T^-, the extended processing period, the System 2 algorithmic trader receives advanced K information. In its totality, advanced K information communicates $K_0 = s - e_s$ and concerns the noise in a forthcoming value relevant signal (see Section 5.4.2, Box 5.4). However, advanced K information (concerning a forthcoming signal error/noise) can only be used in conjunction with the forthcoming signal itself. In other words, advanced K information alone is not informative

about the firms liquidating value v at T^{-}. There is no trading in T^{-} and market clearing conditions are automatically satisfied.[161]

In the trading period, time T (beginning at $t = 0$) the System 2 algorithmic trader receives the value relevant signal $s = v + e_s$. The signal is only partially revealing of firm value because it contains a significant amount of noise, e_s. This is where advanced K information – gathered in the previous period – comes to the fore. At time $t = 0$ the advanced K information can be used in conjunction with s, to correct for the noise in the value relevant signal. Essentially by the time trading commences, the System 2 algorithmic trader has a revised and perfect forecast of v – the fundamental liquidating value of the asset.

Recall, our Bayesian Nash Equilibrium is derived from time $t = 0$ onwards.

In the trading period T ($t = 0$, $t = 1$) the System 2 algorithmic trader determines what the time $t = 1$ realization of v will be (the process of how the System 2 determines this value is detailed above).

We emphasise here that very specific functional form assumptions underlie our equilibrium analysis.

[161]We have already explained this no-trade condition in the preceding section. Particularly, Dugast and Foucault (2016) provide a similar example of this no-trade condition. These authors presume that in period 0, speculators can only acquire information about the asset and do not trade until period 1.The rationale for this no-trade condition has also been motivated in part by De long et al., (1990) who include an initial reference period (Period 0) in their model. Contextually, Period 0 provides a benchmark against which traders can measure the appreciation or depreciation of stock from period 0 to period 1. De long et al., (1990) articulate this no trade condition precisely and state that "since there is no trading in period 0, market clearing conditions are automatically satisfied" (p. 387). Conversely, in terms of notation, we have suggested that 'period 0' could easily be misconstrued to imply that the model is a multi-period model. To avoid this, we use the term $\boldsymbol{T^{-}}$ to denote the extended processing phase in our model(s).

More precisely, by assuming (1) that the relevant random variables are normally and independently distributed and (2) that the objective function has a specific form, the benchmark model acquires a tractable linear structure. Indeed, under these assumptions (that u and v are independent random variables) we are able to justify that there is a *linear* solution for our strategic game. The formal equilibrium of the benchmark model is analysed below.

5.5.1 Formal Equilibrium of the Benchmark Model

There exists an equilibrium (X, P)*, in which the System 2 algorithmic trader's trading strategy X and the market maker's pricing rule P are linear functions. Here, the equilibrium P and X are given by:*

$$X(v) = \beta(v - p_0), \quad P(y) = p_0 + \lambda y,$$

where

$$\beta^* = \sqrt{\sigma_u^2/\sigma_v^2}$$

and:

$$\lambda^* = \frac{1}{2}\sqrt{\sigma_v^2/\sigma_u^2}$$

The proof is provided below.

To prove that there exists a linear equilibrium for this strategic game one needs to conjecture linear strategies for both the System 2 algorithmic trader and the market maker and then verify that these conjectures were actually the best response to one another's strategies.

The following caveat should be highlighted: in order to exemplify the narrative that p_0 typifies both the prior, and the initial stock price of the traded asset, some notational modifications were necessary (relative to the notation used in the overview above). These modifications were done solely for the sake of explanatory convenience and have no theoretical implications. Consider the following:

5.5.2 Computation of Equilibrium

Conjecture a linear strategy for the System 2 algorithmic trader in the form $X(v) = \alpha + \beta v$ and a linear pricing rule for the market maker in the form $P(y) = \mu + \lambda\,(y)$. Because the coefficients β and λ will only be determined in equilibrium, our analysis relies on the method of undetermined coefficients to solve for equilibrium.

Also note:

$$P(y) = p = \mu + \lambda(y)\,, \qquad y = x + u$$

and

$$X(v) = x = \alpha + \beta v$$

Accordingly, one can suffice in using equations 5.5.1 and 5.5.2 to denote the pricing and order submission rules respectively:

$$p = \mu + \lambda\,(x + u), \tag{5.5.1}$$

$$x = \alpha + \beta v, \tag{5.5.2}$$

Based on the 'generalisations'[162] relating to System 2 algorithmic traders (which we have modelled explicitly above), one can write the System 2 algorithmic trader's profits as:

$$\pi^{S_2} = (v - p)x \tag{5.5.3}$$

With (5.5.1) – (5.5.3) we find the expected profits of the System 2 algorithmic trader to be

$$E[\pi|v] = E[v - \mu - \lambda(x + u))x|v] \tag{5.5.4}$$
$$= (\, v - \mu - \lambda x)x.$$

Maximizing (5.5.4) to determine the optimal order size of the algorithmic trader gives the following first order condition:

$$v - \mu - \lambda x - \lambda x = v - \mu - 2\lambda x = 0. \tag{5.5.5}$$

Rearranging yields

$$x = -\frac{\mu}{2\lambda} + \frac{v}{2\lambda}. \tag{5.5.6}$$

Comparing coefficients with (5.5.2) we get

$$\beta = \frac{1}{2\lambda}, \tag{5.5.7}$$
$$\alpha = -\frac{\mu}{2\lambda} = -\mu\beta.$$

The second order condition for a maximum,

[162] We have suggested that the System 2 algorithmic trader observes v noiselessly in equilibrium – this assumption follows directly from the assumptions made in our final dual-process taxonomy of algorithmic trading – a taxonomy that was supported by an instantiation of the concept and consolidated with empirical evidence on algorithmic trading activities.

$$-2\lambda < 0, \tag{5.5.8}$$

states that we only have to consider positive λ. Using $p = E[v|u + x]$ we get with (5.5.1) and (5.5.2) and the results of the conditional mean of jointly normal random variables (i.e., the projection theorem):[163]

$$p = E[v|x + u] \tag{5.5.9}$$

$$= E[v] + \frac{Cov[v,x+u]}{Var[x+u]}(x + u - E[x + u])$$

$$= p_0 + \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}(\alpha + \beta p_0) + \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}(x + u).$$

By comparing coefficients with (5.5.1) we see that

$$\lambda = \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}, \tag{5.5.10}$$

$$\mu = p_0 - \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2 + \sigma_u^2}(\alpha + \beta p_0) = p_0 - \lambda(\alpha + \beta p_0).$$

Solving (5.5.7) and (5.5.10) we get with (5.5.8):

---

[163] The projection theorem is very useful for deriving the conditional mean and variance. Here, the projection theorem follows quite directly from the ordinary least squares regression rule. Consider two joint normally distributed random variables, $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$, and denote their covariance σ_{xy}. A property of the bivariate normal distribution is that the conditional density of Y given $X = x$ is itself normal with conditional mean:$E[Y|x] = \mu_y + \frac{\sigma_{xy}}{\sigma_x^2}(x - \mu_x) = \left(\mu_y - \frac{\sigma_{xy}}{\sigma_x^2}\mu_x\right) + \frac{\sigma_{xy}}{\sigma_x^2}x$, which is the predicted value of Y from an ordinary least squares (OLS) regression of the equation $Y = a + bX$, upon setting the explanatory variable $X = x$. The slope coefficient $\frac{\sigma_{xy}}{\sigma_x^2}$ is precisely the OLS estimate of b. In our case the expected value of v conditional on y is provided by the projection theorem: $E[v|y] = E(v) + \frac{cov(v,y)}{var(y)}(y - E(y))$. The proof of the projection theorem can be found in almost any statistics book (see for example, Goldberger, 1991).

$$\beta = \sqrt{\frac{\sigma_u^2}{\sigma_v^2}}, \tag{5.5.11}$$

$$\lambda = \frac{\sqrt{\sigma_v^2}}{2\sqrt{\sigma_u^2}}$$

$$\mu = p_0,$$

$$\alpha = -\beta p_0.$$

As a result we can rewrite (5.5.1) and (5.5.2) as

$$p = p_0 + \lambda(x + u), \tag{5.5.12}$$

$$x = \beta(v - p_0). \tag{5.5.13}$$

This confirms that the linear equilibrium exists and is unique.[164]

From (5.5.12) we derive λ, which measures the influence an additional unit of an order has on price:

$$\frac{\partial p}{\partial(x+u)} = \lambda \tag{5.5.14}$$

[164] See Boulatov, Kyle and Livdan (2012) for a lucid description of what it means for an equilibrium to be 'unique'.

We define a market as liquid if by placing an additional order the price does not change – this is Kyle's definition. Therefore λ measures the liquidity of the market, the closer it is to zero the more liquid the market is. Usually, $1/\lambda$ is taken as a measure of liquidity as by this definition a larger value corresponds to higher liquidity. Naturally, lambda cannot be zero.

5.6 EQUILIBRIUM ANALYSIS

In order to compute the equilibrium solution for the benchmark model, we used a concept of equilibrium that embodied the strategic interaction of traders with incomplete information. More precisely, under the model's assumptions, a Bayesian Nash Equilibrium was obtained by considering (1) that the informed System 2 algorithmic trader chooses a demand function that maximises expected profits,[165] given his expectation of the impact of his order on the market price and (2) that market-makers set prices based on their Bayesian interpretation of the information content of the aggregate order flow. Precisely, we showed that there is an equilibrium in which the System 2 algorithmic trader's quantity x is a linear function of his private information on v, and the equilibrium asset price p chosen by the market maker is a linear function of $x + u$.

As is customary (à la Kyle), we allow traders to hold conjectures on one another's strategies. Contextually, in the benchmark model, the market maker conjectures that the System 2 algorithmic trader is using a linear strategy, satisfying $x = \alpha + \beta v$. Likewise, the System 2 algorithmic trader conjectures that the market maker is using a price adjustment rule that is linear in aggregate order

[165] The System 2 algorithmic trader is risk neutral. Thus, risk is not considered in the optimal behaviour of the agent as maximising expected profits and expected utility are equivalent (De-Jong & Rindi, 2009). Specifically, under the assumption of risk neutrality, the expected utility of end-of-period profits is simply equal to expected end-of-period profits. Consider the following proof: With an exponential utility function, we obtain: $Max\ E[U(\pi_i)] = Max\left[E(\pi_i) - \left(\frac{A}{2}\right)Var(\pi_i)\right]$. Adding risk neutrality ($A = 0$), maximisation of expected utility simplifies to maximisation of expected profits.

flow: $p = \mu + \lambda\,(x + u)$. Notably, we were able to prove that if players hold the above mentioned conjectures, such conjectures are actually the best response to one another's strategies – an important property of game theoretic Nash equilibria (see Box 4.6, particularly Step 3).

In the following subsection, we analyse the benchmark equilibrium vis-à-vis each agent (i.e., from the perspective of the System 2 algorithmic trader and the market maker respectively).

The following subsection may appear a bit missionary and repetitive to some readers. However, it provides the necessary foundation for Section 5.7, where we present a non-trivial *illustrative* example of the benchmark model.

5.6.1 The System 2 Algorithmic Trader

Suppose that the System 2 algorithmic trader conjectures that the market maker uses a price adjustment rule that is linear in aggregate order flow: $p = \mu + \lambda\,(y)$ where $y = x + u$. Also, here, the System 2 algorithmic trader acquires a revised and perfect estimate of the *ex-post* fundamental value v. Knowledge of v implies that the System 2 algorithmic traders (random) profits can be expressed as $\pi^{S2} = (v - p)x$. Substituting in for the price conjecture p, and for aggregate order flow y, we get *realized* random profits of :

$$\pi^{S2} = (v - p)x$$

$$= [v - (y\lambda + \mu)]x$$

$$= \big[v - \big((x + u)\lambda + \mu\big)\big]x$$

As seen in equation (5.5.4), conditional on his perfect information on v, the System 2 algorithmic traders *expected* profits are:

$$E[\pi|v] = E[v - \mu - \lambda(x + u))x|v]$$

$$= (\, v - \mu - \lambda x)x.$$

As equation (5.5.5) demonstrates, the first order condition is $\frac{\partial E(\pi|x,v)}{\partial x} = v - \mu - 2\lambda x = 0$. The solution can thus be expresses as $x = \frac{v-\mu}{2\lambda}$. Concomitantly, the second order condition is $\frac{\partial^2 E(\pi|x,v)}{\partial x^2} = -2\lambda x < 0$. See equation (5.5.8). The System 2 algorithmic trader maximises expected profits by trading this x. As intuition suggests the System 2 algorithmic trader makes more profit that the informed trader in Kyle (1985). As will become clear from the proof below, the lower the noise in the informed (S2 algorithmic trader's) signal, the higher the profit, also the higher the S2's trading intensity β; since the S2 trader would trade more aggressively due to his perfect knowledge about the asset value v.

5.6.2 The Market Maker

Suppose that the market maker conjectures that the System 2 algorithmic trader is using a strategy satisfying the linear $x = \alpha + \beta v$. Congruent with Kyle's intuition, the market maker knows the functional form of the System 2 algorithmic trader's information and he can equate this to his conjecture: $\frac{v-\mu}{2\lambda} = x = \alpha + \beta v$. This holds for all v. Therefore, $\alpha = -\frac{\mu}{2\lambda}$ and $\beta = \frac{1}{2\lambda}$. See equation (5.5.7). We can then solve for λ and μ.

The market maker sets a price p and trades the quantity that makes markets clear. As noted above, the market maker observes the aggregate order flow as y. The market maker knows $y = x + u$, but does not view either x or u individually. Given this information, the market maker sets an equilibrium market price p that he expects to give him zero expected profits.

This zero profit assumption satisfies the market efficiency condition described above (that is, definition 5.1 (Equilibrium conditions)). Specifically, the market efficiency condition states that the market maker's expected profits equal to zero,[166]conditional on observing the order flow, and taking the System 2 algorithmic traders strategy as given, i.e.,

$$p = P(y) = E[v|X(v) + u = y].$$

Expected profit of zero implies that the market maker sets the equilibrium price p as a function of $x + u$ such that: $p = E[v|y]$. [167]

Provided that the informed trader's order x is linear in v, then, given v and u are normally distributed, the total market order, $y = x + u$, will be a linear combination of normals, implying that it will be normally distributed as well. That v and y are bivariate normal guarantees that the pricing rule - the expectation of the asset's value v conditional upon the market order y - will be a linear function of the market order. Consequently, y is said to provide a noisy signal of v. In particular, the net order flow provides a signal of the asset value v, so that the market maker can

[166] This zero profit assumption is consistent with the free entry of competing market makers; whereby no single market maker can exercise monopoly power. The underlying Bertrand competition with potential rival market makers is not explicitly modelled in Kyle (1985). Rather Kyle (1985) suggests that the market efficiency condition could be replaced with an explicit Bertrand auction between at least two risk neutral bidders, each of whom observes the order flow $x + u$ and nothing else. The results of this explicit procedure would be the market efficiency condition, in which profits of market makers are driven to zero. Since the implied Bertrand competition between market makers is sufficient enough reason for Kyle (1985) to presuppose a single market maker who expects to earn zero profits, it should also suffice for our analysis here. Notably, in our model, a single market maker assumed to set semi-strong informationally efficient prices that give him zero expected profits. Although modelling how market makers can earn the positive frictional profits necessary to attract them into the business of market making is an interesting topic, Kyle (1985) argues that such analysis would "take us away from the main objective of studying how price formation is influenced by the optimising behaviour of an informed trader in a somewhat idealized setting" (p.1318).

[167] This equation can also be expressed as $p = E[v|x + u]$. Notice from the equations above that the market maker's pricing function depends on the System 2 algorithmic trader's demand function $X(v)$. If there were only noise traders, then net demand would be $y = u$. The pricing function that results would reduce to $P(y) = p = E[v|u]$ and the market would be infinitely deep.

use the noisy signal y to form their expectation of v. The resulting conditional expectation $E[v|y]$ will generally differ from their unconditional expectation p_0. As the aggregate order size from noise traders' u is normally distributed and independent of v, the expected value of v conditional on y is provided by the projection theorem. Here, the projection theorem follows quite directly from the ordinary least squares regression rule (See for example, Goldberger, 1991):

Consider two joint normally distributed random variables, $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$, and denote their covariance σ_{xy}. A property of the bivariate normal distribution is that the conditional density of Y given $X = x$ is itself normal with conditional mean: $E[Y|x] = \mu_y + \frac{\sigma_{xy}}{\sigma_x^2}(x - \mu_x) = \left(\mu_y - \frac{\sigma_{xy}}{\sigma_x^2}\mu_x\right) + \frac{\sigma_{xy}}{\sigma_x^2}x$, which is the predicted value of Y from an ordinary least squares (OLS) regression of the equation $Y = a + bX$, upon setting the explanatory variable $X = x$. The slope coefficient $\frac{\sigma_{xy}}{\sigma_x^2}$ is precisely the OLS estimate of b. In our case the expected value of v conditional on y is provided by the projection theorem: $E[v|y] = E(v) + \frac{cov(v,y)}{var(y)}(y - E(y))$.

Therefore,

$$p = E[v|y] \qquad (5.5.9\ \text{A})$$

$$= E(v) + \frac{cov(v, y).[y - E(y)]}{var(y)}$$

$$= p_0 + \frac{cov(v, u + x).[y - E(u + x)]}{var(u + x)}$$

$$p_0 + \frac{cov(v, u + (\alpha + v\beta)).[y - E(u + (\alpha + v\beta))]}{var(u + (\alpha + v\beta))}$$

$$p_0 + \frac{\beta\sigma_v^2.(y-\alpha-p_0\beta)}{\sigma_u^2+\beta^2\sigma_v^2} \qquad (5.5.9.1\ A)$$

$$= \left(\frac{\beta\sigma_v^2}{\sigma_u^2+\beta^2\sigma_v^2}\right)y + p_0 + \frac{\beta\sigma_v^2.(-\alpha-p_0\beta)}{\sigma_u^2+\beta^2\sigma_v^2}$$

$$= \left(\frac{\beta\sigma_v^2}{\sigma_u^2+\beta^2\sigma_v^2}\right)y + \frac{-\alpha\beta\sigma_v^2+p_0\,\sigma_u^2}{\sigma_u^2+\beta^2\sigma_v^2}$$

Thus, $\lambda = \frac{\beta\sigma_v^2}{\beta^2\sigma_v^2+\sigma_u^2}$, and $\mu = \frac{-\alpha\beta\sigma_v^2+p_0\,\sigma_u^2}{\sigma_u^2+\beta^2\sigma_v^2}$. See equation (5.5.10) for an analogous expression.

We have already deduced in the overview above, in equation (5.5.7), that $\alpha = -\frac{\mu}{2\lambda}$ and $\beta = \frac{1}{2\lambda}$. Therefore, we were able to deduce in equation (5.5.11) that:

$$\alpha = -\frac{\sqrt{\sigma_u^2}\,p_0}{\sqrt{\sigma_v^2}}$$

$$\mu = p_0$$

$$\lambda = \frac{\sqrt{\sigma_v^2}}{2\sqrt{\sigma_u^2}}$$

$$\beta = \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}$$

By implication we were able to express equilibrium in equation (5.5.12) and (5.5.13) as:

$$p = p_0 + \lambda(x + u)$$

$$x = \beta(v - p_0)$$

Equilibrium can also be expressed using the analogous equations (5.5.12 A) and (5.5.13 A) whereby;

$$p = p_0 + \lambda(y) = p_0 + \frac{\sqrt{\sigma_v^2}}{2\sqrt{\sigma_u^2}}\, y \qquad (5.5.12\ \text{A})$$

$$x = \alpha + \beta v = \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v - p_0) \qquad (5.5.13\ \text{A})$$

5.6.3 Properties of the Solution

The System 2 algorithmic trader's order flow x satisfies $sgn(x) = sgn(v - p_0)$. If the signal exceeds the unconditional prior, the System 2 trader buys, and vice versa. The more uncertainty there is surrounding the noise traders' order flow, the larger is the System 2 algorithmic trader's trade *ceteris paribus*. In other words, the greater is the amount of noise trading σ_u^2, the larger is the magnitude of the order submitted by the System 2 algorithmic trader for a given deviation of v from its unconditional mean. Hence, the System 2 algorithmic trader trades more actively the greater is the 'camouflage' provided by noise traders. More noise trading makes it more difficult for the market maker to extract the signal of the System 2 algorithmic trader from the noise.

Note that

$$p = p_0 + \lambda(y)$$

$$= p_0 + \lambda(x + u)$$

$$= p_0 + \lambda[\frac{1}{2\lambda}(v - p_0) + u]$$

$$= \frac{v + p_0}{2} + \lambda u.$$

That is, the market maker's market clearing price is the average of the prior price p_0 and the signal v plus the uninformed order flow u times the (inverse) liquidity parameter. So price is average of prior plus signal plus the move caused by liquidity traders as a function of the degree of liquidity.

Profits

From the model we see that the System 2 algorithmic trader's realised profit is:

$$\pi^{S2} = (v - p)x$$

$$= \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}\,(v - p_0)[v - \left(\frac{\sqrt{\sigma_v^2}}{2\sqrt{\sigma_u^2}}\,y + p_0\right)]$$

$$= \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}\,(v - p_0)^2 - \frac{1}{2}y(v - p_0)$$

$$= \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}\,(v - p_0)^2 - \frac{1}{2}y(x + u)(v - p_0)$$

$$= \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}\,(v - p_0)^2 - \frac{1}{2}[\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v - p_0) + u](v - p_0)$$

$$= \frac{1}{2}\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v - p_0)^2 - \frac{1}{2}u(v - p_0)$$

It follows that the System 2 algorithmic trader's conditional profit is:

$$E[\pi|v] = \frac{1}{2}\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v - p_0)^2$$

Moreover, his ex-ante 'unconditional profit'[168] is:

$$E(\pi) = \frac{1}{2}\sqrt{\sigma_u^2}\sigma_v^2$$

where,

$$E[\pi|v] = \frac{1}{2}\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v - p_0)^2$$

$$= \frac{(v - p_0)^2}{4\lambda}$$

That is, conditional expected profit is linear in depth (1/ λ) because double the depth doubles the System 2 algorithmic traders opportunity for profit. Notice that compared to Kyle's original formulation (as seen in Appendix III) , β is larger, and the System 2 algorithmic trader trades more aggressively. At the same time, λ increases and depth decreases. As before, the market maker loses on trades with the informed (System 2 algorithmic trader) and gains money on trades with the uninformed traders. Hence, relative to Kyle, both trading profits and costs are higher in our

[168] The System 2 trader's *unconditional* profits are given by $E(\pi) = \frac{1}{2}\sqrt{\sigma_u^2}\sigma_v^2$. The difference between his unconditional profits $E(\pi) = \frac{1}{2}\sqrt{\sigma_u^2}\sigma_v^2$ and his conditional profits $E[\pi|v] = \frac{(v-p_0)^2}{4\lambda}$ is as follows: unconditional profits increase with asset value variance, but conditional profits decrease with it. This is because, unconditionally, a high asset value variance makes it more likely that the difference between v and p_0 will be large. Conditionally, this variance does not affect the observed difference, but still affects the market maker's reaction to the observed order volume. More details are given below in the proceeding Section 5.6.4 (interpretation of equilibrium).

formulation (Refer to Appendix III for detailed analysis of Kyle's original solution). We elaborate further below.

5.6.4 Interpretation of Equilibrium

From our solution to the equilibrium order submitted by the System 2 algorithmic trader (equation 5.5.13 A), we can see that the greater the amount (volatility) of noise trading, the greater the magnitude of the order submitted by the System 2 trader for a given deviation of v from its unconditional mean p_0. Hence, the informed System 2 algorithmic trader trades more actively on his private information the greater is the 'camouflage' provided by uninformed traders. More uninformed noise trading makes it more difficult for the market maker to extract information from noise. Also consider the following: If equation (5.5.13 A) is substituted into equation (5.5.12 A) one obtains

$$p = p_0 + \frac{1}{2}\frac{\sqrt{\sigma_v^2}}{\sqrt{\sigma_u^2}}u + \frac{1}{2}(v - p_0) \tag{5.6.4}$$

$$= \frac{1}{2}\left(\frac{\sqrt{\sigma_v^2}}{\sqrt{\sigma_u^2}}u + p_0 + v\right)$$

Notably, one-half of the System 2 algorithmic trader's perfect private information on v, (namely; $\frac{1}{2}v$) is reflected in the equilibrium price, so that, like Kyle (1985), the price is *not fully revealing*. (A fully revealing price would be $p = v$.) This insight is qualified below. Before we substantiate further, a review the difference between the System 2 algorithmic trader's conditional and unconditional profits is provided:

Using (5.5.13 A) and (5.6.4) we show now that the System 2 trader's expected profits can be expressed as

$$E[\pi] = E[x(v-p)] = E[\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v-p_0)\frac{1}{2}\left(v - p_0 - \frac{\sqrt{\sigma_v^2}}{\sqrt{\sigma_u^2}}u\right)]$$

As emphasised in the overview of the model, at the beginning of the trading period T 'conditional'[169] on knowing v (that is after using advanced K information to forecast v perfectly), the informed System 2 trader expects profits of

$$E[\pi|v] = \frac{1}{2}\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v-p_0)^2$$

$$= \frac{(v-p_0)^2}{4\lambda}$$

Hence, the larger is v's deviation from p_0, the larger the System 2's expected profit. Moreover, unconditional on knowing v, that is prior to the start of trading, the System 2 trader expects a profit of

$$E[\pi] = \frac{1}{2}\frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}E[(v-p_0)^2]$$

$$\frac{1}{2}\sqrt{\sigma_u^2}\sigma_v^2.$$

[169] Recall, as per standard game theoretic nomenclature, the symbol " $|$ " is used to denote "conditional upon" (Rasmusen, 2001).

5.6.5 Information Revelation

In line with convention, we use the variance of the market maker's posterior distribution for v to formally measure how much of the informed System 2 algorithmic trader's perfect information is revealed through trading.

Initially, the market maker has an expectation that is distributed about v with a variance we will denote as σ_0^2 – this is the prior variance. However, after observing the aggregate order flow (i.e., $= x + u$), the market maker's expectation is distributed about v with a variance σ_1^2; whereby,

$$\sigma_1^2 = \frac{1}{2}\sigma_0^2 \qquad (5.6.5)$$

The result above (equation 5.6.5) was obtained using Bayes' rule for updating conditional variances.[170] A succinct exposition of this result (i.e., its derivation) is provided below. Implications are also noted. Consider the following:

Bayes' formulation for the determination of conditional variances of jointly distributed random variables can be expressed as:

$$var(Y|X) = var(Y) - \frac{[cov(X,Y)]^2}{var(X)}$$

For the purpose of the benchmark model, this formula can be restated as:

$$Var[v|p] = Var[v] - \frac{Cov[v,p]^2}{Var[p]}$$

Now, define $\sigma_0^2 = \sigma_v^2$ and

$$\sigma_1^2 = var[v|p]$$

[170] The proof of the Bayes' theorem can be found in almost any statistics book. Therefore, we use it here without derivation. To see its formal derivation, we refer readers to the appropriate literature (i.e., Maurer & Ralston, 2005; McCarthy, 2000; Nanda, 2002; Platen & Heath, 2006).

It is then simple to show using the formulation above that ($\sigma_1^2 = \frac{1}{2}\sigma_0^2$).

Notably, the updated variance is exactly one-half of the prior variance. This means that informed System 2 trader's strategy results in exactly half of his private information being revealed by the market price; that is, the new conditional variance of the true value of the stock is only half of the original unconditional variance.[171]

Interestingly, like Kyle (1985), half the informed trader's private information has been revealed through trading. Perhaps more importantly however, because the System 2 algorithmic trader has more accurate information on *v* than the informed trader in Kyle (1985); the 2 algorithmic trader in our benchmark model effectively reveals more accurate information in the trading process than the informed trader in Kyle (1985). The illustrative example below should clarify this concept.

5.7 AN ILLUSTRATIVE EXAMPLE OF THE BENCHMARK MODEL

Suppose that the unconditional value of a stock in the benchmark framework is normally distributed with a mean of R50 and a variance equal to 30. (We use the local South African currency R (Rand) to denote value here. However, the economic role of currency is trivial and (R) was used solely for illustrative convenience. We could have very easily used $ to denote value in this example)

We begin by considering the following scenario:

[171] The intuition for this result starts with the fact that we expect the variance of asset price conditional upon observing the order flow to be lower than the unconditional variance σ_0^2 (or σ_v^2) because some of the informed trader's signal should have been revealed through trading. One would not expect σ_1^2 to collapse to zero: In that case, the System 2 would make no profit, because all the trades would clear at the perfectly revealing price. Nor would one expect σ_1^2 to remain at σ_0^2: In that case the market maker learns nothing, and the informed trader could make infinite profits.

Suppose that in T^{-} (the the extended processing phase/period), the System 2 algorithmic trader obtains and observes uniquely private information about the noise in a forthcoming, value relevant signal (this value relevant signal satisfies $s = v + e_s$, and is set for a forthcoming period) We define the extended processing phase/period private information as advanced K information (advanced K information comprises of K_2, K_1 and K_0). In its totality this information communicates $s - e_s$. The structure of K information is presented above in Chapter 5, Section 5.4.2 (see particularly, Box 5.4). As emphasized throughout, this information is about the noise in a forthcoming signal and is not informative about the firms liquidating value v at phase (T^{-}).

T denotes the trading period. Here a single auction occurs in the interval $t = 0$ and $t = 1$. As referenced above, the System 2 trader receives a value relevant signal in the form $s = v + e_s$ at time $t = 0$. The signal is only partially revealing of firm value because it contains a significant amount of noise e_s. This assumption is consistent with prior microstructure models of trading (See e.g. Grossman & Stiglitz, 1980; Kyle, 1985).

However, in contrast to prior microstructure models of trading, the informed agent in our model is able to correct for the noise in this signal; forming a revised and perfect estimate of the *ex-post* fundamental value of the risky asset v. Conceptually, the information gathered by the System 2 algorithmic trader in the previous period/phase (i.e., advanced K information) is used to provide a context or interpretation to the signal received in the trading period. In other words, at time $t = 0$ the advanced K information (gathered in the extended processing phase) can be utilised, in conjunction with s, to correct for the noise (e_s) in the signal $s = v + e_s$. (See Section 5.4.2)

Consider the following illustration:

5.7.1 Equilibrium Illustration

As previously indicated, our illustrative example considers a normally distributed stock with a mean of R50 and a variance equal to 30.

In line with conventional approaches, the informed agent receives private information in the form a signal about the *ex-post* liquidating value of the stock. This signal satisfies $s = v + e_s$. For illustrative convenience let us assume that this signal communicates $s = v + e_s \equiv 44.5$. Let us also assume (again, only for illustrative purposes) that $0 < e_s < 1$. In the majority of microstructure models of trading this simplified, one-shot signal $s = v + e_s$ would be the informed trader's private information. The reason is that $s = v + e_s \equiv 44.5$ $(0 < e_s < 1)$ provides a better estimate of the time $t = 1$ realization of *v* than its initial mean of R50.

However, we depart from this traditional, single-signal approach by assuming that, in addition to this one-shot signal $s = v + e_s$, the informed agent in our model has access to other forms of private information; information that can be combined and filtered into a perfect forecast of an assets true/fundamental value *v*. Indeed, consistent with the reflective/effortful processing characteristics of System 2 algorithmic traders in our taxonomy, we include an additional phase of information processing in our model. Again, the System 2 algorithmic trader is granted this additional phase of information processing in our model to account for the inherent complexities associated with filtering out noise from a signal. The extended processing phase, denoted T^-, occurs in the period immediately preceding the trading round. Here, through a cumulative three-step filtration process, the System 2 algorithmic trader produces uniquely private information about the noise (e_s) in the forthcoming value relevant signal $s = v + e_s$. Knowledge of e_s in the trading period means that the System 2 algorithmic trader observes *v* perfectly in equilibrium. In

other words, the System 2 algorithmic trader determines that the fundamental value of the stock is actually $v =$ R45 per share.

Barring these additional assumptions, the System 2 traders signal would have simply been $s = v + e_s \equiv 44.5$.

We are now in a better position to illustrate the equilibrium of the benchmark model. Let us recap:

- The securities value at the *end of trading* is denoted by *v* which is assumed to be normally distributed with a mean p_0 (R50) and variance σ_v^2 (30)

- Give the modelling assumptions mentioned above, the System 2 algorithmic trader determines that the fundamental value of the stock is actually $v =$ (R45) per share.

- Suppose that uninformed investor trading is random and normally distributed with an expected net share demand of zero σ_u^2 equal to say, 5000.

- The market maker can observe the total level of order volume $x + u$ where u reflects noise trader transactions and x reflects informed demand but cannot distinguish between x and u.

What would be the System 2 algorithmic traders demand for the stock?

We determine the level of the System 2 traders demand by solving for x using the following equation (5.5.13 A):

$$x = \alpha + \beta v = \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}}(v - p_0)$$

$$= \frac{\sqrt{5000}}{\sqrt{30}}(45 - 50) = -64.5497$$

With $p_0 = (R50)$ and $v = (R45)$, the System 2 algorithmic trader would like to sell as many shares as possible in order to earn a R5 profit on each, but cannot because the market maker would correctly infer that his share sales convey meaningful information.

At what level does the market maker set his price?

Given the total demand that the market maker observes $(x + u) = -64.5497 + 0,$ we can calculate the level that he sets the price using the following parameters:

$$\alpha = -\frac{\sqrt{\sigma_u^2}\ p_0}{\sqrt{\sigma_v^2}} = -\frac{\sqrt{5000}.50}{\sqrt{30}} = -645.49$$

$$\beta = \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}} = \frac{\sqrt{5000}}{\sqrt{30}} = 12.91$$

$$\lambda = \frac{\sqrt{\sigma_v^2}}{2\sqrt{\sigma_u^2}} = \frac{\sqrt{30}}{2.\sqrt{5000}} = 0.0387$$

Using equation (5.5.9.1 A) here for illustrative convenience, we can determine the market maker sets the price as follows:

$$p = p_0 + \frac{\beta\sigma_v^2.(y - \alpha - p_0\beta)}{\sigma_u^2 + \beta^2\sigma_v^2}$$

$$= 50 + \frac{12.91 \cdot 30 \cdot (-64.5497 + 0 + 645.49 - 50 \cdot 12.91)}{5000 + 12.91^2 \cdot 30}$$

$$= 47.50$$

Thus, the results of our illustrative example are consistent with and confirm the models solution. After a single round of trading, approximately half of the System 2 algorithmic trader's *perfect* private information ($v = 45$) is revealed by the equilibrium price.

5.8 CONCLUDING REMARKS

In this chapter we presented the benchmark model with only one type/system of algorithmic trader; a System 2 algorithmic trader with an inherent accuracy advantage.

The literature confirms, somewhat ubiquitously, that algorithmic traders outperform their human counterparts in terms of predictive accuracy. Thus, our suggestion that the accuracy advantage is exclusive to algorithmic traders is a reasonable one.

However, a crucial premise of this thesis is that algorithmic trading is not an isolated phenomenon. Thus, by including only a certain portion of algorithmic trading strategies (strategies that fall under the rubric of System 2 algorithmic trading) the benchmark model functions exclusively as an introductory concept.

Although the benchmark concept builds the initial link between our dual process construct of algorithmic trading and market microstructure, it neglects crucial interactions underlying our dual process formulation. The overarching dual process microstructure model of algorithmic trading, presented in the following chapter, is the focus of this thesis.

CHAPTER 6

ALGORITHMIC TRADING, MARKET QUALITY AND INFORMATION: A THEORETICAL DUAL-PROCESS MODEL

6.1 FOREWORD

The benchmark model functions as a preparatory concept, laying the microstructural foundations for our final model. In this section we present the final model. The model, to follow, fulfills the precepts underlying our dual process decomposition of algorithmic trading.

Essentially, our taxonomy (Section 5.2) suggests a distinction between two separate types/systems of algorithmic trading. This dual process approach incorporates all known algorithmic trading strategies within a single unified microstructure framework. Inevitably, all the various strategies associated with algorithmic trading (as seen in Chapter 2, Section 2.4) correspond with their own respective system, and by implication, informational advantage. Table 6.1 is provided here for illustrative purposes.

TABLE 6.1: COMPARATIVE OVERVIEW OF THE DUAL-PROCESS MOEL

COMPONENT	**SYSTEM 1 TRADING**	**SYSTEM 2 TRADING**
Private Information	Speed Advantage	Accuracy Advantage
Strategies	• Spread Capturing Algorithms • Rebate Trading Algorithms • Time Weighted Average Price (TWAP) Algorithms • Volume Weighted Average Price (VWAP) Algorithms • Implementation Shortfall Algorithms • Adaptive Execution Algorithms • Liquidity Detection Algorithms	• Data Mining Algorithms • Text Mining Algorithms • Neural Network Algorithms • Support Vector Machine Algorithms

Our all-inclusive dual process microstructure model will look at the effects of these two algorithmic trading systems individually and concurrently (that is, their constituent elements and dynamic interactions), but within the same framework. It is hoped that this integrative model – a model that integrates and synthesizes the multitude of algorithmic trading strategies within a single workable framework – may shed light on contradictory findings, and on previously unknown market variables.

6.2 THE FINAL MODEL: INTRODUCTION

This section presents the final model with two algorithmic traders – System 1 and System 2 algorithmic traders. Our equilibrium characterisation follows the benchmark model.

It is important to note that the benchmark model *nests* the Kyle (1985) model. In a limiting case of the benchmark model, where the System 2 algorithmic trader does not have access to his particular form of private information, the equilibrium reduces to the equilibrium of the static model of Kyle (1985). Likewise, here, in the final model, when the System 1 algorithmic trader has no information, the model collapses to the benchmark model. This will be explored below.

Like the benchmark model, the final model considers two 'periods' – (1) an extended processing period/phase and (2) a single trading period. The System 2 algorithmic trader, the market maker and noise traders' have the same action timing as in the benchmark model. However, we modify the timeline of the benchmark model to accommodate the System 1 algorithmic trader. More precisely, we introduce delays in trading and quoting into the *trading period* of the benchmark model and add a new type of trader, the System 1 algorithmic trader, who is fast enough to exploit the short delays (the rationale for which will be clarified below).

Assets and Traders Traders trade two assets: a risk free numeraire asset with zero interest rate and a risky asset. We begin by assuming that the risky asset to be traded in the trading period has an *ex-post* payoff of v where the value is drawn according to $v \sim N(p_0, \sigma_v^2)$. In other words, the *ex-post* liquidation value of the risky asset is normally distributed with mean p_0 and variance σ_v^2. All traders are risk neutral.

Four kinds of traders exist: (1) uninformed noise traders randomly trade normally distributed $u \sim N(0, \sigma_u^2)$, assets for exogenous non-informational reasons; (2) market makers set the pricing function, absorb the residual order flow imbalances, and make zero expected profits; (3) System 2 algorithmic trader with private, sequential observations of the *ex-post* liquidation value of the risky asset; and (4) System 1 algorithmic trader who anticipates incoming market orders, trades rapidly and has a short holding horizon.[172]

The fourth type of trader – the System 1 algorithmic trader - is relatively new here and represents a novel addition relative to the benchmark model. Before we formally introduce the System 1 algorithmic trader, let us briefly recap the assumptions regarding the *other* three agents as they appear in benchmark model. A refresher on these other traders will provide the contextual framework for the discussion to follow. System 1 algorithmic traders are defined and discussed in what proceeds.

[172] This is just a broad overview of the agents in our model. As such, we only note the key characteristics of System 1 algorithmic trader's here. The assumptions made for the System 1 algorithmic trader are clarified by the model itself, presented in Section 6.5.

System 2 Algorithmic Traders

System 2 algorithmic traders are thoroughly addressed in the benchmark model (Section 5.4). As System 2 algorithmic traders in the final model are identical to the System 2 algorithmic traders in the benchmark model, the following discussion will be brief in nature:

Like the benchmark model, the System 2 algorithmic trader is granted an additional phase of information processing in the final model. The extended processing phase, denoted T^- , occurs in the period immediately preceding the trading round.

At T^-, the extended processing period, the System 2 algorithmic trader receives advanced K information. In its totality, advanced K information provides $K_0 = s - e_s$ and concerns the noise in a forthcoming value relevant signal (see Section 5.4.3 Box 5.4). However, advanced K information (concerning a forthcoming signal error/noise) can only be used in conjunction with the forthcoming signal itself. In other words, advanced K information alone is not informative about the firms liquidating value v at T^-. Thus, there is no trading in T^- and market clearing conditions are automatically satisfied.

Trading commences at time $t = 0$. Here the System 2 algorithmic trader receives the value relevant signal $s = v + e_s$. Again, the signal is only partially revealing of firm value because it contains a significant amount of noise e_s. This is where advanced K information – gathered in the previous period – comes to the fore. In the trading period, the advanced K information can be used in conjunction with s, to correct for the noise in the value relevant signal. Essentially by the time trading commences, the System 2 algorithmic trader has a revised and perfect estimate of v – the fundamental liquidating value of the asset.

Again, because we have already exhausted the description of System 2 algorithmic traders in the benchmark model itself, we direct interested readers to Section 5.4.3 of Chapter 5, where we delineate the precepts underlying this type of private information – advanced K information.

Uninformed Noise Traders

Uninformed noise traders are present in nearly all market microstructure models of trading.[173] Noise traders trade for liquidity, tax purposes, hedging, or other reasons which are exogenous to a given model. In line with Kyle (1985), the benchmark model, and other extant literature on the topic of market microstructure, we include uninformed noise traders in the final model.

In accordance with the benchmark model, once trading commences, uninformed noise traders randomly trade normally distributed $u \sim N(0, \sigma_u^2)$, assets for exogenous non-informational reasons. Like both Kyle's (1985) and the benchmark model, u is exogenously determined (simplifying inference).

Market Makers

There is extensive market microstructure literature on the topic of price formation in the presence of market makers and the present study forms part of this ongoing discourse. In both Kyle (1985) an the benchmark model above, an informed trader (a System 2 algorithmic trader in the benchmark model) trades against uninformed noise traders , in the presence of a price setting competitive market maker.

The market maker sets a price and trades the quantity that makes markets clear. The prices determined by the market maker are assumed to equal the expectation of the liquidation value of

[173] See Bagehot (1971) for a detailed discussion on the topic of noise trading.

the commodity, conditional on the market maker's information sets at the dates the prices are determined. Their information consists of the aggregate 'order flow' (Chapters 4 & 5).

The aggregate order flow provides a signal about the liquidation value of the asset to the market maker. Based on this signal, the market maker revises her/his beliefs and sets price such that it equals the expected liquidation value given the observed order flow. The resulting equilibrium price change is an increasing linear function of net order flow, whose slope represents a measure of the *market depth*. Although defined inconsistently in the literature (e.g., Lyons, 2001), practitioners often define market depth as an orders price impact - denoted λ.

The market maker in the final model serves the same function as those of the Kyle (1985) and the benchmark models. Thus, for the sake of the discussions to follow, let review some of the key insights that emerge from the benchmark model with regards to the market maker.

Review of the Market Maker in the Benchmark Model

In Kyle (1985) and the benchmark model, the market makers pricing rule is pinned down by the assumption that he expects to earn a profit of zero. This assumption is consistent with free entry of competing market makers, a condition under which the single market maker cannot exercise monopoly power.[174]

[174] Kyle (1985) suggests that the market efficiency condition could be replaced with an explicit Bertrand auction between at least two risk neutral bidders, each of whom observes the order flow $x + u$ and nothing else. The results of this explicit procedure would be the market efficiency condition, in which profits of market makers are driven to zero. Since the implied Bertrand competition between market makers is sufficient enough reason for Kyle (1985) to presuppose a single market maker who expects to earn zero profits, it should also suffice for our analysis here. Notably, a *single* market maker assumed to set semi-strong informationally efficient prices that give him zero expected profits.

The competition between market makers drives their (conditional) expected profits down to zero. So, the market clearing price they arrive at is just the expected price of the asset conditional upon the information they possess.[175]

The assumption that market makers earn zero profit is important to both the benchmark and the final model and is shared by many other models within microstructure. Contextually, the market maker makes a loss on the trades with the informed trader, but recoups this loss on trades with the noise traders, making zero profit on average.

Expected profit of zero implies that the market maker sets price p as a function of the sum $x + u$ such that

$$p = E[v|x + u].$$

Price depends on the sum $x + u$ because the market maker does not observe x and u individually. Again, the u component of that sum is exogenous which simplifies inference.[176] The complication emerges from the x component, which depends on the trading strategy X of the informed trader – the System 2 algorithmic trader in the benchmark model (Section 5.4 distinguishes between the informed traders trading strategy X and his order x).

[175] The underlying Bertrand competition with potential rival market makers is not explicitly modelled in Kyle (1985). However, to add some concreteness, we highlight Kyles intuition with the following example: Recall that in both the benchmark model and in Kyle (1985), the market makers receive a net order flow equal to $y = x + u$, which they can only observe in aggregate. Holding this information they compete on prices, offering a price function $p = E[v|y]$. The underlying Bertrand competition that drives the market makers' expected profit to zero can be expressed concisely, whereby; $E[\pi_M|y] = E((p(y) - v)y|y) = E((p(y) - E(v|y))y) = 0$. However, we emphasise (like Kyle) that it suffices to posit a single market maker who sets prices according to a semi-strong efficiency condition. Specifically, because Bertrand competition and the market efficiency condition are equivalent, we, like Kyle, speak in terms of a single market maker. See Chapter 5, Definition 5.1, for the given market efficiency condition.

[176] Notice that if there were only noise traders, then net demand would be $y = u$. The pricing function would reduce to $P(y) = p = E[v|u]$ and the market would be infinitely deep.

An important feature of both Kyle and the benchmark model is that the informed trader trades strategically, meaning that he takes into account the effect of his orders on price. This involves conditioning on the behaviour of the other players, such as noise traders, whose trades are exogenous, and the market maker.

Because the informed trader (System 2 algorithmic trader in the benchmark model) is risk neutral, he will choose a strategy that maximises his expected profit (this is Kyle's inference).[177] That is, he chooses a demand (or supply) x that satisfies:

$$\max_{x} E[\pi] = E[(v - p)x|v]$$

for each possible realisation of v. The interaction between the market makers problem and the informed trader's problem is clear from the last two equations. The market makers pricing rule depends on the contribution of x to order flow, but the informed System 2 algorithmic traders choice of x depends on the impact orders have on the market makers price p. In equilibrium this circularity is resolved. See below.

The pricing and trading rules that produce equilibrium convey the benchmark model's essential lessons. Like Kyle (1985), the benchmark model focuses on a Bayesian Nash Equilibrium (BNE), where all strategies are linear in equilibrium (namely, we conjecture linear strategies for both the System 2 algorithmic trader and the market maker and then verify that these conjectures are actually the best response to one another's strategies). Section 4.6.2.1 ('Kyles Solution Procedure') highlights the rationale for focusing on a linear equilibrium.

[177] Specifically, under the assumption of risk neutrality, the expected utility of end-of-period profits is simply equal to expected end-of-period profits. Consider the following proof: With an exponential utility function, we obtain: $Max\, E[U(\pi_i)] = Max\left[E(\pi_i) - \left(\frac{A}{2}\right)Var(\pi_i)\right]$. Adding risk neutrality ($A = 0$), maximisation of expected utility simplifies to maximisation of expected profits.

We were able to prove that a linear equilibrium exists for our strategic game such that the following pricing and order submission rules hold:

$$p = \lambda\,(x + u),$$

$$x = \beta v,$$

With strictly positive parameter λ and β which take the following form (not derived here):[178]

$\beta = (\,\sigma_u^2/\sigma_v^2)^{1/2}$ *and* $\lambda = 1/2(\,\sigma_u^2/\sigma_v^2)^{1/2}$.

Notice that the pricing rule and the trading rule depend on the same two parameters – the variance of the uninformed order σ_u^2 and the variance in the payoff σ_v^2 – a natural consequence of being determined jointly. Notice also that the ratio of these two parameters is inverted in the two rules. This is quite intuitive: when λ is high, meaning that orders have a high price impact, then β is low, meaning that the informed System 2 algorithmic trader trades less aggressively (to avoid the impact of his own trades). Recall, β fully characterises the System 2 algorithmic traders trading intensity. The constituent variance parameters are also easily interpreted. When σ_v^2 is high, the System 2 algorithmic traders information is more likely to be substantial, ceteris paribus, inducing the market maker to adjust price more aggressively.

To clarify, λ determines the price increase of an additional buy order. The reciprocal of λ can be viewed as, what Kyle (1985) refers to, as market depth. If λ is low, an additional order will not lead to a large price change and, thus, the market is very liquid. The small price impact of an additional order reflected by a low λ induces the informed trader to trade more aggressively.

[178] For formal derivation refer to equation 5.5.1.

New Trader – Introducing the System 1 Algorithmic Trader

The benchmark model – presented in the previous chapter – introduces System 2 algorithmic traders into the static model of Kyle (1985).[179] Chapter 5 clarifies the methodological procedure used to incorporate System 2 algorithmic traders into the static Kyle (1985) model. In the present chapter (Chapter 6), we extend the benchmark model and introduce a new type of trader – a System 1 algorithmic trader. Like System 2 algorithmic traders in the benchmark case, theoretical considerations for these agents emanate from the key assumptions underlying our dual process taxonomy of algorithmic trading.[180]

In Section 5.2 of Chapter 5, we detail the final integrative dual process *taxonomy* of algorithmic trading. A systematic review of Section 5.2 may prove useful for readers who require further granularity. However, for sake of brevity, we highlight here, only the key insights reflected in the taxonomy - as they apply to System 1 algorithmic traders.

Consider the following: in the context of theoretical market microstructure finance, our dual process categorisation of algorithmic trading suggests that System 1 algorithmic traders possesses and inherent speed advantage relative to the rest of the market.

However, a pertinent question remains: *How do System 1 algorithmic traders acquire this advantage?*

---

[179] It is important to note that the benchmark model nests the Kyle (1985) model. In a limiting case of the benchmark model, where the System 2 algorithmic trader does not have access to his particular form of private information, the equilibrium reduces to the equilibrium of the static model of Kyle (1985).

[180] A key assumption made in our taxonomy is that System 1 algorithmic traders' private information relates to incoming order flow. This means that they have no independent source of information; their information is dependent on the activity (orders) of other traders –including, but not limited to the System 2 algorithmic trader. Thus, it seems inappropriate to model System 1 algorithmic traders without first modelling System 2 algorithmic trading.

In order for our model to be as rich as a description of real markets as possible, we provided an instantiation of System 1 algorithmic trading in Section 3.5. An instantiation refers to the creation of a real instance or particular realization of an abstraction or class of objects and processes (Gregor, 2006).

Following, the instantiation of System 1 algorithmic trading in Section 3.5, we suggest that co-location infrastructure subsists as the mechanism by which System 1 algorithmic traders attain their speed advantage.

As the term suggests, co-*location* involves *locating* ones 'server'[181] as close as possible to the exchanges' matching engine. Again, the matching engine establishes a channel of communication between the market and the trading algorithm. It determines how orders are processed, and thus controls the automation of trades. Placing one's server adjacent to the exchanges matching engine significantly reduces the time it takes to access the central order book – where electronic information on quotes to buy and sell as well as current market prices are warehoused. It also decreases the time it takes to transmit trade instructions and execute matched trades. Essentially, placing ones server adjacent to the exchanges matching engine means that **incoming order flow** can reach System 1 algorithmic traders platform almost instantaneously. Thus, in this thesis, we assume that System 1 algorithmic trader's private information relates to incoming order flow and not about the fundamental value of the asset. This is contrary to System 2 traders who possess private information about the fundamental value of the asset.

[181]. Strictly speaking, a server is a computational device which is dedicated to the running of a certain program. With regards to algorithmic trading, a server functions as an electronic communication device or electronic infrastructure that manages access to a centralised resource.

Strictly speaking, our assertion - that System 1 algorithmic traders utilise colocation facilities - provides ample grounds on which to base the assumption that their private information relates to *incoming order flow* and not about the fundamental value of the asset.

However, additional rationale for this assumption – the assumption that advanced order flow information comprises System 1 algorithmic trader's private information - is provided below in Section 6.2.2.2

6.2.2 System 1 Algorithmic Trading

A Focused Overview

Here we highlight key intuitions regarding System 1 algorithmic trading. We follow immediately (Section 6.2.2.1 & 6.2.2.2) with a detailed description of each concept.

Speed, Information and Holding Horizon

1. Speed: System 1 algorithmic traders have a speed advantage over all other traders (including System 2 algorithmic traders, noise traders, and market makers).

2. Information: advanced information relates to incoming order flow.

3. Short holding horizon: liquidates position by the end of the trading round.[182]

[182]As clarified above, co-location subsists as the mechanism by which the System 1 algorithmic trader attains his/her speed advantage. The SEC (2014) suggests that those algorithmic traders who subscribe to co-location facilities "end the trading day in as close to a flat a position as possible, that is, not carrying significant positions over-night"(p.4). In congruence with this suggestion, we will assume that the System 1 algorithmic trader liquidates his/her inventory at the end of the trading period.

6.2.2.1 Discussion on Speed – (1)

In many models of strategic trading, such as Foster and Viswanathan (1996), Holden and Subrahmanyam (1992), and Vayanos (1999), all traders move at the same speed: they act one per round. An interesting feature of System 1 algorithmic traders, however, is not *just* that they are fast but that they are faster than others. In the final model, the System 1 algorithmic trader has speed advantage over all other traders in various ways.

First, the System 1 algorithmic trader is able to trade *twice* without latencies in the trading round, the System 2 algorithmic trader and noise traders' can only trade *once.*

Second, the System 1 algorithmic trader is also faster than the market maker. This is a particularly interesting feature of the final model and requires further clarification.

(I) Capturing the System 1 algorithmic trader's speed advantage over the market maker

The market maker in the final model sets a price and trades the quantity that makes markets clear. Specifically, at the beginning of a trading round, the competitive market maker posts a linear pricing function $P(\cdot)$ against which others can trade. Like both Kyle (1985) and the benchmark model, the slope of the pricing function – denoted by λ – represents how much the market maker adjusts the price in response to the net order flow.

Interestingly, we are able to capture the System 1 algorithmic trader's speed advantage relative to the market maker using the slope of the market makers pricing function. In other words, during the single trading round of the final model, the slope of the market makers pricing function is *fixed*. The rationale for fixing the pricing function at the beginning of the trading round is as follows:

The System 1 algorithmic trader can trade *twice* without latencies in the single trading round – we have already alluded to, and justified this assumption.

In an ideal world, market makers would adjust the slope of the pricing function (λ) in light of this new information/ trade innovation (this adjustment mechanism for market makers is discussed at length above and does not merit further discussion here. See Chapters 4 and 5 for more detail)

By fixing λ in the final model, the market maker cannot adjust the pricing function instantaneously. The pricing function $P(\cdot)$ is set at time 0 and cannot be updated until the end of the trading.

Effectively, lags in adjusting the pricing function $P(\cdot)$ prevents the market maker from using all information to price the orders in the trading window and thus represents the market maker's speed disadvantage relative to the System 1 algorithmic trader.

6.2.2.2. Discussion on Information (2) and Holding Horizon (3)

From a theoretical perspective, our dual process categorisation of algorithmic trading suggests that System 1 algorithmic traders possess and inherent speed advantage relative to the rest of the market.

In practice, faster speeds may also imply that they can trade more frequently, have smaller inventories and shorter holding horizons (Froot, Scharfstein & Stein, 1992). The assumption of shorter holding horizons, although restrictive, is also consistent with the stylised facts of System 1 algorithmic traders (Section 5.2).

The short holding horizon has important implications for the System 1 algorithmic trader's behaviour and choice of information. Unlike the other traders, the strategic fast System 1 algorithmic trader incurs price impact twice in one trading round. When establishing a position, the System 1 trader must have a plan to exit within a short time window. His profit is not determined by the difference between the entry price and the fundamental value, but by the difference between his entry and exit prices.

Hence, System 1 algorithmic traders do not try to infer the long-term fundamental value of the risky asset but focus on predicting short-term price dynamics.

We suggested that co-location facilities subsist as the mechanism behind this speed advantage.

Isolating co-location infrastructure as the physical source/mechanism behind the System 1 algorithmic trader's speed advantage enabled us to make additional assumptions regarding System 1 algorithmic traders. Although these assumptions are thoroughly addressed in the proceeding subsections, we provide a brief overview below:

In practical terms, speed or *latency* refers to the time it takes to access and respond to market information. Locating ones server adjacent to the exchanges matching engine significantly reduces the time it takes to access and respond to the central order book – where electronic information on quotes to buy and sell are warehoused. Essentially, the reduced latency provided by colocation infrastructure means that those who utilise this technology can view, and react to incoming order flow in the context of microseconds.

Because System 1 algorithmic trader's speed advantage is predicated on co-location facilities, it seems reasonable to suggest that their private information concerns incoming order flow.

In congruence with the suggestion made above, we will assume that the fast System 1 trader obtains a signal I_y about the aggregate incoming order flow $y = x + u$, not about the fundamental value v.[183]

183 System 1 algorithmic trader's focus on predicting short-term price dynamics (i.e., incoming order flow). Conversely, System 2 algorithmic trader's try to infer the long-term fundamental value of the risky asset.

6.3 SYSTEM 1 ALGORITHMIC TRADER: KEY ASSUMPTIONS

The key assumptions made with regards to the System 1 algorithmic trader emanate from the discussion above, they are as follows:

- In the single trading round, the System 1 algorithmic trader obtains a signal I_y about the aggregate incoming order flow $y = x + u$.
- They rapidly trade *twice* in *one* trading round and do not carry inventory when the trading round ends.

6.4 THE FINAL MODEL: INSIGHTS

The benchmark model - presented in the previous chapter – introduces System 2 algorithmic traders into the static model of Kyle (1985). System 2 algorithmic traders are introduced into an extended Kyle framework with an adjusted information structure – we modify the sequence and acquisition process of the informed trader's private information. Interested readers are referred to Section 5.4, where we clarify the formal methodological procedure used to incorporate this type of trader into the static Kyle (1985) model.

Although the benchmark model provides a rich description of System 2 algorithmic trading it is still based on the basic framework of Kyle (1985). Thus, the benchmark model serves an important function - it allows us to keep our analysis (i.e., the final model) tractable. We emphasise that in a limiting case of the benchmark model, where the System 2 algorithmic trader does not have access to the particular form of private information described above; the equilibrium reduces to the equilibrium of the static model of Kyle (1985). Inexorably, the baseline/benchmark model *nests* the Kyle (1985) model. The formal model, to follow, builds on the benchmark model and introduces the System 1 algorithmic trader into the benchmark model.

Interestingly, in the final model, when the System 1 algorithmic trader has no information, the equilibrium reduces to the equilibrium of the benchmark model.

Insights

The final model presented below generates many novel insights. Relative to both the benchmark model and Kyle (1985), three of the most important insights include:

1. Timing: The System 1 algorithmic trader's speed advantage emerges within a single trading round. More specifically, we introduce latencies (delays) between trades and quotes into the single trading period of the benchmark model. The System 1 algorithmic trader is the only trader fast enough to exploit these latencies.

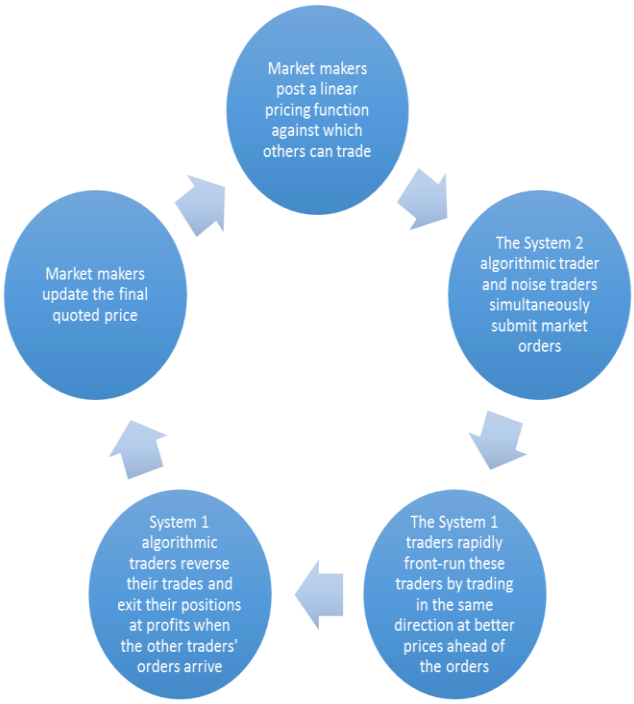
2. Private information: in keeping with theoretical market microstructure finance, we can consider both System 1 and System 2 algorithmic traders informed traders – the difference between these informed traders lies in the nature of their private information. Contextually, System 2 algorithmic traders have an accuracy advantage over all other market participants. Again, this accuracy advantage emanates from, and concerns, a firm's fundamentals (i.e., they are informed about the long-term fundamental value v of the risky share). On the other hand, System 1 algorithmic traders have a speed advantage over the rest of the market. Crucially, their private information concerns incoming order flow y , not the fundamental value v. To summarise, in the final model, we will assume that System 1 algorithmic traders are informed about short-term price dynamics (i.e., incoming order flow), and System 2 algorithmic traders; long-term fundamentals (i.e., the *ex-post* liquidating value of the risky share v)

3. Price discovery is deeply related to the nature of information heterogeneity between informed traders (algorithmic traders in our case): Our findings (with respect to asymmetric/private information, algorithmic trading and price discovery) are particularly profound and convey some of the model's essential lessons.

Though the model itself will enrich each of these insights, let us offer a few thoughts at this stage. Insight (1) emphasises that delays between trading and quoting emerge within a single trading round. The assumption of a single trading round is strong in the sense that it allows us to convey some fundamental aspects of System 1's relative information advantage – a speed advantage. Consider the intuition regarding our supposition of a single trading round:

There are multiple rounds of trading ($t = 1, 2, 3$) in multi-period models. However, we argue that multiple rounds of trading would ultimately be a less rich description of an agents speed advantage. Indeed, the term *speed* evokes images of short time-scales. *Ergo, a single round of trading (like the one presented here) would be more suggestive of an agent's ability to trade faster than others.* In other words, we suggest that things can be properly illustrated with a single trading round. Consequently, in order to reflect the System 1 algorithmic traders advantage within a short time window, we conform to the time decomposition $[0.1]$ – typical of *single* period models. However, we add some granularity to the timing of events within a single trading round framework. We introduce delays between quotes and trades within the single period and add a trader fast enough to exploit these short delays (again, the System 1 algorithmic trader). More specifically, the single trading round, which begins at $\mathrm{t} = 0$ and ends at $\mathrm{t} = 1$ is further decomposed into more fine sub-intervals/instances denoted $t = 0,\ t = 0^{+},\ t = 1^{-},\ t = 1,\ t = 1^{+}$. See the model itself for further clarity. A broad visual summary of a typical *trading round* is provided below in Figure 6.1.

FIGURE 6.1: A TYPICAL TRADING ROUND



Insight (2) introduces a type of private information that is not present in the benchmark model. Notably, in the single trading round of the final model the System 1 algorithmic trader anticipates the size of incoming market orders – he obtains a signal I_y about the aggregate incoming order flow $y = x + u$.

In the final model, the System 1 algorithmic trader will be faster than the System 2 algorithmic trader but the System 2 algorithmic trader will have better information about the fundamentals. To be more precise, System 1 algorithmic traders differ from System 2 algorithmic traders (as seen in the benchmark model) in two respects: (a) they have higher speeds and shorter holding horizons than the informed System 2 algorithmic trader and (b) the System 1 algorithmic trader has no independent source of information, his information concerns incoming order flow.[184]

[184] Recall, incoming order flow depends on the activities of other traders, including the System 2 algorithmic trader.

Insight (3) - that price discovery is deeply related to the nature of information heterogeneity between algorithmic traders is a fascinating message. To understand this insight one needs to understand how the final model relates to the literature on informed trading. Existing attempts at characterising speculative informed trade have taken one of two basic approaches. The approach adopted by Kyle (1985), Back (1992), Brunnermeier (2002) and Baruch (2002) is to assume that (i) there is a single informed trader; and (ii) the single informed trader has either a speed advantage *or* an accuracy advantage over other agents (see e.g. Section 4.3 where we delineate the dynamics of each advantage). The alternative approach considers multiple informed traders. However, models with multiple informed traders typically limit assumptions to cases of 'symmetrically informed' informed traders – i.e., informed traders have identical private information.

Interestingly, the *two* informed traders in our model have heterogeneous information advantages.

The nature of information asymmetry between the two informed algorithmic traders in our model contrasts the findings of other models with multiple informed traders' vis-à-vis the price discovery process. Consider the following example:

Holden and Subrahmanyam (1992) posit that in financial markets, at least a few players will have access to private information and will trade in the knowledge that they will face competition with other informed agents in the market. Motivated by the above, Holden and Subrahmanyam (1992) develop a market microstructure model in which multiple, identically informed 'informed' traders strategically exploit their common private information. They show that such traders compete aggressively and cause most of their common private information to be revealed very *rapidly*. Evidently, the aggressive competition between informed traders translates into more private information being incorporated into price. Back, Cao and Willard (2000), Bernhardt and Miao (2004), Foster and Viswanathan (1996), and Li (2013) all confirm the assumption that informed

trader's trade very intensely on their common information. In these models, an informed trader will increase his/her trading intensity when faced with the competition from other informed traders with *highly positively correlated* information. Again, an elevated informed trading intensity leads to more private information being revealed in prices i.e., prices become more informative.

It seems that the informed trader's trading intensity is the single most important driving force behind the results of each model respectively. Indeed, the informed trader's trading intensity has a major effect on the price discovery process because it reflects how much the informed trader's private information is revealed through the trading process.

Essentially, when an informed trader chooses a trading intensity, he faces the trade-off of price impact and information decay. If an informed trader reduces trading intensity, his private information is traded away by competing informed traders; if an informed trader increases trading intensity, his price impact is higher.

Critically, this trade-off between price impact and information decay is not reflected in our final model. Readers should keep in mind that this insight can only be appreciated following the formal presentation of the model itself. Only a summary of this insight is provided below:

In contrast to existing models with multiple informed traders, where an informed trader will increase his trading intensity when faced with competition from other informed traders, the System 2 algorithmic trader (one of the informed traders) in our final model reduces rather than increases his trading intensity when faced with competition from the System 1 algorithmic trader (the other informed trader). In other words, the System 2 algorithmic trader has no incentive to increase trading intensity when faced with the competition from System 1 algorithmic trader in our model.

The rationale is as follows:

Firstly, the System 2 algorithmic trader cannot reduce information decay by trading more intensely because the System 1 algorithmic trader always anticipates the order flow – the System 1 algorithmic trader always observes a private signal I_y about the incoming order flow $y = x + u$. Secondly, if the System 2 algorithmic trader does not trade at all, his private information is not traded away by the System 1 algorithmic trader because the System 1 algorithmic trader does not have an independent source about fundamentals.

Why does the System 2 algorithmic trader *reduce* trading intensity?

The answer lies in the nature of the System 1 algorithmic trader's advantage over the market maker. Like Kyle (1985), the benchmark model, and other extant literature on market microstructure finance, the assumption that market makers earn zero profits is key to our analysis here. In the context of the benchmark model, the market maker makes losses on the trades with the informed System 2 algorithmic trader, but recoups these loses on trades with the noise traders - making zero profit on average.

However, in the final model the market maker cannot update the pricing function fast enough to keep up with the System 1 algorithmic trader (i.e., the market makers pricing function is fixed in the single trading round).

Because market makers cannot adjust the pricing function instantaneously, they suffer additional losses when trading with the System 1 algorithmic trader in the final model. To make up for the additional losses, market makers have to increase the slope of the pricing function ex-ante in order to *break-even* (i.e., they charge more for absorbing order flow imbalance.)

Effectively, the System 1 algorithmic trader extracts rent from market makers and market makers shift the burden to System 2 and noise traders by making liquidity more costly. In response to the higher cost of liquidity, the System 2 algorithmic trader reduces trading intensity.

To summarise, unlike in existing models with multiple informed traders such as Back et al., (2000), Bernhardt and Miao (2004), Holden and Subrahmanyam (1992), Li (2013), and Vayanos (1999); in our model, the better informed System 2 algorithmic trader (by better informed we mean that he/she has better information about fundamentals) always reduces his/her trading intensity in the presence of the fast System 1 algorithmic trader, and prices become less informative.

6.5 A DUAL PROCESS MICROSTRUCTURE MODEL OF ALGORITHMIC TRADING

6.5.1 Model Setup

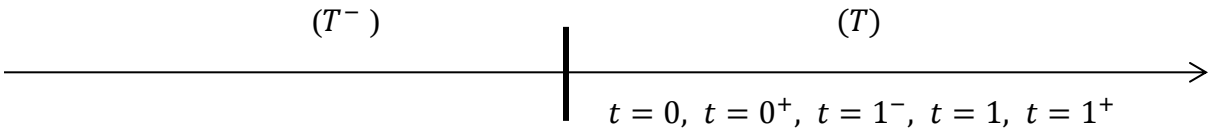
Below we present our final dual process microstructure model of algorithmic trading. This section presents the final model with two algorithmic traders – System 1 and System 2 algorithmic traders. In the particular model we investigate, one risky asset is exchanged for a riskless asset among four kinds of traders: a single System 2 algorithmic trader who has unique access to sequential private observations of the ex-post liquidation value of the risky asset; uninformed noise traders who trade randomly for exogenous non-informational reasons; a continuum of competitive market makers who passively absorb order flow imbalance; and a single System 1 algorithmic trader who anticipates incoming market orders, trades rapidly and has a short holding horizon.

Like the benchmark model, the final model considers two 'periods' – (1) an extended processing period/phase and (2) a single trading period. The System 2 algorithmic trader, market makers and noise traders have the same action timing as in the benchmark model. However, we modify the timeline of the benchmark model to accommodate the System 1 algorithmic trader.

More precisely, we introduce delays in trading and quoting into the *trading period*; the System 1 algorithmic trader is the only trader fast enough to exploit the short delays

Timeline and Information Structure This thesis presents models of *one* trading round. Figures 6.2 and 6.3 illustrate the timeline of the generalized model. Like the benchmark model, the System 2 algorithmic trader is granted an additional phase of information processing in the final model. The extended processing phase, denoted T^- , occurs in the period immediately preceding the trading round. T denotes the single trading round. In this single trading round, trading takes place in the interval $[0.1^+]$.

FIGURE 6.2



In the extended processing period T^- we assume that the System 2 algorithmic trader obtains and observes uniquely private information about noise in a forthcoming, value relevant signal (this value relevant signal is set for a forthcoming period T). The private information concerning a future signal noise and is defined as advanced K information. In its totality this information communicates $K_0 = s - e_s$. The structure of K information was presented above in Section 5.4.3, Box 5.4.

At T we initialize trading. More formally, at time $t = 0$ the trading round starts. Here, market makers set a publicly observable pricing function $P(\cdot,\cdot)$. An order of y_t shares arrives at time $t \in (0,1]$ and is filled by the market makers at the average price of $p_t = P(y_t, \mathcal{F}_t)$.

$\mathcal{F}_t$ denotes market makers' information by time t. To model market maker's latency, the pricing function $P(\cdot,\cdot)$ is fixed in time interval $(0,1^+)$.

At time 0 the System 2 algorithmic trader receives the signal $s = v + e_s$. This signal communicated firm value with noise. At this point the extended processing period private information, in the form $K_0 = s - e_s$, can be utilized to correct for the noise in the signal. Thus, by time 0 , the System 2 algorithmic trader has perfect information about the fundamental value of the asset.

At time 0^+, after observing v and the pricing function $P(\cdot,\cdot)$, the System 2 algorithmic trader submits a market order of x shares. At the same time noise traders submit a market order of u shares. Both orders suffer from a short latency and will not arrive in the market until time 1. Trades and quote updates may take place between time 0^+ and time 1. The delay is so short that no one other than the System 1 algorithmic trader can exploit it.

Right after time 0^+ the fast System 1 algorithmic trader observes a private signal I_y about the incoming order flow $y = x + u$ where $e_y \sim N(0, \sigma_e^2)$ denotes the normally distributed observation error. The quality of the signal I_y is represented by ρ, the squared correlation between I_y and y, i.e.,

(6.1) $$\rho \equiv Corr^2(y, I_y) \in (0,1]$$

We take information quality ρ as exogenously given.[185] A more informative signal I_y has a higher ρ. If I_y reveals y precisely, ρ = 1; if I_y is almost all noise, $\rho \to 0$. Moreover, the projection theorem for normal random variables implies that $\acute{y} = E[y|I_y] = \rho.I_y$.

185 A fixed ρ implies that variance of the observation error. For example, if it is know that the informed trader almost does not trade ($\sigma_x^2 \to 0$), the observation error is almost entirely about noise trading size u.

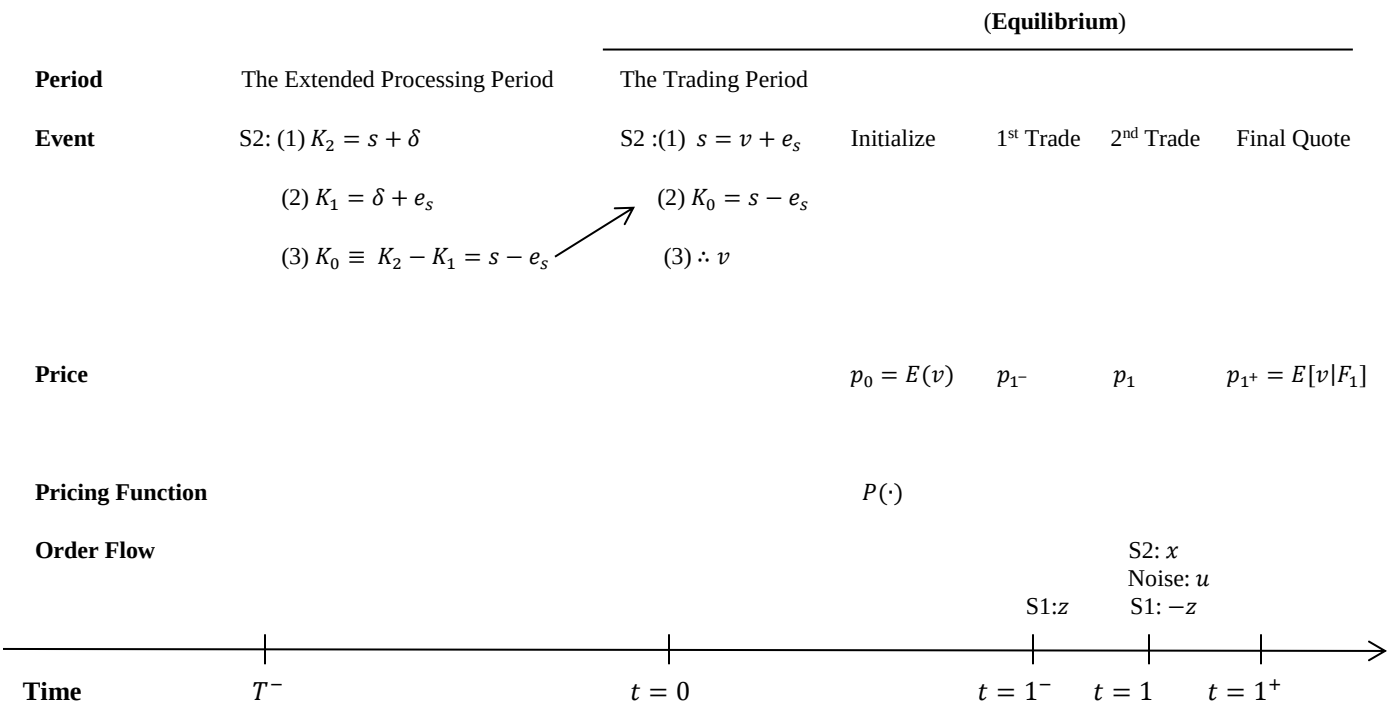
At time 1^{-}, based on his signal I_y the System 1 algorithmic trader trades z shares. Market makers fill the order at the price of $p_{1^-} = P(z, \mathcal{F}_{1^-})$.

At time 1, System 2 trader's order x and the noise traders' order u arrive in the market. At the same time, the System 1 trader submits an order of $-z$ shares to liquidate his position because he is not allowed to carry inventory when the trading round ends. Trades are all anonymous and market makers fill all orders at the same price of $p_1 = P(x + u - z, \mathcal{F}_1)$.

Finally, at time 1^{+}, right after time 1, market makers look back at the order flow history of the trading round and update the final quoted price to be p_1^+. The trading round ends.

It is common knowledge that value of the risky asset v, noise traders' order size u, and the observation error e_y are mutually independent.

FIGURE 6.3: TIMELINE OF THE FINAL MODEL



Below we provide an overview of the equilibrium concept. We highlight some key assumptions and define the equilibrium conditions. While the overview does not provide detailed proofs, it does attempt to outline the important steps and highlight the key intuition.

Some Intuition for the Equilibrium

Both the benchmark model and the final model - presented here - are based on Kyle (1985). Recall, Kyle (1985) focuses on the Bayesian Nash Equilibrium (BNE), where all strategies are linear in equilibrium. Likewise, only a linear equilibrium is considered in the final model (i.e. we 'conjecture and verify' only linear strategies)

Equilibrium in the final model is defined by four functions $Z(\cdot)$,$X(\cdot)$,$P(\cdot)$ and $Q(\cdot)$. They denote the System 1 algorithmic trader's trade size strategy, the System 2 algorithmic trader's trade size strategy, market makers' pricing function, and the final quote function respectively.

In equilibrium, four conditions have to be satisfied:

1. System 1 algorithmic traders profit maximisation.

2. System 2 algorithmic traders profit maximisation.

3. Competitive pricing function.

4. Informationally efficient quotes.

Like Kyle (1985), the benchmark model, and other extant literature on the topic of market microstructure we conjecture that the pricing function is linear. Given that the linear pricing function, market makers fill the order at the average price of

$$p_{t_j} = P\left(y_{t_j}, \mathcal{F}_{t_j}\right) = p_{t_{j-1}} + \lambda^S y_{t_j}, \qquad j \geq 1$$

λ^S denotes the market depth (price impact) factor which is *fixed* in the trading round. Given the assumptions above, there is a unique equilibrium where

- System 1 trading size: $z = Z(I_y; p_0, \lambda^S) = \alpha\acute{y} = \alpha\rho I_y$
- System 2 trading size: $x = X(v; p_0, \lambda^S) = \beta(v - p_0)$
- Market order pricing: $p_{t_j} = P\left(y_j, p_{t_{j-1}}\right) = p_{t_{j-1}} + \lambda^S y_j$
- Initial quote: p_0
- Final Quote: $p_{1^+} = Q(z, y - z) = p_0 + \lambda^L y$

Under the assumption that P(·) is linear, Z(·),X(·) and Q(·) are all linear. (This follows the assumptions underlying linear equilibrium – discussed at length above. See Chapter 4, specifically the heading *linear equilibrium* for further clarity)

The equilibrium is then fully characterized by the four endogenous parameters α, β, λ^S *and* λ^L

- α is The System 1 algorithmic trader's trading intensity
- β is the System 2 algorithmic trader's trading intensity.
- λ^S is the short-term (temporary) price impact per share of a market order on the transaction price p_t.
- λ^L is the long-term (permanent) impact per share on the final quote p_{1+} of the aggregate order size $x + u$.

The formal equilibrium of the final model is analysed below.

6.6 FORMAL EQUILIBRIUM OF THE FINAL MODEL

Definition 6.1 (Equilibrium conditions). Based on the generalizations mentioned above, we assume that the System 2 algorithmic trader observes v noiselessly in equilibrium. The fast System 1 trader chooses his trade size using a strategy function $Z(\cdot)$ and the System 2 trader chooses his trade size using a function $X(\cdot)$. Market makers commit to a pricing function $P(\cdot)$ and set the final quote using a function $Q(\cdot)$. The equilibrium is defined by four functions $Z(\cdot)$,$X(\cdot)$,$P(\cdot)$ and $Q(\cdot)$ such that the following conditions hold:

1. System 2 algorithmic trader profit maximisation. Given $P(\cdot)$,$Z(\cdot)$,and the asset's true value v, the System 2 algorithmic trader's profit $\pi^{S2} = x(v - p_1)$ is maximised if he trades $x\ shares, i.e.,$

$$x = X(v; Z(\cdot), P(\cdot)) = \underset{x}{argmax}\ E[\pi^{S2} | v, P(\cdot), Z(\cdot)] \tag{6.2}$$

where p_1 is the execution price of his trade.

2. System 1 algorithmic trader profit maximisation. Given $P(\cdot), X(\cdot)$, and a signal about the incoming order flow $I_y = x + u + e_y$ the System 1 trader's profit $\pi^{S1} = z(p_1 - p_{1^-})$ is maximised if he trades z shares at time 1^- and liquidates at time 1, i.e.,

$$z = Z(I_y; P(\cdot)\, X(\cdot)) = \underset{z}{argmax}\ E\left[\pi^{S1} \middle| I_y, P(\cdot), X(\cdot)\right] \tag{6.3}$$

where $p_1 - p_{1^-}$ is the difference between his entry and exit prices.

3. Competitive pricing function. Given $X(\cdot)$ and $Z(\cdot)$, market makers choose a pricing function $P(\cdot)$ such that their expected profit $E[\pi^M]$ at time 0 equals zero, i.e.,

(6.4) $$0 = E[\pi^M | P(\cdot), X(\cdot), Z(\cdot)]$$

4. *Informationally efficient quotes*. Market makers set quotes p_0 and p_{1+} to be their expected value of v conditional on available information $\mathcal{F}_0$ and $\mathcal{F}_{1^+}$.

(6.5) $$p_0 = E[v]$$

(6.6) $$p_{1+} = E[v|\mathcal{F}_{1^+}, X(\cdot), Z(\cdot)] = Q(\mathcal{F}_{1^+}; X(\cdot), Z(\cdot))$$

Remark 6.1. Market makers' profits in equation (6.4) is $\pi^M = zp_{1-} + (x + u - z)p_1 - (x + u)v$ because they trade $-z$ shares at price p_{1-} and $-(x + u - z)$ shares at price p_1.

Remark 6.2. Setting p_{1+} has no impact on equilibrium because the game ends at time 1^+

Assumption 6.1 (Linear pricing function). Upon receiving the j-th market order of y_{t_j} shares at $t_j \in (0,1]$, market makers fill the order at the average price of

(6.7) $$p_{t_j} = P\left(y_{t_j}, \mathcal{F}_{t_j}\right) = p_{t_{j-1}} + \lambda^S y_{t_j}, \qquad j \geq 1$$

If y_{t_j} is the first arriving order ($j = 1$), the reference price p_{t0} is the initial quote p_0; if $j > 1$, $p_{t_{j-1}}$ is the average price of the previous traded market order. The price impact (market depth) factor λ^S is fixed in the trading round.

Assumption 6.1 reduces the choice of the pricing function $P(\cdot)$ to the choice of two parameters: the initial quote p_0 and the price impact factor λ^S. As discussed above, a fixed λ^S captures the market maker's speed disadvantage relative to the System 1 algorithmic trader.

Lemma 6.1. *In the final model,* $t_0 = 0$, $t_1 = 1^-$ *and* $t_2 = 1$. *Given Assumption 6.1, the traded prices are*

(6.8) $$p_{1^-} = p_0 + \lambda^S z$$

(6.9) $$p_1 = p_{1^-} + \lambda^S(x + u - z) = p_0 + \lambda^S(x + u)$$

It might seem that the execution price p_1 for the System 2 algorithmic trader and noise traders are not affected by the System 1 algorithmic trader's trading *z* because the System 1 trader completely liquidates his position at time 1. The observation, however, is not correct because in equilibrium the price impact factor λ^S is endogenously determined by the System 1 trader's trading intensity. We now examine the equilibrium in more detail.

Theorem 6.1 (Equilibrium of the final model). *Given the linear pricing function, there is a unique equilibrium where*

(6.10) System 1 trading size: $z = Z(I_y; p_0, \lambda^S) = \alpha\acute{y} = \alpha\rho I_y$

(6.11) System 2 trading size: $x = X(v; p_0, \lambda^S) = \beta(v - p_0)$

(6.12) Market order pricing: $p_{t_j} = P\left(y_j, p_{t_{j-1}}\right) = p_{t_{j-1}} + \lambda^S y_j$

(6.13) Initial quote: p_0

(6.14) Final quote: $p_{1^+} = Q(z, y - z) = p_0 + \lambda^L y$

The endogenous parameters $\alpha, \beta, \lambda^S$and $\lambda^L$are:

(6.15) $$\alpha = \frac{1}{2}, \quad \beta = \frac{\sigma_u}{\sigma_v}\theta, \quad \lambda^S = \frac{\sigma_v}{\sigma_u}\frac{1}{2\theta}, \quad \lambda^L = \frac{\sigma_v}{\sigma_u}\frac{\theta}{1+\theta^2}$$

Where the market quality parameter

(6.16) $$\theta \equiv \sqrt{\frac{1-\rho/4}{1+\rho/4}} \ \epsilon[\sqrt{0.6}, 1]$$

6.7 EQUILIBRIUM ANALYSIS

The strategy functions $X(\cdot)$, $Z(\cdot)$, and $Q(\cdot)$ are all linear if the pricing function $P(\cdot)$ is linear.[186] The equilibrium is then fully characterized by the four endogenous parameters α, β, λ^S and λ^L illustrated in Figure 6.4:

The parameter α is the System 1 algorithmic trader's trading intensity. The System 1 trader first observes the signal I_y of the incoming order flow y and estimates the order flow to be $\acute{y} = \rho I_y$. Then, the System 1 trader chooses to trade $z = \alpha\acute{y}$ shares at time 1^- and $-\alpha\acute{y}$ at time 1. A higher intensity α indicates that the System 1 trader trades more given an estimate of order flow $\acute{y}$.

As in the benchmark model, trading intensity β characterizes the System 2 algorithmic trader's strategy in equilibrium. The conditions set in the previous analysis imply that System 2 traders have perfect information in equilibrium. In other words, the System 2 algorithmic trader is able to combine K_0 and s (refer to Section 5.4.3 for further clarity) to form uniquely private information about the fundamental value of the risky asset once trading is initialized. Given his perfect observation of v, we can characterize System 2 activity in equilibrium as follows: The System 2 algorithmic trader first calculates the pricing error $v - p_0$ using his perfect private information on v and the initial quote p_0. He then submits a market order of $x = \beta(v - p_0)$ shares. A higher β indicates that the System 2 trader trades more aggressively based on the same pricing error $v - p_0$.

[186] In line with Kyle (1985) and BNE in general, we have assumed that the pricing function is linear. Note that in Kyle's (1985) proof, a linear pricing rule implies linear strategies and vice versa. Boulatov, Kyle and Livdan (2005) prove that the linear equilibrium in the single-period trading model of Kyle (1985) is unique; suggesting that equilibria with a nonlinear structure cannot exist in Kyle's setting. Invariably, all models following Kyle conjecture and verify only linear trading and pricing rules respectively. Our model is consistent with convention here.

The parameter λ^S represents the short-term (temporary*)* price impact per share of a market order on the transaction price. Transaction price responds to the order flow according to the pricing function $\Delta p_t = \lambda^S y_t$. Effectively, market makers charge $\lambda^S y^2$ to execute a market order of *y* shares. Competitive market makers set λ^S just enough such that their revenues for executing trades exactly offset their loss in trading with the System 1 and System 2 algorithmic traders. A higher λ^S means that it costs more to execute a market order of any given size.

The slope λ^L represents the long-term (permanent*)* price impact per share on the final quote p_1^+ of the aggregated order size $x + u$. The difference between the closing quote p_1^+ and the opening quote p_0 equals $\lambda^L(x + u)$. Because market makers are competitive, the quote update $p_1^+ - p_0$ is determined by the information content of the order flow $x + u$. A higher λ^L indicates that the aggregate order flow $x + u$ contains more information about the fundamental value *v* and thus the quote update $p_1^+ - p_0$ is more sensitive to $x + u$.

FIGURE 6.4: EQUILIBRIUM STRATEGIES

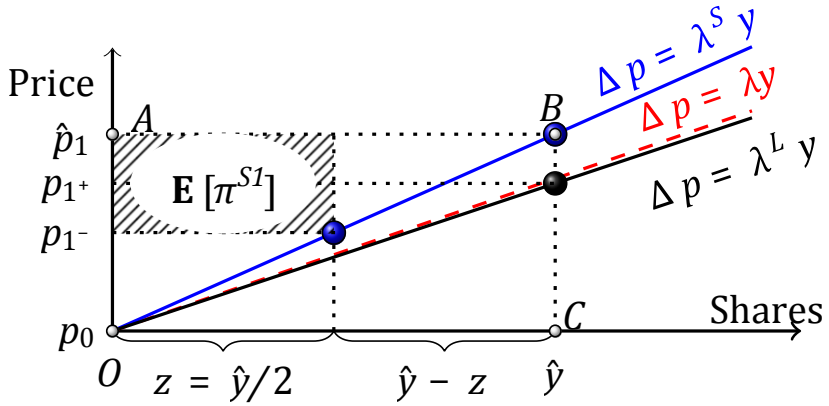


Figure 6.4. Equilibrium strategies. $\Delta p_t = \lambda y$ is the pricing function with only a System 2 algorithmic trader. When the System 1 algorithmic trader is present as well, market makers raise λ^S and lower λ^L. At time 0, market makers set p_0 and λ^S; at time 1^-, the System 1 algorithmic trader trades $z = \acute{y}/2$ shares at price $p_{1^-} = p_0 + \lambda^S z$; at time 1, the System 2 algorithmic trader trades x shares, the noise traders u shares, and the System 1 $(-z)$ shares at the price $p_1 = p_{1^-} + \lambda^S(x + u - z) = p_0 + \lambda^S y$; finally at time 1^+, market makers set the quote to $p_{1^+} = p_0 + \lambda^L y$. The shaded rectangle is the System 1 traders profit π^{S1}. Its area is 1/4 of rectangle *OABC* which corresponds to market makers' price impact surplus.

The other three endogenous parameters β, λ^S, and λ^L are determined by three exogenous parameters: volatility of the fundamental value σ_v, volatility of noise trading σ_u, and information quality of the System 1 algorithmic trader $\rho \in [0,1]$. The equilibrium effects of σ_v and σ_u, are similar to both the benchmark model and to Kyle (1985). In addition, we can set $\sigma_v = \sigma_u = 1$ by choosing certain units of currency and trade size. System 1 trader's information quality ρ, however, is invariant to change of units. When the System 1 algorithmic trader has no information ($\rho = 0$),

the equilibrium reduces to the benchmark model.[187] Likewise, barring System 2 algorithmic traders in the benchmark model, the benchmark model collapses to Kyle's (1985).

Overall, the models simplicity and elegance is a significant contribution to the novelty of the thesis.

Corollary 6.1. *Other things equal, when the System 1 traders information becomes more accurate ($\rho \uparrow$), short-term price impact increases ($\lambda^S \uparrow$), long-term price impact decreases ($\lambda^L \downarrow$), System 2 trading intensity declines ($\beta \downarrow$), and the System 1 algorithmic traders intensity α is unchanged.*

In the presence of System 1 algorithmic traders, market makers raise the short-term or temporary price impact ($\lambda^S \uparrow$) to break-even. System 2 algorithmic traders respond by reducing trading intensity ($\beta \downarrow$). Because System 2 algorithmic trader's order x is the only informative component of the order flow $x + u$, the aggregate order flow $x + u$ becomes less informative ($\lambda^L \downarrow$) in the presence of the System 1 algorithmic trader.

6.8 DUAL-PROCESS MODEL INSIGHTS

In the final model, the System 1 algorithmic trader is faster than the System 2 algorithmic trader but the System 2 algorithmic trader has better information about the fundamentals. To be more precise, System 1 algorithmic traders differ from System 2 algorithmic traders (as seen in the benchmark model) in two respects: (a) they have higher speeds and shorter holding horizons than the better informed System 2 algorithmic trader and (b) the System 1 algorithmic trader has no independent source of information, his information concerns incoming order flow.[188]

[187] When the System 1 trader has some information $\rho > 0$, the equilibrium differs.

[188] Recall, incoming order flow depends on the activities of other traders.

The nature of information asymmetry between the two informed algorithmic traders in our model contrasts the findings of other models with multiple informed traders' vis-à-vis the price discovery process. Consider the following example:

Holden and Subrahmanyam (1992) posit that in financial markets, at least a few players will have access to private information and will trade in the knowledge that they will face competition with other informed agents in the market. Motivated by the above, Holden and Subrahmanyam (1992) develop a market microstructure model in which multiple, identically informed 'informed' traders strategically exploit their common private information. They show that such traders compete aggressively and cause most of their common private information to be revealed very rapidly. Evidently, the aggressive competition between informed traders translates into more private information being incorporated into price. Back et al., (2000), Bernhardt and Miao (2004), Foster and Viswanathan (1996), and Li (2013) all confirm this assumption that informed traders trade very intensely on their common information. In these models an informed trader will increase his trading intensity when faced with the competition from other informed traders with highly positively correlated information. Again, an elevated informed trading intensity leads to more private information being revealed in prices i.e., prices become more informative.

Invariably, in the models described above, when choosing a trading intensity, the informed traders face a trade-off between price impact and information decay. If an informed trader reduces trading intensity, his private information is traded away by competing informed traders; if an informed trader increases trading intensity, his price impact is higher.

Critically, this trade-off between price impact and information decay is not reflected in our final model.

Indeed, in contrast to existing models with multiple informed traders where an informed trader will increase his trading intensity when faced with competition from other informed traders, the System 2 algorithmic trader (one of the informed traders) in or final model *reduces* rather than *increases* his trading intensity when faced with competition from the System 1 algorithmic trader (the other informed trader). In other words, the System 2 algorithmic trader has no incentive to increase trading intensity when faced with the competition from System 1 algorithmic trader in our model. The rationale is as follows:

Firstly, the System 2 algorithmic trader cannot reduce information decay by trading more intensely because the System 1 algorithmic trader always anticipates the order flow. Secondly, if the System 2 algorithmic trader does not trade at all, his private information is not traded away by the System 1 algorithmic trader because the System 1 algorithmic trader does not have an independent source about fundamentals.

Why does the System 2 algorithmic trader reduce trading intensity?

The answer lies in the nature of the System 1 algorithmic trader's advantage over the market maker. Like Kyle (1985), the benchmark model, and other extant literature on market microstructure finance, the assumption that market makers earn zero profits is key to our analysis here. In the context of the benchmark model, the market maker makes losses on the trades with the informed System 2 algorithmic trader, but recoups these losses on trades with the noise traders – making zero profit on average.

However, in the final model the market maker cannot update the pricing function fast enough to keep up with the System 1 algorithmic trader (i.e., the market makers pricing function is fixed in the single trading round).

Because market makers cannot adjust the pricing function instantaneously, they suffer additional losses when trading with the System 1 algorithmic trader in the final model. To make up for the additional losses, market makers have to increase the slope of the pricing function ex-ante in order to *break-even* (i.e., they charge more for absorbing order flow imbalance.)

Effectively, the System 1 algorithmic trader extracts rent from the market maker and the market maker shifts the burden to System 2 and noise traders by making liquidity more costly. In response to the higher cost of liquidity (transacting), the System 2 algorithmic trader reduces trading intensity.

Therefore, unlike in existing models with multiple informed traders such as Back et al., (2000), Bernhardt and Miao (2004), Foster and Viswanathan (1996), Holden and Subrahmanyam (1992), Li (2013), and Vayanos (1999), in our model, the better informed System 2 algorithmic trader (by better informed we mean that he has better information about fundamentals) always reduces his trading intensity in the presence of the other informed fast System 1 algorithmic trader and prices become less informative.

6.9 FINDINGS: MARKET QUALITY

Other things equal, when the System 1 traders information becomes more accurate ($\rho \uparrow$), short-term (temporary) price impact increases ($\lambda^S \uparrow$), long-term (permanent) price impact decreases ($\lambda^L \downarrow$), System 2 trading intensity declines ($\beta \downarrow$), and the System 1 algorithmic traders intensity α is unchanged.

Figure 6.5 illustrates the equilibrium impact of the System 1 algorithmic trader's information quality ρ. Due to the existence of System 1 algorithmic traders; market makers cannot break-even if they set the price p_1 to equal their posterior expectation. The System 1 algorithmic trader intercepts z shares of the order flow $x + u$: he acquires z shares from market makers at a discounted price p_{1^-} and supplies the shares back to the System 2 and noise traders at a profit. To make up for the loss to the System 1 trader, market makers have to charge more to absorb the same order imbalance. They raise the temporary price impact factor λ^S above the permanent price impact λ^L implied by the informativeness of the order flow. Market maker's break-even price at time 1 thus differs from their posterior conditional expectation.

When the signal I_y is more accurate signal *($\rho \uparrow$)*, the System 1 trader makes more profit. Again, market makers raise the temporary price impact *($\lambda^S \uparrow$)* more to break-even. The System 2 trader responds by reducing his trading intensity *($\beta \downarrow$).* Because the System 2 trader's order x is the only informative component of the order flow $x + u$, the aggregate order flow $x + u$ becomes less informative *($\lambda^L \downarrow$).*

In the final model, the System 1 algorithmic trader can profit on a less accurate signal I_y if he is faster than others. In addition to the 'information rents', market makers have to pay the 'speed rents'. Because market makers are not the fastest, they protect themselves by setting a steeper pricing schedule. System 2 traders reduce their trading intensity faced with a higher cost of transacting.

FIGURE 6.5: EQUILIBRIUM PARAMETERS

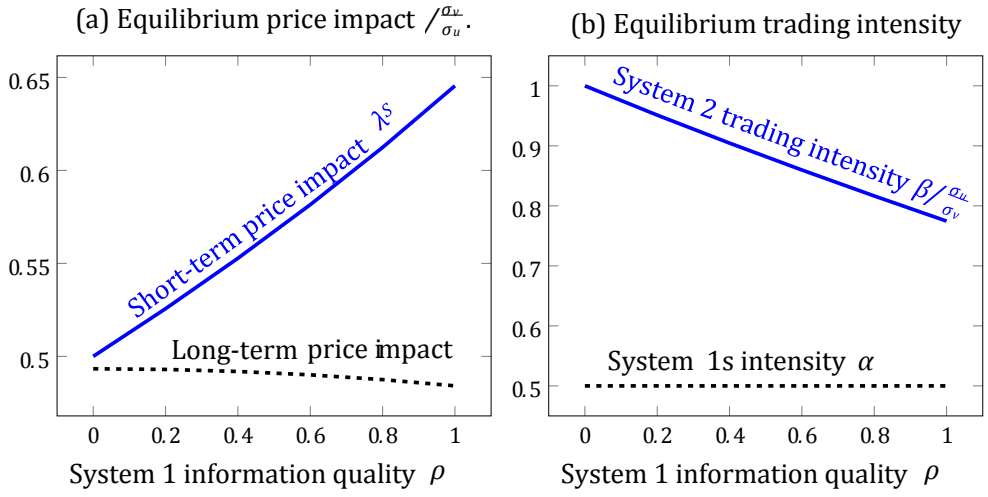


Figure 6.5. Equilibrium parameters of the final model normalized by volatility of fundamental value σ_v and volatility of noise trading σ_u. Theorem 6.1 implies that when the System 1 algorithmic trader's information I_y is more informative (higher ρ) about the incoming order, temporary price impact per share λ^S increases, permanent price impact per share λ^L decreases, System 2 trading intensity β decreases, and System 1 trading intensity α stays the same.

6.9.1 Price Informativeness

Interestingly, this thesis highlights an omission by empirical studies on the topic of algorithmic trading (specifically empirical studies that suggest algorithmic traders have a speed advantage). Hendershott and Riordan (2011) argue that algorithmic traders have a speed advantage over the rest of the market and that this speed advantage results in better price discovery.[189] Other empirical papers conclude that faster traders make prices more informative (e.g., Boehmer, Fong, & Wu, 2015; Brogaard, 2010; Brogaard, Hendershott, & Riordan, 2014; Carrion, 2013; Chaboud, et al., 2009; Hendershott & Moulton, 2011).

[189] In their view, improving the reaction times to new information implies that quotes and trading prices are also incorporating innovations faster.

Overall, the studies above have documented that information, conditional on being incorporated into prices, is incorporated more rapidly when traders are faster. They have interpreted this as evidence that speed makes prices more informative. Those studies, however, do not control for the effect of speed on the amount of information that *ultimately* becomes incorporated into prices. In contrast, this thesis finds that the net effect of an informed traders speed advantage is to make prices less informative in the *long-run*.

Consider the narrative: in the final model, the fundamental value of the risky asset v is the System 2 algorithmic trader's private information. If people learn more about v by observing the trading process, the market is more informationally efficient. Concomitantly, price informativeness can be determined by the fraction of the System 2 algorithmic trader's orders in the aggregate order flow. Because the System 2 algorithmic trader lowers trading intensity in response to the higher costs of liquidity – induced by System 1 algorithmic traders – the aggregate order flow becomes less informative in the presence of (fast) System 1 algorithmic traders.

6.9.2 Market Liquidity

In this thesis, noise traders' trade for non-informational motives; a market is less liquid if noise traders expect to lose more to trade the same number of shares. Vayanos and Wang (2012) point out that different measures of market liquidity are designed to capture different market frictions. Ultimately, however, all the measures attempt to capture the impact of market friction on traders' economic profits. Thus, in addition to the traditional measure λ^S, we also use traders' expected profit to measure market liquidity.

Recall, in the presence of System 1 algorithmic traders, market makers raise the short-term price impact λ^S so that they can still break-even. As a result, it becomes harder for the System 2

algorithmic trader to extract rent based on the same information. In the market microstructure literature, noise traders face adverse selection when they trade with an informed trader with more information. However, in the final model, noise traders face less adverse selection because the System 2 (informed) trader trades less. Nonetheless, noise traders suffer more losses to trade the same number of shares. Effectively, noise traders must pay information rent to the System 2 trader and speed rent to the System 1 trader. The reduction in information rent is not enough to cover the higher speed rent. Hence, the market is less liquid for the System 2 and noise traders in the presence of System 1 algorithmic trading.

Based on the way they trade, it might seem that System 1 algorithmic traders improve liquidity. In the model, System 1 algorithmic traders 'take liquidity' during time $(t_1, 1)$ when liquidity is cheap and 'provide liquidity' at time 1 to the System 2 and noise traders when liquidity is expensive. One might do a 'reduced-form counter-factual analysis'[190] and find that the price would have been much worse for the liquidity demanders if System 1 algorithmic traders were not trading at time 1. In the context of the model, the conclusion is incorrect because if System 1 traders were not present, the temporary price impact λ^S would have been much lower. Effectively, liquidity demanders, including the System 2 and noise traders, would have been better off without System 1 algorithmic traders.

[190] In statistics and econometrics counterfactuals are used in policy evaluations (e.g. Heckman, 2008 & 2010). By a counterfactual we mean: What would have occurred if some observed characteristics or aspects of the processes under consideration were different from those prevailing at the time?

6.9.3 Volatility

Our findings with regards to algorithmic trading and its impact on market volatility are particularly interesting. We find that algorithmic traders (as a whole) increase short-term price volatility but reduce long-term price volatility. These two predictions are both related to the nature of our characterisation of algorithmic trading. More formally, we define market volatility as the divergence of temporary price impact λ^S and permanent price impact λ^L. In our final model, the higher short-term price volatility is caused by a higher temporary price impact λ^S – the result of System 1 algorithmic trading. Long-term volatility (for example volatility calculated with daily closing prices), is mainly determined by the permanent price impact λ^L , which is determined by the trading intensity of the System 2 algorithmic trader. In the final model, the System 2 trader trades slower when faced with a higher price impact λ^S. Therefore in the long run, the System 2 trader produces less information, resulting in lower daily price volatilities.

CHAPTER 7

ALGORITHMIC TRADING, MARKET QUALITY AND INFORMATION: A DUAL PROCESS ACCOUNT – SUMMARY AND CONCLUSIONS

7.1 FINAL SUMMARY AND CONCLUSIONS

This thesis presents the first 'all-inclusive' theoretical model of algorithmic trading (all-inclusive in that it accounts for the multitude of algorithmic trading strategies within a single theoretical framework).

We hope to set some of the groundwork necessary for a new research agenda; one that recognizes that current accounts of algorithmic trading are deficient and that both the empirical and theoretical methods that we traditionally employed may no longer be appropriate. Interestingly, we have attempted to bring the tools of modern cognitive science to bear on the very core of financial market research. More precisely, using the principles of contemporary cognitive science, we have argued that the dual-process paradigm- the most prevalent contemporary interpretation of the nature and function of human decision making – lends itself well to a unique taxonomy of algorithmic trading. Our taxonomy effectively synthesises the multitude of algorithmic strategies within a single framework. One could argue that this taxonomy alone is a novel contribution to the literature. However, in this thesis, the aforementioned taxonomy serves primarily as a heuristic device; enabling us to confront a more complex and important issue: that is, determining the impact of algorithmic trading on financial market quality. More formally, our taxonomy of algorithmic trading serves primarily as a heuristic to inform a theoretical microstructure model; with the view of explaining the evolving nature of market quality as a consequence of algorithmic trading.

Given that algorithmic traders are almost certainly the fastest and the most accurately informed participants in the market, it seems reasonable to suggest that our taxonomy is exclusive to algorithmic trading. Indeed, the literature on algorithmic trading indicates that these agents consistently outperform their human counterparts in terms of trading speed as well as in terms of predictive accuracy. The prominence of this practice – accounting for a much larger proportion of trades relative to traditional human traders – underscores the relevance of our supposition.

The results of our model are by no means trivial and represent an important theoretical contribution to the literature. However, in order to appreciate the significance of our findings one needs to understand the relative impact of each System of algorithmic trading (i.e., the relative impact of System 1 and System 2 algorithmic trading). This is discussed below.

7.2 AN OVERVIEW OF RESULTS

The results of this thesis suggest that one portion of algorithmic trading strategies (those strategies that fall under the rubric of System 2 algorithmic trading) advance financial market quality; whilst another portion of algorithmic trading strategies (those strategies that fall under the rubric of System 1 algorithmic trading) impede financial market quality. Consider the following discussion:

The benchmark model introduces a System 2 algorithmic trader into the static model of Kyle (1985). Evidently, the results of the benchmark model indicate that, relative to the informed trader in Kyle (1985), the System 2 algorithmic trader incorporates more accurate private information into prices i.e., prices become more informative in the presence of System 2 algorithmic trading.

In other words, System 2 algorithmic traders reveal more accurate information in the trading process than the informed trader in Kyle (1985).[191]

The formal dual process microstructure model, presented in Chapter 6, builds on the benchmark model and introduces the System 1 algorithmic trader. Interestingly, relative to the benchmark model, our analysis indicates that System 1 algorithmic traders have a negative impact on market quality. Prices are ultimately less informative when System 1 algorithmic traders are present. In the brief time before the System 2 and noise traders' orders arrive, System 1 algorithmic traders rapidly 'front-run' [192] these trades and bring information to the market and the intermediate prices are more informative. Information value of the intermediate prices, however, is quickly superseded by the more informative orders from the System 2 and noise traders. Ultimately, price informativeness is determined by the fraction of informed trader's orders in the aggregate order flow. Because the informed trader lowers trading intensity in response to higher costs of liquidity, the aggregate order flow becomes less informative in the long-run.

Therefore the final model of this thesis indicates that in the presence of System 1 algorithmic traders, market quality is unambiguously worse: prices are less (ex-post) informative and liquidity is more costly.

[191] The System 2 algorithmic trader has more accurate information on *v* than the informed trader in Kyle (1985). For further clarity, refer to Section 5.7 (An illustrative example of the benchmark model).

[192] Contextually, System 1 algorithmic traders trade earlier during time $(t_1, 1)$ on a noisy signal I_y. This practice has also been referred to as order anticipation in the literature. Technically speaking the term 'front running' applies to the illegal practice addressed in FINRA Rule 5270, which prohibits a broker-dealer from trading for its own account while taking advantage of knowledge of an imminent client block transaction. However, we use the term front running in a less formal sense in this thesis.

7.3 EMPIRICAL IMPLICATIONS

This thesis suggests that algorithmic trades are informative and improve short-term intermediate price informativeness. Consistent with this, Brogaard, Hendershott and Riordan (2013) find that the marketable orders of algorithmic traders have high predictive power about future price changes in less than 5 seconds. Moreover, Hendershott and Riordan (2011) find that algorithmic traders have an informational advantage and quickly incorporate information into prices. Zhang (2012) finds that algorithmic traders profit on 'hard' information and their profits realize quickly. Interestingly, however, our model also suggests that algorithmic traders can reduce information efficiency in the long-run.[193] Differentiating between these two information efficiencies poses new empirical insight:

This thesis indicates that many empirical characteristics of algorithmic trading could be interpreted in different ways. First, in our model, consistent with many empirical studies, algorithmic trading volume is high. Market liquidity, however, would be better if System 1 algorithmic traders have no order flow information ($\rho = 0$) and do not trade at all.

Second our model is consistent with empirical studies that suggest fast traders trades (System 1 algorithmic trades in our case) are informative. Here, System 1 algorithmic trades are informative because they bring some of the System 2 algorithmic trader's information to the market slightly earlier. Nevertheless, the existence of System 1 algorithmic trading reduces the overall information

[193] Contextually, System 1 algorithmic traders make intermediate prices more informative by trading earlier during time $(t_1, 1)$ on a noisy signal I_y. The closing quote p_1+, however, is less informative because the informed System 2 trader reduces trading intensity.

efficiency of prices. Indeed, without System 1 algorithmic trading, the System 2 algorithmic trader would trade more aggressively and improve the overall informativeness of prices.

Contextually, System 1 algorithmic traders make intermediate prices more informative by trading earlier during time $(t_1, 1)$ on a noisy signal I_y. The closing quote p_1+, however, is less informative because the better informed System 2 trader reduces trading intensity.

The above phenomenon can be described concisely using a simple metaphor: suppose a messenger (System 1 algorithmic trader) glances at a letter (order flow) and summarizes it to the receiver (market) right before delivering it. The summary (a System 1 trade) is informative but its information value is fleeting: the letter itself is much more informative than the summary. Furthermore, the sender (the informed System 2 algorithmic trader) is less likely to write clearly ex-ante worrying about privacy issues and ultimately the receiver (market) is less informed.

7.4 POLICY DISCUSSION

Where should resources go in the real world: Towards co-location or complex software for analysis?

In the final model System 1 algorithmic trades are informative because they bring some of the System 2 algorithmic trader's information to the market slightly earlier. Thus, those strategies that rely on co-location facilities i.e., System 1 algorithmic traders, seems to increase *intermediate* information efficiency.

Nevertheless, the social value of intermediate information efficiency is questionable. Intermediate information efficiency could be socially valuable if (1) people can use the intermediate information to make a welfare enhancing economic decision and (2) the cost of delaying the decision from time t_1 to time 1^+ is very high.

Both conditions are unlikely to be true when holding horizons are at the microsecond level. It is hard to imagine a case when other agents must use p_{1^-} to make economic meaningful decisions. Even if the information is crucial, they could wait until time 1^+ and use a more informative price p_1^+. After all, System 1 algorithmic trading is profitable because other traders cannot react fast enough in the time window $[1^-, 1^+]$.

Regardless of one's desired balance of liquidity and price informativeness, when System 1 algorithmic traders have a more informative signal ρ, market quality is unambiguously worse: prices are less (*ex-post*) informative and liquidity is more costly. The only possible social value is a more informative intermediate price which quickly becomes obsolete. Considering its short life, it hardly improves social welfare.

In conclusion, when considering a potential policy about algorithmic trading, policy makers should focus System 1 trader's information precision ρ.[194] A policy that reduces ρ is going to improve ex-post price informativeness and market liquidity.

[194] System 1 algorithmic trader's would not trade at all if they have no order flow information ($\rho = 0$).

References

Abarbanell, J. S., Lanen, W. N., & Verrecchia, R. E. (1995). Analysts' forecasts as proxies for investor beliefs in empirical research. *Journal of Accounting and Economics, 20*(1), 31-60.

Abergel, F., Bouchaud, J.-P., Foucault, T., Lehalle, C.-A., & Rosenbaum, M. (2012). *Market Microstructure: Confronting Many Viewpoints*: Wiley.

Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial studies, 1*(1), 3-40.

Alagidede, P. (2008). *Market Efficiency and Stock Return Behaviour in Africa's Emerging Equity Markets.* Unpublished Doctoral Thesis, Loughborough University.

Alagidede, P., & Panagiotidis, T. (2009). Modelling stock returns in Africa's emerging equity markets. *International Review of Financial Analysis, 18*(1), 1-11.

Albuquerque, R., & Miao, J. (2014). Advance information and asset prices. *Journal of Economic Theory, 149*, 236-275.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*(4), 569.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56.

Angel, J. J., Harris, L. E., & Spatt, C. S. (2011). Equity trading in the 21st century. *The Quarterly Journal of Finance, 1*(01), 1-53.

Australian Securities & Investment Commission (ASIC), 2010, Report 215: Australian equity market structure, November. http://www.asic.gov.au/asic/pdflib.nsf/LookupByFileName/rep-215.pdf/$file/rep-215.pdf (accessed November 10, 2015).

Australian Securities & Investments Commission (ASIC), 2013, Report 331: Dark liquidity and high-frequency trading.

Authority for the Financial Markets (AFM), 2010, High frequency trading: The application of advanced trading technology in the European marketplace. http://www.afm.nl/layouts/afm/default.aspx~/media/files/rapport/2010/hft - reportengels.ashx (accessed December 10, 2014).

Bachelier, L. (1900) Theory of Speculation, In *The Random Character of Stock Prices*, P.H. Cootner, ed. Cambridge, MA: MIT Press (1964), 17-78.

Back, K. (1992). Insider trading in continuous time. *Review of financial Studies, 5*(3), 387-409.

Back, K., Cao, C. H., & Willard, G. A. (2000). Imperfect competition among informed traders. *The Journal of Finance, 55*(5), 2117-2155.

Back, K. E. (2017). *Asset Pricing and Portfolio Choice Theory*: Oxford University Press, Incorporated.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal, 27*(2), 12-14.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of accounting research*, 159-178.

Ballard, D. H. (1999). *An introduction to natural computation*: Mit Press.

Banner, S. (1998). The origin of the New York stock exchange, 1791–1860. *The Journal of Legal Studies, 27*(1), 113-140.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307-343.

Baron, M., Brogaard, J., & Kirilenko, A. The trading profits of high frequency traders.

Barron, O. E., Harris, D. G., & Stanford, M. (2005). Evidence that investors trade on private event-period information around earnings announcements. *The Accounting Review, 80*(2), 403-421.

Barrouillet, P. (2011). Dual-process theories and cognitive development: Advances and challenges. *Developmental Review, 31*(2), 79-85.

Bartov, E. (1993). The timing of asset sales and earnings manipulation. *Accounting Review*, 840-855.

Baruch, S. (2002). Insider trading and risk aversion. *Journal of financial markets, 5*(4), 451-464.

Bem, D. J. (1965). An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology*.

Benartzi, S., & Thaler, R. H. (2007). Heuristics and biases in retirement savings behavior. *The Journal of Economic Perspectives, 21*(3), 81-104.

Bencivenga, V. R., & Smith, B. D. (1991). Financial intermediation and endogenous growth. *The Review of Economic Studies, 58*(2), 195-209.

Bennett, K. P., & Campbell, C. (2000). Support vector machines: hype or hallelujah? *ACM SIGKDD Explorations Newsletter, 2*(2), 1-13.

Berk, J. B., Green, R. C., & Naik, V. (1999). Optimal investment, growth options, and security returns. *The Journal of Finance, 54*(5), 1553-1607.

Bernhardt, D., & Miao, J. (2004). Informed trading when information becomes stale. *The Journal of Finance, 59*(1), 339-390.

Berson, A., Smith, S., & Thearling, K. (2000). *Building data mining applications for CRM*: McGraw-Hill New York.

Beaulac, G. (2010). A two speed mind?: for a heuristic interpretation of dual-process theories= L'esprit à deux vitesses?: pour une interprétation heuristique des théories à processus duaux.

Biais, B., & Foucault, T. (2014). HFT and market quality. *Bankers, Markets & Investors, 128*, 5-19.

Biais, B., Foucault, T., & Moinas, S. (2011). Equilibrium High Frequency Trading.

Biais, B., Foucault, T., & Moinas, S. (2012). *Equilibrium high-frequency trading.* Paper presented at the AFA 2013 San Diego Meetings Paper.

Biais, B., & Woolley, P. (2011). High frequency trading. *Manuscript, Toulouse University, IDEI*.

Black, F. (1971). Toward a fully automated stock exchange, part I. *Financial Analysts Journal, 27*(4), 28-35.

Black, F. (1986). Noise. *The journal of finance, 41*(3), 528-543.

Boden, M. (2008). *Mind as Machine: A History of Cognitive Science*: Oxford University Press.

Boehmer, E., Fong, K. Y., & Wu, J. J. (2015). International evidence on algorithmic trading.

Boulatov, A. (2005). Uniqueness of equilibrium in the single-period Kyle'85 Model: working paper.

Boulatov, A., Kyle, A.S., & Livdan, D. (2012). Uniqueness of Equilibrium in the Single - Period Kyle 1985 Model.

Boyd, J., & Smith, B. (1996). The coevolution of the real and financial sectors in the growth process. *The World Bank Economic Review, 10*(2), 371-396.

Brownlees, C. T., Cipollini, F., & Gallo, G. M. (2011). Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics, 9*(3), 489-518.

Brunnermeier, M. K. (2001). *Asset pricing under asymmetric information: Bubbles, crashes, technical analysis, and herding*: Oxford University Press on Demand.

Brunnermeier, M. K. (2002). Buy on rumors–sell on news. *Unpublished working paper*.

Brogaard, J. (2010). High frequency trading and its impact on market quality. *Northwestern University Kellogg School of Management Working Paper*.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial studies, 27*(8), 2267-2306.

Brown, L. D., Richardson, G. D., & Schwager, S. J. (1987). An information interpretation of financial analyst superiority in forecasting earnings. *Journal of accounting research, 25*(1), 49-67.

Brownlees, C. T., Cipollini, F., & Gallo, G. M. (2011). Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics, 9*(3), 489-518.

Carmo, & Luis. (2015). *Handbook of Research on Global Competitive Advantage through Innovation and Entrepreneurship*: IGI Global.

Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ. *Journal of Financial Markets, 16*(4), 680-711.

Carruthers, P., Stich, S., & Siegal, M. (2002). The cognitive basis of science: Cambridge University Press.

Cartea, Á., & Penalva, J. (2010). Where is the Value in High Frequency Trading?

Cartea, Á., & Penalva, J. (2012). Where is the value in high frequency trading? *The Quarterly Journal of Finance, 2*(03), 1250014.

Castiglione, F. (2001). Forecasting price increments using an artificial Neural Network. *Advances in Complex Systems, 4*(01), 45-56.

Chaboud, A., Hjalmarsson, E., Vega, C., & Chiquoine, B. (2009). Rise of the machines: Algorithmic trading in the foreign exchange market. *FRB International Finance Discussion Paper*(980).

Chalmers, J. M., & Kadlec, G. B. (1998). An empirical examination of the amortized spread. *Journal of Financial Economics, 48*(2), 159-188.

Chan, H. W., & Faff, R. W. (2003). An investigation into the role of liquidity in asset pricing: Australian evidence. *Pacific-Basin Finance Journal, 11*(5), 555-572.

Chen, C.-M. (2016). Security Market Microstructure. *SSRN eLibrary*.

Chordia, T., Roll, R., & Subrahmanyam, A. (2008). Why has trading volume increased? *SSRN eLibrary*.

Chung, D. Y., & Hrazdil, K. (2012). Speed of convergence to market efficiency: The role of ECNs. *Journal of Empirical Finance, 19*(5), 702-720.

Chung, K. H., & Lee, A. J. (2016). High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World. *Asia-Pacific Journal of Financial Studies, 45*(1), 7-33.

Committee of European Securities Regulators (CESR), 2010, Call for Evidence. Microstructural issues of the European equity markets (April 1). http://www.cesr.eu/data/document/10_142.pdf (accessed January 19, 2014).

Conrad, J., Wahal, S., & Xiang, J. (2015). High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics, 116*(2), 271-291.

Cootner, H. (1964). 1964, The random character of stock market prices. *MIT Press, 17*, 78.

Cukierman, A. (1992). *Central Bank Strategy, Credibility, and Independence: Theory and Evidence*: ICS Press.

Cvitanic, J., & Kirilenko, A. (2010). High frequency traders and asset prices. Available at SSRN 1569075.

Cvitanić, J., & Zapatero, F. (2004). *Introduction to the economics and mathematics of financial markets*: MIT press.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *The Journal of Finance, 53*(6), 1839-1885.

Das, R., Hanson, J. E., Kephart, J. O., & Tesauro, G. (2001). *Agent-human interactions in the continuous double auction.* Paper presented at the International Joint Conference on Artificial Intelligence.

De Jong, F., & Rindi, B. (2009). *The Microstructure of Financial Markets*: Cambridge University Press.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990a). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45*(2), 379-395.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990b). Noise Trader Risk in Financial Markets. *The Journal of Political Economy, 98*(4), 703-738.

Deb, P., Koo, B., & Liu, Z. (2014). Competition, premature trading and excess volatility. *Journal of Banking & Finance, 41*, 178-193.

Demski, J. S., & Feltham, G. A. (1994). Market response to financial reports. *Journal of Accounting and Economics, 17*(1), 3-40.

Dolvin, S. D., Jordan, B. D., & Miller Jr, T. W. (2012). *Fundamentals of investments: valuation and management*.

Domowitz, I., & Yegerman, H. (2006). The cost of algorithmic trading: A first look at comparative performance. *The Journal of Trading, 1*(1), 33-42.

Dugast, J., & Foucault, T. (2016). Data Abundance and Asset Price Informativeness.

Easley, D., Lopez De Prado, M., & O'Hara, M. (2012). The volume clock: Insights into the high frequency paradigm. *The Journal of Portfolio Management,(Fall, 2012) Forthcoming*.

Eiben, A. E., & Smith, J. E. (2003). *Introduction*: Springer.

Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. American psychologist, 28(5), 404.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. American psychologist, 49(8), 709.

Epstein, S. (2003). Cognitive-experiential self-theory of personality. Handbook of psychology.

European Commission, 2010, European Commission Directorate General Internal Market and Services, Public consultation: Review of the Markets in Financial Instruments Directive(MiFID).http://ec.europa.eu/internal_market/consultations/docs/2010/mifid/consultation_paper_en.pdf (accessed January 19, 2014).

Evans, J. S. B. (1989). Bias in human reasoning: Causes and consequences: Lawrence Erlbaum Associates, Inc.

Evans, J. S. B., & Over, D. E. (1996). Rationality and reasoning: Psychology Press.

Evans, J. S. B. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation.

Evans, J. S. B. (2007). Hypothetical thinking: Dual processes in reasoning and judgement (Vol. 3): Psychology Press.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. Annu. Rev. Psychol., 59, 255-278.

Evans, J. S. B., & Frankish, K. E. (2009). In two minds: Dual processes and beyond: Oxford University Press.

Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing, 11*(3), 7-22.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics, 49*(3), 283-306.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383-417.

Farmer, J., & Lillo, F. (2004). On the origin of power laws in financial markets. *Quantitative finance, 314*, 7-10.

Farmer, J. D., Gillemot, L., Lillo, F., Mike, S., & Sen, A. (2004). What really causes large price changes? *Quantitative Finance, 4*, 1-15.

Fodor, J. A. (1975). The language of thought (Vol. 5): Harvard University Press.

Fodor, J. A. (1981). Representations: Philosophical essays on the foundations of cognitive science.

Fodor, J. A. (1987). Psychosemantics: The problem of meaning in the philosophy of mind: The MIT Press.

Fodor, J. A. (1990). A theory of content and other essays: The MIT press.

Fodor, J. A. (1995). The elm and the expert: Mentalese and its semantics: MIT press.

Foster, E. A. (2002). *Commodity futures price prediction: an artificial intelligence approach.* uga.

Foster, F. D., & Viswanathan, S. (1994). Strategic trading with asymmetrically informed traders and long-lived information. *Journal of financial and Quantitative Analysis, 29*(04), 499-518.

Foster, F. D., & Viswanathan, S. (1996). Strategic trading when agents forecast the forecasts of others. *The Journal of Finance, 51*(4), 1437-1478.

Foucault, T., Hombert, J., & Rosu, I. (2012). News Trading and Speed. *Available at SSRN.*

Foucault, T., Hombert, J., & Roşu, I. (2016). News trading and speed. *The Journal of Finance, 71*(1), 335-382.

Fox, M. B., Glosten, L. R., & Rauterberg, G. V. (2015). The new stock market: sense and nonsense. *Duke LJ, 65*, 191.

Francioni, R., Hazarika, S., Reck, M., & Schwartz, R. A. (2008). Equity market microstructure: taking stock of what we know. *The Journal of Portfolio Management, 35*(1), 57-71.

Frankish, K. (2007). Mind and supermind: Cambridge University Press.

Friedenberg, J., & Silverman, G. (2011). Cognitive science: an introduction to the study of mind: Sage

Frino, A., Mollica, V., & Webb, R. I. (2014). The Impact of Co-Location of Securities Exchanges' and Traders' Computer Servers on Market Liquidity. Journal of Futures Markets, 34(1), 20-33.

Froot, K. A., Scharfstein, D. S., & Stein, J. C. (1992). Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation. *Journal of Finance*, 1461-1484.

Fudenberg, D., & Tirole, J. (1991). *Game Theory*: MIT Press.

Gallant, A. R., Rossi, P. E., & Tauchen, G. (1992). Stock prices and volume. *Review of financial studies, 5*(2), 199-242.

Gamzo, R. A. (2014). *Algorithmic Trading, market efficiency and the momentum effect*. Masters Thesis, University of the Witwatersrand, Johannesburg.

Geisst, C. R. (2012). *Wall Street: A History, Updated Edition*: OUP USA.

Gillemot, L., Farmer, J. D., & Lillo, F. (2006). There's more to volatility than volume. *Quantitative finance, 6*(5), 371-384.

Goldberger, A. S. (1991). *A Course in Econometrics*: Harvard University Press.

Goldstein, M. A., Kumar, P., & Graves, F. C. (2014). Computerized and High-Frequency Trading. Financial Review, 49(2), 177-202.

Gomber, P., Arndt, B., Lutat, M., & Uhle, T. (2011). High-frequency trading. Available at SSRN 1858626.

Gosnell, T. F., Keown, A. J., & Pinkerton, J. M. (1996). The intraday speed of stock price adjustment to major dividend changes: Bid-ask bounce and order flow imbalances. *Journal of Banking & Finance, 20*(2), 247-266.

Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 611-642.

Griffin, J. M., Ji, X., & Martin, J. S. (2003). Momentum investing and business cycle risk: Evidence from pole to pole. *The Journal of Finance, 58*(6), 2515-2547.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), 393-408.

Grossman, H. I. (1979). Adverse selection, dissembling, and competitive equilibrium. *The Bell Journal of Economics*, 336-343.

Grossman, S. (1976). On the efficiency of competitive stock markets where trades have diverse information. *The Journal of Finance, 31*(2), 573-585.

Groth, S. S. (2011). Does Algorithmic Trading Increase Volatility? Empirical Evidence from the Fully-Electronic Trading Platform Xetra.

Groß-Klußmann, A., & Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. Journal of Empirical Finance, 18(2), 321-340.

Gsell, M. (2008). Assessing the impact of Algorithmic Trading on markets: A simulation approach: CFS Working Paper.

Hanson, T. A. (2014). *Real Effects of High Frequency Trading.* Kent State University.

Hasbrouck, J., & Saar, G. (2011). Low-latency trading. *Johnson School Research Paper Series*(35-2010).

Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets, 16*(4), 646-679.

Hagströmer, B., & Nordén, L. (2013). The diversity of high-frequency traders. Journal of Financial Markets, 16(4), 741-770.

Hayek, F. A. (1945). The use of knowledge in society. *The American economic review*, 519-530.

Heckman, J. J. (2008). Econometric causality. *International statistical review, 76*(1), 1-27.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance, 66*(1), 1-33.

Hendershott, T., & Moulton, P. C. (2011). Automation, speed, and stock market quality: The NYSE's hybrid. *Journal of Financial Markets, 14*(4), 568-604.

Hendershott, T., & Riordan, R. (2009). Algorithmic Trading and information. Manuscript, University of California, Berkeley.

Hendershott, T., & Riordan, R. (2011). High frequency trading and price discovery. *Manuscript, University of California, Berkeley, 3*.

Hendershott, T., & Riordan, R. (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis, 48*(04), 1001-1024.

Hoyningen-Huene, P. (2008). Systematicity: The Nature of Science. Philosophia, 36, 167-180.

Hirshleifer, J. (1971). The private and social value of information and the reward to inventive activity. *The American economic review, 61*(4), 561-574.

Hirshleifer, D., Subrahmanyam, A., & Titman, S. (1994). Security analysis and trading patterns when some investors receive information before others. *The Journal of Finance, 49*(5), 1665-1698.

Hoffmann, P. (2014). A dynamic limit order market with fast and slow traders. *Journal of Financial Economics, 113*(1), 156-169.

Holden, C. W., & Subrahmanyam, A. (1992). Long-lived private information and imperfect competition. *The Journal of Finance, 47*(1), 247-270.

Holden, C. W., & Subrahmanyam, A. (2002). News Events, Information Acquisition, and Serial Correlation. *The Journal of Business, 75*(1), 1-32.

Holthausen, R. W., & Verrecchia, R. E. (1990). The effect of informedness and consensus on price and volume behavior. *Accounting Review*, 191-208.

Hong, D., Lee, C., & Swaminathan, B. (2003). Earnings momentum in international markets. *Available at SSRN 390107*.

Humphery-Jenner, M. (2011). High Frequency Trading, Information, and Takeovers: Tilburg University, Center for Economic Research.

Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance, 49*(3), 851-889.

Indjejikian, R. J. (1991). The impact of costly information interpretation on firm disclosure decisions. *Journal of accounting research, 29*(2), 277-301.

Jain, P. (2005). Financial Market Design and the Equity Premium: Electronic vs. Floor Trading. Journal of Finance 60:2955-2985.

Jarnecic, E., & Snape, M. (2014). The Provision of Liquidity by High-Frequency Participants. *Financial Review, 49*(2), 371-394.

Johnson, B. (2010). *Algorithmic Trading & DMA: An introduction to direct access trading strategies*: 4Myeloma Press.

Johnson, S. A., Wang, Q. E., & Zhang, J. (2014). The Effect of Algorithmic Trading on Firm Value. Available at SSRN.

Jones, C. M. (2013). What do we know about high-frequency trading?

Jovanovic, B., & Menkveld, A. (2012). Middlemen in securities markets. *Unpublished working paper. VU Amsterdam, Amsterdam*.

Kang, J., Liu, M. H., & Ni, S. X. (2002). Contrarian and momentum strategies in the China stock market: 1993–2000. *Pacific-Basin Finance Journal, 10*(3), 243-265.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis, 22*(01), 109-126.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics, 17*(2), 357-390.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. Cognition, 11(2), 123-141.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions.

Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes. Behavioral and Brain Sciences, 23(05), 681-683

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. American psychologist, 58(9), 697.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment.

Kahneman, D., & Frederick, S. (2005). A Model of Heuristic Judgment.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. American economic review, 1449-1475.

Kahneman, D. (2011). Thinking, fast and slow: Macmillan.

Kim, O., & Verrecchia, R. E. (1991). Trading volume and price reactions to public announcements. *Journal of accounting research*, 302-321.

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of accounting and economics, 17*(1), 41-67.

Kim, O., & Verrecchia, R. E. (1997). Pre-announcement and event-period private information. *Journal of Accounting and Economics, 24*(3), 395-419.

Kirilenko, A. A., & Lo, A. W. (2013). Moore's Law Versus Murphy's Law: Algorithmic trading and its discontents. *The Journal of Economic Perspectives, 27*(2), 51-72.

Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2011). The flash crash: The impact of high frequency trading on an electronic market. *Manuscript, U of Maryland*.

Kissell, R. (2013). The science of algorithmic trading and portfolio management: Academic Press.

Korsmo, C. (2014). High-Frequency Trading: A Regulatory Strategy.

Kuhn, T. S. (1962). The structure of scientific revolutions: University of Chicago press.

Kyle, A.S. (1985a). *International Agricultural Trade: Advanced readings in Price Formation,Market Structure, and Price Instability*, Chapter Market structure, information, futures market, and price formation, pp. 45-64. Boulder and London: West view Press.

Kyle, A. S.(1985b), Equilibrium in a Speculative Market With Strategic Informed Trading, Princeton University. Working Paper.

Kyle, A. S. (1985c). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315-1335.

Lamoureux, C. G., & Lastrapes, W. D. (1990). Heteroskedasticity in stock return data: volume versus GARCH effects. *The Journal of Finance, 45*(1), 221-229.

LeBaron, B. (2006). Agent-based computational finance. *Handbook of computational economics, 2*, 1187-1233.

Leinweber, D. J. (2009). *Nerds on Wall Street: Math, machines and wired markets*: John Wiley and Sons.

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *Review of financial studies, 12*(5), 1113-1141.

Lewellen, J., & Shanken, J. (2002). Learning, asset-pricing tests, and market efficiency. *The Journal of Finance, 57*(3), 1113-1145.

Li, S. (2013). Imperfect Competition, Long Lived Private Information, and the Implications for the Competition of High Frequency Trading. *SSRN eLibrary*.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics, 47*(1), 13-37.

Lieberman, M. D. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach.

Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol., 58*, 259-289.

Lieberman, M. D. (2009). What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness.

Lo, A. W., & MacKinlay, A. C. (1999). A Non-Random Walk Down Wall Street.

Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45*(2), 379-395.

Lutat, M. (2010). The effect of maker-taker pricing on market liquidity in electronic trading systems–empirical evidence from European equity trading.

Lyons, R. K. (2001). *The microstructure approach to exchange rates* (Vol. 333): Citeseer.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets, 3*(3), 205-258.

Makarov, I., & Rytchkov, O. (2012). Forecasting the forecasts of others: Implications for asset pricing. *Journal of Economic Theory, 147*(3), 941-966.

Martinez, V., & Rosu, I. (2011). High frequency traders, news and volatility. *News and Volatility (December 29, 2011)*.

Maurer, S. B., & Ralston, A. (2005). *Discrete Algorithmic Mathematics, Third Edition*: CRC Press.

Mazza, P. (2015). Rethinking Zero Returns in the Liquidity Puzzle of a Limit Order Market. *Finance, 36*(2), 7-36.

McCarthy, J. G. (2000). *Existence of equilibrium and other fundamental properties in a class of insider trading models*: University of California, Berkeley.

McKelvey, B. (1978). Organizational systematics: Taxonomic lessons from biology. *Management Science, 24*(13), 1428-1440.

McNamara, S. R. (2016). The Law and Ethics of High-Frequency Trading. *Minnesota Journal of Law, Science & Technology, 17*(1), 71.

McNichols, M., & Trueman, B. (1994). Public disclosure, private information collection, and short-term trading. *Journal of Accounting and Economics, 17*(1), 69-94.

Melamed, L. (2009). For crying out loud: From open outcry to the electronic screen: John Wiley & Sons.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics, 8*, 1-24.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information.

Mike, S., & Farmer, J. D. (2008). An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics and Control, 32*(1), 200-234.

Mills, T. C. (1991). Nonlinear time series models in economics. *Journal of Economic Surveys, 5*(3), 215-242.

Moore, A. B. (1962). *A statistical analysis of common stock prices.* University of Chicago.

Moses, O. D. (1987). Income smoothing and incentives: Empirical tests using accounting changes. *Accounting Review*, 358-377.

Muga, L., & Santamaria, R. (2007). The momentum effect: omitted risk factors or investor behaviour? Evidence from the Spanish stock market. *Quantitative Finance, 7*(6), 637-650.

Nanda. (2002). *Discrete Mathematics*: Allied Publishers.

Narang, R. K. (2013). *Inside the Black Box: A Simple Guide to Quantitative and High Frequency Trading* (Vol. 883): John Wiley & Sons.

O'Hara, M. (1995). *Market Microstructure Theory*: Wiley.

O'Hara, M. (2014). High Frequency Market Microstructure. Johnson Graduate School of Management Working Paper, New York.

Obstfeld, M. (1992). Risk-taking, global diversification, and growth: National bureau of economic research.

Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations research, 7*(2), 145-173.

Osborne, M. J., & Rubinstein, A. (1994). *A Course in Game Theory*: MIT Press.

Oesterle, D. A. (2005). Regulation NMS: Has the SEC Exceeded its Congressional Mandate to Facilitate a" National Market System" in Securities Trading? *The Ohio State University Moritz College of Law Working Paper Series*, 12.

Pagnotta, E., & Philippon, T. (2011). Competing on speed: National Bureau of Economic Research.

Parlour, C. A. (1998). Price dynamics in limit order markets. *Review of financial studies, 11*(4), 789-816.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of political Economy, 111*(3), 642-685.

Piccinini, G., & Bahar, S. (2013). Neural computation and the computational theory of cognition. Cognitive science, 37(3), 453-488.

Platen, E., & Heath, D. (2006). *A Benchmark Approach to Quantitative Finance*: Springer Berlin Heidelberg.

Prix, J., Loistl, O., & Huetl, M. (2007). Algorithmic trading patterns in Xetra orders. *The European Journal of Finance, 13*(8), 717-739.

Pylyshyn, Z. W. (1978). Computational models and empirical constraints. Behavioral and Brain Sciences, 1(01), 91-99.

Pylyshyn, Z. W. (1980). Computation and cognition: Issues in the foundations of cognitive science. Behavioral and Brain Sciences, 3(01), 111-132.

Pylyshyn, Z. W. (1984). Computation and cognition: Cambridge Univ Press.

Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*: Clarendon Press Oxford.

Qin, X. (2012). Making use of the big data: next generation of algorithm trading *Artificial Intelligence and Computational Intelligence* (pp. 34-41): Springer.

Rajan, R. G., & Zingales, L. (1996). Financial dependence and growth: National bureau of economic research.

Rasmusen, E. (2001). *Games and Information: An Introduction to Game Theory*: Blackwell.

Rochet, J.-C., & Vila, J.-L. (1994). Insider trading without normality. *The Review of Economic Studies, 61*(1), 131-152.

Ronen, J., & Sadan, S. (1981). Smoothing income numbers: Objectives, means, and implications: Addison-Wesley Pub. Co.(Reading, Mass.).

Rouwenhorst, K. G. (1998). International Momentum Strategies. *The Journal of Finance, 53*(1).

Saint-Paul, G. (1992). Technological choice, financial markets and economic development. *European Economic Review, 36*(4), 763-781.

Salmon & Stokes (2010). *Algorithms Take Control of Wall Street*, WIRED. Retrieved, 27 December, 2012, from https://www.wired.com/2010/12/ff_ai_flashtrading/.

Samuels, R. (2009). The magical number two, plus or minus: Dual-process theory as a theory of cognitive kinds.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance, 44*(5), 1115-1153.

Securities and Exchange Commission (SEC), 2010, Findings Regarding The Market Events Of May 6, 2010. *Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues*.

Securities and Exchange Commission (SEC), 2014, Equity Market Structure Literature Review Part II: High Frequency Trading. *Staff of the Division of Trading and Markets*. Available from: https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf.

Shapiro, C., & Varian, H. R. (2013). *Information rules: a strategic guide to the network economy*: Harvard Business Press.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance, 19*(3), 425-442.

Shah, V. H. (2007). Machine learning techniques for stock prediction. Foundations of Machine Learning| Spring.

Singleton, K. J. (1988). Econometric issues in the analysis of equilibrium business cycle models. *Journal of Monetary Economics, 21*(2), 361-386.

Snape, M., & Jarnecic, E. An Analysis of Trades by High Frequency Participants on the London Stock Exchange.

Snyman, H. A. (2011). *Investigating momentum on the Johannesburg Stock Exchange.* Stellenbosch: University of Stellenbosch.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. Psychological bulletin, 119(1), 3.

Sloman, S. A. (2002). Two systems of reasoning.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. Personality and social psychology review, 4(2), 108-131.

Smith, R. D. (2010). Is high-frequency trading inducing changes in market microstructure and dynamics? arXiv.org.

Stanovich, K. E. (1999). Who is rational?: Studies of individual differences in reasoning: Psychology Press.

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. In two minds: Dual processes and beyond, 55-88.

Stanovich, K. (2011). Rationality and the reflective mind: Oxford University Press.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. Behavioral and Brain Sciences, 23(05), 701-717.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? Behavioral and Brain Sciences, 23, 645-726.

Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. H., Garfield, J. L., & Rissland, E. L. (1995). Cognitive Science An Introduction Second Edition.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. Personality and social psychology review, 8(3), 220-247.

Suhas Yadav, M., & Patil, S. (2011). Intelligent Learning Of Fuzzy Logic Controllers Via Neural Network And Genetic Algorithm.

Tauchen, G. E., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society*, 485-505.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*: CUP Archive.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. science, 185(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological review, 90(4), 293.

Van Gestel, T., Suykens, J. A., Baestaens, D.-E., Lambrechts, A., Lanckriet, G., Vandaele, B., Vandewalle, J. (2001). Financial time series prediction using least squares support vector machines within the evidence framework. *IEEE Transactions on neural networks, 12*(4), 809-821.

Vandewalle, J. (2001). Financial time series prediction using least squares support vector machines within the evidence framework. *Neural Networks, IEEE Transactions on, 12*(4), 809-821.

Varian, H. R. (1989). Differences of Opinion in Financial Markets. *Financial Risk: Theory, Evidence and Implications*.

Vayanos, D. (1999). Strategic trading and welfare in a dynamic market. *The Review of Economic Studies, 66*(2), 219-254.

Vayanos, D., & Wang, J. (2012). Market Liquidity--Theory and Empirical Evidence: National Bureau of Economic Research

Vives, X. (2010). *Information and learning in markets: the impact of market microstructure*: Princeton University Press.

Vo, M. T. (2008). Strategic trading when some investors receive information before others. *International Review of Economics & Finance, 17*(2), 319-332.

Walczak, S. (2001). An empirical analysis of data requirements for financial forecasting with neural networks. *Journal of management information systems, 17*(4), 203-222.

Wang, J. (1994). A model of competitive stock trading volume. *Journal of political Economy*, 127-168.

Weber, P., & Rosenow, B. (2006). Large stock price changes: volume or liquidity? *Quantitative finance, 6*(1), 7-14.

Wilkinson, V. E., Jackson, M. L., Westlake, J., Stevens, B., Barnes, M., Swann, P., . . . Howard, M. E. (2013). The accuracy of eyelid movement parameters for drowsiness detection. *J Clin Sleep Med, 9*(12), 1315-1324.

Wimsatt, W. C. (1986). Forms of aggregativity *Human nature and natural knowledge* (pp. 259-291): Springer.

Wolfson, N., & Russo, T. A. (1970). The stock exchange specialist: An economic and legal analysis. *Duke Law Journal, 1970*(4), 707-746.

Wu, H. (2010). An Empirical Bayesian Approach to Misspecified Covariance Structures. The Ohio State University.

Ye, M., Yao, C., & Gai, J. (2013). The externalities of high frequency trading.

Yoon, M. H. (2010). Trading in a flash: implication of high-frequency trading for securities regulators worldwide. *Emory Int'l L. Rev., 24*, 913.

Zhang, F. (2010). High-Frequency Trading, Stock Volatility, and Price Discovery. SSRN Electronic Journal.

Zhang, F., & Powell, S. B. (2011). The impact of high-frequency trading on markets. *CFA Magazine, 22*(2), 10-11.

Zhang, S. S. (2012). Need for speed: An empirical analysis of hard and soft information in a high frequency world.

APPENDIX I

Introduction

Appendix I evaluates existing empirical research on the infrastructural aspects of algorithmic trading in order to determine the extent to which empirical research supports our dual process supposition of algorithmic trading. However, readers are advised to note the following: Algorithmic trading infrastructure and financial market regulation are not mutually exclusive constructs. Inexorably, the regulatory environment that prefaced (and enabled) algorithmic trading merits some discussion before an investigation on the infrastructural aspects of algorithmic trading (vis-à-vis co-location and high-end capability computer infrastructure) can be justified.

The discussion to follow provides some historical context to the practice of algorithmic trading by detailing the economic conditions that led to the advent of the practice. The regulatory environment that prefaced algorithmic trading is defined and discussed below - we highlight some key structural (i.e., regulatory) changes that fostered the development of algorithmic trading. Co-location and high-end capability computer infrastructure are addressed explicitly in subsections (A) and (B) of this appendix.

Historical Overview of the Changing Trading Environment

The history of the stock markets in the United States stretches back to 1792, when stock trading was said to have been organized under a buttonwood tree on Wall Street. The infamous 'Buttonwood Agreement' was signed by twenty-four brokers and specified the conditions for trading. Accordingly, brokers would be allowed to trade only amongst themselves; with a minimum commission rate of 0.25%.

These two features of trading on what became the New York Stock Exchange (NYSE) formed the cornerstone of the trading system that persisted until the twenty-first century.

Banner (1998) chronicles the NYSE's growth from its Buttonwood beginnings to what became the centre of the American economy. Under the Securities Exchange Act of 1934 (henceforth, the Exchange Act) the NYSE became a 'predictable'[195] system for buying and selling equity in public companies. Trading was traditionally limited to 'broker-dealers'[196] holding seats on the exchange, who traded in the context of *specialist* system. McNamara (2016) outlines the role of the *specialist* system in the stock market. Briefly, a specialist in a company's stock was obligated to support trading when liquidity dried up (see also, Wolfson & Russo, 1970).[197] Thus, in addition to broker-dealers who were allowed to trade on behalf of their clients, specialists supported trading in NYSE-listed stocks. And while specialists could see the orders in the 'order book', and thus had knowledge of where the market was headed, 'front-running' was strictly prohibited. Technically speaking, the term front-running applies to the illegal practice addressed in FINRA Rule 5270, which forbids a specialist from trading for its own account while taking advantage of knowledge of an imminent transaction. As noted by Geisst (2012): "specialists were in a privileged position to see prices before executing for the public and would often act for themselves before filling an order from the public being executed through a broker"(p.207).

[195] To paraphrase Loy and Chi (2014): "The original conception of federal securities laws ushered in an unprecedented era of financial stability" (p.669).

[196] Contextually, a broker buys or sells a stock on behalf of a third-party, typically an institutional investor. Dealers, also known as market-makers, ensure trading by purchasing at the bid price and selling at the ask price.

[197] A specialist has to manage the trading and quote process and has to guarantee the provision of liquidity, when necessary, by taking the other side of the market. Essentially, a specialist acts as either a broker or dealer but only for a specific list of stocks that he or she is responsible for

In contemporary financial markets, front-running is prosecuted as a violation of Section 10(b) of the Exchange Act (see e.g. Fox, Glosten & Rauterberg, 2015). However, it is important to recognise that the legal restrictions surrounding front-running apply only to specialists and/or broker-dealers who trade on behalf of clients or customers (algorithmic traders do not fall under the specialists and/or broker-dealers rubric, and are thus not subject to Section 10(b) of the Exchange Act prohibiting front running. This regulatory doctrine has important implications on the legality of algorithmic trading and is addressed explicitly in the discussion to follow.)[198]

The classic twentieth century form of the exchange was for the most part an 'open outcry system'. In the open outcry auction system, trading occurred at the post on the floor of the exchange where a specialist was located, with brokers *shouting* bids or offers to one another.

By the 1960s, the volume of traded shares was overwhelming the traditional paper systems that brokers, dealers, and specialists on the trading floor used. This 'paperwork crisis' put significant pressure on floor traders and seriously affected operations on the NYSE. This crisis led to the Securities and Exchange Commission (SEC)[199] recommending changes to the industry in 1963.

In response to the SEC's call to reform, the National Association of Securities Dealers created the NASDAQ in 1968. The National Association of Securities Dealers Automated Quotations (NASDAQ) represented a fundamentally new mode of equity trading both in its use of technology

[198] With respect to algorithmic trading – the central topic of this thesis - some caveats are in order. As suggested in the narrative to follow, algorithmic traders are proprietary enterprises. Unlike specialists and/or broker-dealers, algorithmic traders do not act on behalf of clients or customers. They are thus exempt from the legal restrictions mandated by Section 10(b) of the Exchange Act.

[199] The Securities and Exchange Commission (SEC) was created by Congress in 1934 by the Securities Exchange Act. Since then, it has acted as the regulator of the stock exchanges and the companies that list on them. Over time, the SEC and Wall Street have evolved together, influencing each other in the process.

and its format.[200] It provided its members with a way to trade stocks that would have difficulty meeting the NYSE listing requirements.[201] It also incorporated the use of computer technology, and so was the precursor to today's electronic trading systems. Orders were displayed electronically on the NASDAQ trading system. Brokers put up quotes to buy or sell stocks on an electronic bulletin board, and waited for counterparties to accept them. However, by offering an alternate market for trading, the introduction of the NASDAQ had some unintended consequences - alas, the lack of a centralised trading venue meant that trading became more fragmented.

In an attempt to reconcile the fragmented state of financial markets, the SEC passed the Regulation of the National Market System (more commonly known as Reg NMS). This legislature mandated the interconnectedness of various markets for stocks. Given the elementary role played by Reg NMS in the progression, and now dominance of algorithmic trading, some discussion on this regulatory framework is warranted.

Regulation National Market System

Foreseeing the problems of fragmented securities markets, Congress authorized the SEC to use its authority to facilitate the establishment of a National Market System (NMS) for securities. Regulation National Market System (henceforth, Reg NMS) has its genesis in Section 11A of the Exchange Act and came into effect in 2007. Reg NMS is "premised on promoting fair competition among individual markets, while at the same time assuring that all of these markets are linked

[200] The NASDAQ was the first financial market to incorporate computers in the trading process (Korsmo, 1971).

[201] These stocks were in most part issued by smaller start-ups; often technology ventures. See generally; https://www.nyse.com/publicdocs/nyse/listing/NYSE_Initial_Listing_Standards_Summary.pdf (last visited Aug. 12, 2015).

together, through facilities and rules, in a unified system that promotes interaction among the orders of buyers and sellers in a particular NMS stock" (McNamara, 2016, p.84).

Under Reg NMS, trading on the many exchanges in the United States would be conducted through the mechanism of a *single* price, known as the National Best Bid and Offer (NBBO).[202]

As allude to in Chapter 2, Section 2.3, Regulation NMS consists of Exchange Act Rules 600-612. The key provisions of Reg NMS were promulgated in 2005, and include: Rule 611 (the Order Protection Rule), Rule 610 (the Access Rule), Rule 612 (the Sub-Penny Rule), and amendments to Rules 601 and 603 (the Market Data Rules).

Rule 611 (*the Order Protection Rule*) is at the cornerstone of Regulation NMS. This rule was established to maintain and enforce written policies and procedures that prevent 'trade-throughs', or trades on one market at a price inferior to one available on another market. In this way, orders are protected from being executed at a price inferior to the NBBO.

Rule 610 (*the Access Rule*) prohibits trading centres from unfairly discriminating against non-members.[203] Specifically, the Access Rule provides for uniform access to quotes on the various trading venues by restricting the price charged for a quote to $0.003 per share.

Rule 612 (*the Sub-Penny Rule*) was designed to limit the ability of a market participant to gain execution priority over a competing 'limit order'[204] by stepping ahead by an economically

[202] The National Best Bid and Offer - NBBO - is the nationwide best available bid or ask price for a security and is determined using consolidated data from U.S. marketplaces. To paraphrase McNamara (2016): "The national best bid and national best offer means, with respect to quotations for an NMS security, the best bid and best offer for such security that are calculated and disseminated on a current and continuing basis by a plan processor pursuant to an effective national market system plan" (p.85).

[203] Technically speaking, non-members access quotations through an exchange member or subscriber.

[204] A limit order is a standing order to purchase or sell a certain amount of a given security at a given price.

insignificant amount. More precisely, Rule 612 restricts the use of limit orders that afford an agent with *undue* execution priority. In the pre-Reg NMS environment, traders would often use minuscule increments to jump ahead of standing limit orders. A trader could, for example, jump ahead of a bid for a given stock at $20.55, by bidding, say, $20.551.

The Sub-Penny Rule (Rule 612) thus prohibits quotes in prices less than $0.01 for securities valued at more than $1.00, and in prices less than $0.0001 for stocks valued at less than $1.00.

Finally, **Rules 601** and **603** (Market Data Rules) form the regulatory backdrop for the information systems (SIPs) that disseminate the NBBO and other information to market participants. According to Oesterle (2005), Securities Information Processors (SIPs) play a fundamental role in the National Market System and are responsible for providing each NMS security with 'core data'.[205] The literature regularly refers to core data as: the data required by Reg NMS to be submitted by a national securities exchange to a SIP (Fox, Glosten & Rauterberg, 2015; McNamara, 2016; Oesterle, 2005).

Under Reg NMS, any and all information dissemination systems that 'qualify'[206] as Securities Information Processors (SIPs) will be subject to Rules 601 and 603. Concomitantly, Market Data

[205] Accordingly, core data encompasses (1) last sale reports ;(2) the current highest bid and lowest offer for the security; and (3) the NBBO (Oesterle, 2005; Coffee, 2014; Fox et al., 2015; McNamara, 2016).

[206] *See* U.S.C. § 78c (a)(22)(A) (2014), for a characteristic based description of Securities Information Processors (SIP's).

Rules (Rules 601 and 603) provide for the contemporary version of the 'consolidated tape',[207] and serve the crucial function of disseminating a single NBBO for all exchange-listed stocks.

Note that Reg NMS in its entirety, including, but not limited to Exchange Act Rules 600-612, have all helped open the door for the practice algorithmic trading. However, it should be noted that Market Data Rules - specifically, Rules 603 (a) (1) and 603 (a) (2) – have become increasingly relevant in the current algorithmic trading environment. According to McNamara (2016), an important distinction needs to be made between Market Data Rules 603 (a) (1) and 603 (a) (2).

As suggested by McNamara (2016), the legal nuances between these respective Market Data Rules, although relatively subtle, have an important bearing on the practice of algorithmic trading. However, a critical evaluation of these legal protocols is beyond the scope of this thesis; rather a precise delineation of Rules 603 (a) (1) and 603 (a) (2) should suffice for purpose of the discussion to follow.

Specifically, Rule 603(a) (1) requires that any 'exclusive processor' [208] distributing information, with respect to quotations in an NMS stock to a securities information processor (SIP), shall do so on terms that are fair and reasonable. Fox et al., (2015) posit that this requirement is meant to allow

[207] The SEC (2012) defines the consolidated tape as an electronic system that reports the latest price and volume data on sales of exchange-listed stocks (See for example, https://www.sec.gov/answers/consolt.htm). There are currently four consolidated information dissemination systems inheriting the role of the consolidated tape in the U.S: The Consolidated Quotation System, the Consolidated Tape System, the NASDAQ System, and the OPRA System (see e.g., McNamara, 2016). However, a critical evaluation of the various consolidated information dissemination systems is beyond the scope of this research. Refer to Oesterle (2005) for a detailed discussion on the topic.

[208] The term "exclusive processor" means any securities information processor or self-regulatory organization which, directly or indirectly, engages on an exclusive basis on behalf of any national securities exchange which engages on an exclusive basis on its own behalf, in collecting, processing, or preparing for distribution or publication any information with respect to (i) transactions or quotations on or effected or made by means of any facility of such exchange or (ii) quotations distributed or published by means of any electronic system operated or controlled by such association (U.S.C. § 78c(a)(22)(B), 2014 - defining "exclusive processor").

a SIP to gather the relevant data it requires. Conversely, Rule 603(a) (2) allows for *exchanges* to disseminate information concerning quotes and transactions to other market participants on terms that are not unreasonably discriminatory. Contextually, Rule 603(a) (2) implies a less stringent standard than the one applicable to the requirement in 603(a) (1) - that an 'exclusive processor' shall provide such information to a SIP.

This stance has acquired far more relevance with the advent of algorithmic trading and the marketing of private data feeds to algorithmic trading firms. Exchanges' actively disseminate machine readable data to algorithmic traders. Notably, this activity - disseminating data to algorithmic traders - falls under Rule 603 (a)(2), not Rule 603 (a)(1). This important regulatory detail is explored in what follows; vis-à-vis co-location and high-end capability computer infrastructure.

An Overview of Algorithmic Trading Infrastructure

Although not ostensibly intuitive,[209] exchanges regularly engage in 'for-profit' activities; namely, they sell their own financial products. Co-location and high-end capability computing arrangements provide just some the more notable examples of these income generating activities.[210]

Co-location is a tangible manifestation of algorithmic trading. It refers to the exchanges' practice of renting space in the facilities that house their computer servers to traders who believe they can benefit from this proximity (McNamara, 2016). Co-location is explicitly defined and discussed in

[209] Exchanges are more commonly viewed as serving a *single* function of enabling trade.

[210] For example, the most recent 10K available of Intercontinental Exchange, Inc., the present operator of the NYSE, reports that data services revenue (which includes the fees charged for co-location facilities) amounted to $631 million in 2014, or 14.9% of total revenues of $4.221 billion. *See* Intercontinental Exch., Inc., 2014 Annual Report (Form 10-K).

the proceeding subsection of this appendix, Section A. Here we focus briefly on the data facilitating this practice. Specifically, one way in which exchanges' facilitate co-location is by offering direct data feeds to 'algorithmic traders'[211] who wish to co-locate.

The use of direct data feeds is permitted under the regulatory provisions of the SEC; provided that the said exchange disseminate such data to their algorithmic trader clients simultaneously with the provision of such data to the SIP.

Under this arrangement, data is released directly to a client (co-located firm) at precisely the same time that it is sent to the SIP. Notably, the inherent delays in transmission and processing by the SIP means that a direct data feeds are often a few milliseconds ahead of the SIP. The literature is congruent with this suggestion. In fact, Yoon (2010) asserts that direct data feeds are on average five to ten milliseconds faster than the NBBO disseminated by the SIP.

The question of the timing of the release of such data came up during the promulgation of Regulation NMS. However, the SEC has determined that the simultaneous release of market data to algorithmic traders and the SIP is consistent with the requirement of Rule 603(a) that information be released on terms that are "fair and reasonable" and "not unreasonably discriminatory."

High-end capability computing is another tangible manifestation of algorithmic trading. High-end capability infrastructure is explicitly defined and discussed in the proceeding subsection of this appendix, Section B. In line with the focus above, we highlight only the data facilitating this practice.

[211] While co-location facilities are completely voluntary and available to all traders on a non-discriminatory basis, the technical details of the practice - rooted in a computer driven system - limits their use to computer driven agents (i.e., algorithmic traders).

Notably, one way in which exchanges' are seen to facilitate this practice is by offering enriched data feeds to algorithmic traders who rely on high-end capability infrastructure. The proprietary, 'enriched' data feeds offered by the exchanges provide a greater breadth of information than those relying solely on the NBBO disseminated by the SIP can obtain. Whereas the NBBO disseminated by the SIP offers only traditional order book data, enriched data feeds will offer a variety of different forms (defined below) of textual, as well as numerical, information. (Leinweber, 2009).

Like in the case with direct data feeds, enriched data feeds engender questions of legality under Reg NMS. However, the SEC has also determined that the release of such market data by exchanges is consistent with the requirement of Rule 603(a) (2) - defined above.

Overall, the discussion above confirms the aforementioned nexus between Reg NMS and algorithmic trading. However, the discussion above serves only as a premise for the narrative to follow where algorithmic trading can be analysed vis-à-vis co-location and high-end capability computer infrastructure.

Empirical Support the Dual Process Proposition – Algorithmic Trading Infrastructure

A) Co-location Infrastructure

All algorithmic trading is strategic because its goal is generally to maximize a particular strategy against a market's matching engine (O'Hara, 2014). Effectively, the matching engine determines how orders are processed, and thus controls the automation of trades. This is where infrastructure comes to the fore as it establishes a channel of communication between the market – the matching engine – and the trading algorithm. To clarify, the ability to conduct algorithmic trading is predicated on infrastructure. As a prerequisite, algorithmic traders rely on *either*, co-location

infrastructure or high-end capability infrastructure (See, for e.g., Johnson, 2010; Gomber et al., 2011; Kissell, 2013; Frino, Mollica & Webb, 2014 and O'Hara, 2014).

The success of many algorithmic trading strategies – highlighted above – is dependent on co-location infrastructure. Essentially, co-location infrastructure allows firms to co-locate their servers next to the exchanges'. In other words, many exchanges now offer co-location services to specific algorithmic traders, allowing them to place their servers as close to the exchanges matching engine as possible.

Placing ones server adjacent to the exchanges matching engine means that real-time market information can reach the algorithmic traders platform almost instantaneously. It therefore, significantly reduces the time it takes to access the central order book – where electronic information on quotes to buy and sell as well as current market prices are warehoused. It also decreases the time it takes to transmit trade instructions and execute matched trades. In exchange for a fee, those who subscribe to co-location services get the infrastructure from the exchange itself. The package includes everything from the actual connection to the matching engine, to server cages, electricity, maintenance, and safety installations. Effectively, co-located firms are able to cut their latency in the access to news about order flow and also their order submission to the matching engine.

As co-location services are a material aspect of the operation of many of the algorithmic trading strategies, we argue that several of the algorithmic trading strategies, highlighted above, are simply predicated on an inherent speed/latency 'advantage'[212]- the time it takes to access and respond to

[212] Relevant to other market participants – those who do not use co-location. They exploit their very quick access to public information in an attempt to analyze the news and trade before everyone else - effectively turning public information into private information.

market information. This is supported by a plethora of other researchers, including Johnson (2010), Gomber et al. (2011), Kissell (2013), Frino, Mollica and Webb (2014) and O'Hara (2014). Remarkably, the current round-trip latency in the case of co-location is estimated to be less than 10 microseconds. According to O'Hara (2014), an ability to process market data (such as prices and order book information) before everyone else effectively allows these traders to turn public information into private information signals. To paraphrase O'Hara (2014): "an algorithmic trader's access to public information signals seconds (or even milliseconds) before they are seen by other traders…effectively turns public information into private information" (p.20). It follows that, even though the various strategies and algorithms can appear to be genuinely diverse, many of these strategies simply rely on high speed access to markets, i.e. the usage of co-location services.

Succinctly, many of the algorithmic trading strategies highlighted above - are simply latency dependent strategies. More precisely, latency, or the speed factor, is a crucial component in many of algorithmic trading strategies. In fact, 'evidence'[213] suggest that the typical strategies that employ co-location infrastructure operate in the 'microsecond'[214] environment - where 1 microsecond is roughly 300000 times the speed of an average blink of an eye.[215] Arguably, this ultra-fast, special class of strategies exploits their very quick access to public information in an attempt to analyse the news and trade before everyone else - effectively turning public information into private information. (O'Hara 2014).

[213] The SEC (2010) notes that: "the speed of trading has increased to the point that the fastest traders now measure their latencies in microseconds" (page 3605).

[214] Where 1 microsecond is an SI unit (International System) of time equal to 1000000th of a second.

[215] It takes on average 300 milliseconds to blink (Wilkinson et al., 2013).

B) High-End Capability Computing Infrastructure

As prior noted, co-location infrastructure may be useful for trading on order book news as it travels fast within the exchange. However, information on the macro-economy or firm fundamentals, which in general are larger and more complex than order book information, travels slower and is exponentially more complex to interpret than order book information. Therefore, it requires more time, other algorithms and infrastructure. This is where high-end capability computing infrastructure comes to the fore.

As the amount, complexity, and rate of generation of market data increase exponentially, high-end capability computing infrastructure is becoming an increasingly relevant tool for algorithmic traders.[216]

Indeed, those strategies that do not necessitate co-location facilities are almost certainly conducted via high-end capability computing infrastructure.

High-end capability computing infrastructure generally prioritises capability computing over capacity computing. Typically capability computing serves as a lever to pry new insight from a mass of big data or complicated mathematical formula. As such, capability computing is not measured in terms of floating point operations per second or number of processors. Rather, from a researcher's perspective, high-end capability computing is measured by its inherent complex, analytical, logical and deductive reasoning capabilities.

[216] Indeed, large data is a fact of today's world and data-intensive processing is fast becoming a necessity, not merely a luxury or curiosity.

More importantly, high-end capability computing infrastructure functions to exploit information beyond the traditional order book data. This includes news, 'pre-news'[217] and other forms of 'textual'[218], as well as numerical information. (Leinweber, 2009). Indeed, major news providers have started offering algorithmic traders access to electronically processable news feeds - providing high-capability computer infrastructure users with valuable numerical and textual information.

According to Johnson (2010), high-end capability computing infrastructure offers the potential to improve predictions for key market variables, "This is because they can incorporate a much wider range of factors in their forecast models. They may also be able to cope with today's more (complex)[219] marketplaces" (p.489).

[217] Pre-news information comes from financial news articles, expert recommendations and even social media. By leveraging high-end capability computing infrastructure algorithmic traders can apply advanced forecasting techniques – such time series analytics, machine learning, neural networks, support vector machines tools, as well as text mining – in order to extract actionable information from these pre-news feeds. Interestingly, evidence indicates that many algorithmic traders – those that leverage high-end capability computing infrastructure – are able to verify and exploit previously unheard forms of information (such as investor sentiment) from sources like twitter – a social media website (Groß-Klußmann & Hautsch 2011).

[218] Nowadays, a huge amount of valuable information related to the financial market is in textual format and available on the web. However, there is a limit to the amount of information a human trader can analyze. By leveraging high-end capability computing infrastructure algorithmic trades can apply advanced forecasting techniques – such time series analytics, machine learning, neural networks, support vector machines tools, as well as text mining – in order to extract actionable information from textual news feeds (Shah, 2007).

[219] Market integration/expansion has undeniably facilitated an increased level of market complexity. However, high end capability computing has created new possibilities that no human trader could ever offer, such as assimilating and integrating vast quantities of data and making multiple accurate trading decisions across multiple venues. Johnson (2010) states that "algorithmic traders are able to cope with today's more complex marketplaces, where trading is fragmented between multiple venues."(p.489) Moreover, their ability process vast, complicated and imprecise data means that they are able to detect patterns and identify trends that are too intricate to be noticed by humans alone (Gamzo 2014). Accordingly, their access to multiple markets allow them to search through trillions of observations and identify elaborate patterns in market activity – this then allows them to implement profitable trading strategies without any direct human intervention.

Overall, those strategies that are performed on high-capability computer infrastructure exploit their superior ability to interpret public information in an attempt to make forecasts that are superior to the forecasts of other traders. In other words, these traders filter public information through an advanced platform (high-capability computer infrastructure), in order to detect private patterns from public information – patterns that signal a firms future performance.

More precisely, high-end capability computing infrastructure is a crucial aspect of the operation of many of the algorithmic trading strategies, therefore, it seems natural to assume that several algorithmic trading strategies – those not reliant on colocation infrastructure- are simply predicated on an ability to make forecasts that are superior – in terms of accuracy – to those of other traders. Indeed, Qin (2012) advocates that high-end capability computer infrastructure has fundamentally influenced the accuracy of forecasting. However, it is important to note that this type of forecasting – based on high end capability computing – is especially intensive, thus, it takes relatively longer to accomplish than other less effortful techniques.[220]

Following key academic and regulatory literature on the subject of algorithmic trading, we indicate below which of the noted strategies rely on co-location infrastructure and which rely on high-end capability infrastructure.

Functional Organization of Algorithmic Trading Strategies: A Comparative Summary

1. **Strategies that rely on co-location infrastructure**

- *Spread Capturing Algorithms*
- *Rebate Trading Algorithms*

[220] Such as those strategies based on co-location infrastructure.

- *Time Weighted Average Price (TWAP) Algorithms*
- *Volume Weighted Average Price (VWAP) Algorithms*
- *Implementation Shortfall Algorithms*
- *Adaptive Execution Algorithms*
- *Liquidity Detection Algorithms*

2. Strategies that rely on high-end capability computer infrastructure

- *Data\Text Mining Algorithms*
- *Neural Network Algorithms*
- *Support Vector Machine Algorithms*

We close this appendix with Table I - a comparative summary of the dual process construct.

Table I: A Comparative Summary

COMPONENT	SYSTEM 1 ALGORITHMIC TRADING	SYSTEM 2 ALGORITHMIC TRADING
Theoretical Relationship	System 1 Cognitive Processing	System 2 Cognitive Processing
Behavioral	Reflexive/Fast (Speed Advantage)	Reflective/Slower (Accuracy Advantage)
Functional/ Infrastructural	Co-location	High-End Capability Computing
Practical	Order Book Information	Fundamental Information
Organizational	• Spread Capturing Algorithms • Rebate Trading Algorithms • Time Weighted Average Price (TWAP) Algorithms • Volume Weighted Average Price (VWAP) Algorithms • Implementation Shortfall Algorithms • Adaptive Execution Algorithms • Liquidity Detection Algorithms	• Data Mining Algorithms • Text Mining Algorithms • Neural Network Algorithms • Support Vector Machine Algorithms

APPENDIX II

Private Information: The Dimensions of Private Information Accuracy

As discussed in Chapter 4, two related mechanisms are referenced in the literature to explain the accuracy phenomenon. The first position holds that, although agents observe the same public information, they *interpret* this public information differently. This idea was first introduced over a century ago by Bachelier (1900) and later developed by Holthausen and Verrechia (1990). The second proposition argues that agent's use public announcement to *infer* new private information from the public information itself (e.g., Kim & Verrecchia, 1994, 1997). These mechanisms are defined as the interpretation mechanism (a) and the inference mechanism (b) of private information accuracy respectively.

a) The Interpretation Mechanism of Private Information Accuracy

Holthausen and Verrecchia (1990)

It is well established in the literature that trading on financial markets is strongly influenced by public firm-specific, macroeconomic and other related information flows. Markets react sensitively to this so called 'news' which is announced on a recurrent and intermittent basis. However evidence in the accounting literature also seems to suggest that these information flows (e.g., accounting earnings) are interpreted heterogeneously by investors. What is relevant is not that each investor sees the same earnings per share figure, for example, but that investors reach varying conclusions about the revised value of the firm after information releases / earnings announcements (Morse et al., 1987). In keeping with this focus, Holthausen and Verrechia (1990) present a partially revealing rational expectations model of competitive trading in order to identify

the effects such information releases on investor decision making. Put succinctly, they focus on the effects heterogeneous interpretations of public information releases have on market prices and volumes.

To facilitate discussion, Holthausen and Verrechia (1990) introduce a rational model of trade in which many agents exchange a single risky asset and a single riskless asset over one period. At the start of the period, agents have homogeneous expectations with respect to the value of the risky asset. Exchange is motivated by agents receiving information about the liquidating value of the risky asset. Information is modelled specifically to assess the extent to which heterogeneous interpretations of a public information release result in price and volume reactions. We elaborate briefly on the particular form of information addressed in the model before discussing its key insights.

Holthausen and Verrechia's (1990) theoretical market is comprised of a continuum of informed and optimally motivated agents, each indexed by $\alpha_{\epsilon}[0,1]$. At the beginning of the period, market participants are endowed with risky and riskless assets. The aggregate supply of risky assets is unknown, and uncertainty about the effect of its behaviour on prices represents the noise in the economy. Within the single period two events occur. First, each agent receives a signal about the liquidating value of the risky asset. Second, agents exchange assets on the basis of their information. This exchange of assets implies an equilibrium price p for the risky asset. When the period concludes, the risky and riskless assets are liquidated, and agents and traders consume their holdings of each. The risky and riskless assets pay off in the economy's single consumption good. The risky asset has a normal distribution and is represented by the random variable v.The riskless asset is a numeraire commodity, one unit of which returns one unit of the single consumption good

when assets are liquidated, independent of when the riskless asset is acquired (i.e., discounting is ignored).

In order to capture the universal elements of the variety of rational expectations trading models, Holthausen and Verrecchia's (1990) rational expectations model was designed to be as basic and generic as possible. Given that rational models of trade are well established in the literature, we note only the unique elements of their model that differ from traditional modelling techniques. Specifically, the nature and distribution of information as well as the heterogeneity underlying agent specific interpretations of these informational variables is what sets this model apart from its contemporary counterparts. Thus, a precise description of informational variables specific to the model will suffice for the purpose of this section.

Information Sets in Holthausen and Verrecchia's (1990) Setup

The notation v represents the liquidating dividend of the risky asset and, thus, represents the risky asset's true, economic cash flow. The parameter σ_v^2 can be thought of as the variance of that cash flow. At the beginning of the period, agents believe that v is normally distributed. Moreover, during the period each agent receives an information signal and *interprets* what the information signal implies about the liquidating value v, such that each investor's interpretation of the signal is given by:

$$A = v + e + \delta$$

Where, δ is an idiosyncratic noise term that has a normal distribution with mean zero and variance s, and e is a common noise term that is normally distributed with mean zero and variance n. The random variables v, e and δ are each independent from one another. Furthermore, the idiosyncratic noise terms δ are independent across agents : that is $E[\delta_\alpha \cdot \delta_k] = 0$ for all $\alpha = k$.

By modelling information with both common and idiosyncratic noise terms,[221] Holthausen and Verrecchia (1990) facilitate a discussion regarding the formation process of private information. Moreover, the model effectively demonstrates that agents can be differentially informed following a commonly-observed signal. Indeed, the literature suggests that while investors all observe the same reported earnings, their interpretations of those earnings for the value of the firm need not be homogeneous. In such an environment, each agent observes the same public signal (e.g., an earnings announcement) but each agent's interpretation of what the signal implies about the value of the liquidating dividend varies because of the agent-specific noise term (idiosyncratic noise term). With respect to an earnings announcement, this is akin to all agents observing the same reported earnings per share figure, and then determining the implication of the reported earnings for the value of the firm. What is relevant is not the earnings number per se, but the implications of the earnings release for the value of the firm.

Another dimension of information accuracy provided in the literature explicates the role of an agent's ability to infer *new* private information from public information.

This inference proposition implies that some agents utilise information gathered in anticipation of public announcement in order to infer new private (and possibly more accurate) information from the announcement itself.

[221] Indjejikian (1988), Lundholm (1988), and Pfleiderer (1984), among others, also model information with both a common and idiosyncratic component of noise.

b) The Inference Mechanism of Private Information Accuracy

Kim and Verrecchia (1994)

Recall, Holthausen and Verrecchia (1990) model private information by assuming that an earnings announcement provides each agent with a *single* signal $A = v + e + \delta$. An alternative way to model private information with regards to information accuracy is provided below. This alternative modelling methodology explicates the role of an agent's ability to infer *new* private information from public information itself. In other words the inference mechanism of private information accuracy suggests that agents are able to utilise information gathered in anticipation of public announcement in order to infer new private, and perhaps more accurate information from the information release. Although closely related to the differential interpretation principal, the inference principal suggests that the process of gathering private information is a cumulative process. Kim and Verrecchia's (1994) rational expectations model is introduced below.

Their paper suggests that information releases provide information that allows certain traders to make judgments about a firm's performance that are superior to the judgments of other traders. Their characterisation of information releases is sufficiently broad to include earnings announcements, management and analysts' forecasts, 10-K filings, and other summaries of detailed financial accounting statistics. For convenience, however, throughout the remainder of the paper, they imagine these disclosures to be specifically earnings announcements.

In their model some market participants process earnings announcements into private, and possibly diverse, information about a firm's performance at some cost (e.g., time and effort). This private information can be thought of as informed judgments or opinions. Market participants who provide informed judgments are those traders willing to bear the cost for engaging in this activity. The

ability of information processors to produce superior assessments of a firm's performance on the basis of an earnings announcement provides them with a comparative information advantage relative to other market participants. Kim and Verrecchia's (1994) model set-up is as follows:

There are four types of risk-neutral agents: a market maker, potential information processors, $\boldsymbol{L}$ nondiscretionary liquidity traders, and $\boldsymbol{T} \cdot \boldsymbol{M}$ discretionary liquidity traders. There are T periods, where $s = 1,2, \dots, T$. There is one risky asset, which they call the firm, and riskless bonds. One bond pays off one unit of consumption good in period $\boldsymbol{T}$**.** The firm generates cash flow of v_s in period s**.** Each v_s is an independent, normally distributed random variable with mean zero and variance α_s.

The liquidating value of the firm, denoted by v, is a random variable defined by

$$v = \sum_{S=1}^{T} v_S$$

At the end of period s, the realization of v_s, becomes common knowledge. Institutionally, this arrangement can be thought of as one in which the firm earns revenue by completing a series of T independent contracts.[222] As each contract is completed, the cash flow generated by that contract becomes known (by the firm reporting this information or otherwise). In this context, a period is not a fixed length of time or a cycle over which a firm must report, but rather the length of time it takes to complete a contract. Before the tth contract is completed, at time τ, say, the firm publicly

[222] E.g., the firm builds customized homes under contract, one at a time.

discloses a signal, A, that contains information about the firm's (anticipated) cash flow for period t. The signal is of the form of

$$A = v_t + \delta$$

Where δ is a normally distributed variable with mean zero and variance d. The variance of v_t is α_t. $\alpha \equiv \alpha_t$ is defined solely for notational convenience. Institutionally, A can be thought of as an earnings announcement that imperfectly communicates, or forecasts, a cash flow whose realization ultimately becomes known. In anticipation of A at time τ, those with the ability to process announcements decide whether or not to do so at a fixed cost C and become information processors. The number of information processors at date s is denoted by N_s and is determined endogeneously. Simultaneous with the dissemination of A, an information processor i observes (at a cost C)

$$K_i = \delta + e_{i,}$$

Where e_i is normally distributed with mean zero and variance e for all i. Note that K_i alone is not an informative signal about the firm's liquidating value v, since both δ and $e_{i,}$ are independent of v. Combined with the announced signal $A = v_t + \delta$, however, K_i generates a signal $A - K_i = v_t - e_{i,}$ and this provides information about the firm's performance. Institutionally, K_i can be thought of as the information a trader gleans about the random error in financial reports by studying the firm. For example, in the case of an earnings announcement characterized by $= v_t + \delta$, the random error δ represents the discrepancy between cash flow in period t (v_t) and the forecast of that cash flow implicit in current accounting profits $(v_t + \delta)$. This discrepancy arises, perhaps,

from the failure of accounting profits to recognize unrealized gains, use consistent accounting procedures, or avoid questionable levels of capitalization. By studying data issued by the firm (at some cost C), traders who process public disclosures are better prepared to translate current accounting profit figures into superior assessments of the firm's cash flow realization. For convenience we assume that the market maker observes only A, consistent with the notion that specialists do no fundamental analysis.

Overall, Kim and Verrecchia's (1994) rational expectations model shows that earnings announcements stimulate informed judgments. These informed judgments, in turn, create or exacerbate information asymmetries between traders and market makers. As a consequence the market becomes less liquid when there is more disclosure.

Discussion

Regarding the dimensions of private information accuracy, two notions currently exist to explain the mechanisms behind how some traders come to be superiorly informed. The first is to assume that agents are capable of interpreting public information differently; and as a consequence some have more accurate information than others. Another explanation for this information superiority - in terms of accuracy - explicates the role of an agent's ability to infer *new* private information from public information. Unfortunately rational models of trade generally restrict themselves to the analysis one type of information variable in isolation. However, one exception is Kim and Verrecchia (1997), who are able to combine both variables in a unified manner. In fact, Kim and Verrecchia (1997) produce a rational trading model that incorporates both, the differential interpretation proposition, and the new inference proposition - more precisely defined by Kim and Verrecchia (1997) as pre-announcement information and event-period information respectively.

In their paper pre-announcement private information refers to the information that investors actively gather prior to a news release. Conversely, event-period information denotes the new information arising from the interaction between information contained in the public announcement and private information gathered prior to the announcement, which becomes useful only in conjunction with the announcement itself.

Kim and Verrecchias (1997) model is defined broadly below. However, the focus here is on the informational aspects of their model. Indeed, the methodology used by Kim and Verrecchia (1997), in order to model information has prescriptive relevance to our own model – however, our model will not have a rational expectations equilibrium as it is an extension of Kyle's (1985) model (Kim and Verrecchia's overall findings are inconsequential here; they are applicable only to rational models of trade) . Therefore, the discussion to follow is restricted to the manner in which private information evolves and the way it is structured in their model in order to highlight its key intuition; insights that may lend themselves well to our own unique model.

Kim and Verrecchia (1997)

Kim and Verrecchia (1997) begin by taking the rational expectations trading model suggested in Kim and Verrecchia (1991), which is based on the existence of pre-announcement private information, and adapt it to include event-period private information of the type suggested in Kim and Verrecchia (1994). There are three points in time, time 1, 2 and 3, and two assets in the economy, a risky asset (firm) and a riskless bond. One unit of riskless bond pays off one unit of consumption good at time 3 when consumption occurs. One unit of risky asset pays off v units of consumption good at time 3. The random variable v is assumed to be normally distributed with mean $\bar{v}$ and precision (i.e., the reciprocal of variance) h. There is a countably infinite number of

informed investors in the economy with constant but differing risk aversion and also a countably infinite number of liquidity traders. That is, investor $i's$ utility function can be written as a negative exponential utility function, $V_i(W_i) = -\exp(-\frac{W_i}{r_i})$), where W_i is his wealth at time 3 and r_i ,is his risk tolerance (r_i is allowed to differ across investors).

In the model, time 1 characterizes the pre-announcement period. Four events occur during the pre-announcement period. First, investor $i, i = 1,2,3, ...,$ is endowed with E_i riskless bonds and zero risky asset. Second, investor i observes a private assessment of firm value, $S_{1i} = v + e_{1i}$, where e_{1i} is independently and normally distributed with mean zero and precision Z_{1i}. The Z_{1i}'s may differ across investors. Third, investor i obtains and observes private information about the error in a forthcoming public announcement, which we assume to be an earnings announcement. This information is represented by $K_i = \eta - e_{2i}$, where the forthcoming earnings announcement is $A = v + \eta$. As discussed below, $K_i = \eta - e_{2i}$ is used only at time 2 (in the event-period) in conjunction with earnings, $A = v + \eta$. Thus, investors actions and equilibrium are unaffected by whether K_i is assumed to be observed at times 1 or 2. Note that η and the e_{2i}'s are all independently and normally distributed with mean zero and precisions n and s_{2i}'s, respectively. The s_{2i}'s may also differ across investors. Fourth, the market opens and investors and liquidity traders buy and sell securities at competitive market prices. The aggregate gross demand for the risky asset by liquidity traders at time 1, denoted by x_1, is a random variable normally distributed with mean zero and precision t_1.

Time 2 characterizes the event-period. During the event-period there is a public earnings announcement. As mentioned above, the earnings announcement communicates firm value with noise, that is $A = v + \eta$. Here, the market reopens and investors and liquidity traders exchange securities a second time. The aggregate gross demand for the risky asset by liquidity traders at time

2, denoted by x_2, is a random variable that is independently and normally distributed with mean zero and precision t_2. Finally, at time 3 the risky payoff v is realized and investors consume their wealth.

Their model incorporates pre-announcement and event-period private information as follows. The S_i's provide private information about firm value at time 1, in the pre-announcement period. This is information investors can use to revise their portfolios in the pre-announcement period in anticipation of a forthcoming public announcement in period 2.

With regard to event-period private information, the K_i's alone are not informative about the firms liquidating value v. Thus, the K_i's cannot be used at time 1, in the pre-announcement period. However, once earnings are announced and A is known, the K_i's generate private information about firm value in the form of $S_{2i} \equiv A - K_i = v + e_{2i}$. This information, in turn, is used by investor i to assess firm value. Institutionally, K_i can be thought of as the information an investor gleans by studying the error in a firms financial reports, where the error arises from the application of random, liberal, or conservative accrual-based accounting practices and estimates.

Notably, earning announcements do not simply relay numerical data, they contain valuable text as well. In fact, the text contained in these financial statements is crucial when it comes to interpreting the data. Sometimes footnotes will even change the meaning of the numbers. The discretionary nature of income recognition inherent in the U.S. generally accepted accounting practice (GAAP) often results in a degree of management manipulation – where the text included in the footnotes of financial statements is frequently the only indication of these activities (White, 2011). Two of the most noticeable trends follow under the guise of either income smoothing or big bath accounting. Empirical evidence indicates that management can and do engage in such behavior (Bartov 1993; Moses, 1987; Ronen & Sadan, 1981). With income smoothing, many firms reduce

earnings in 'good years' and inflate earnings in 'bad years' in order to present stable earnings.[223] With big bath accounting, the hypothesis suggests that, unlike income smoothing, management will report additional losses in bad years in the hope that by taking on all available losses at one time, they will clear the decks once and for all. Crucially, this activity implies that future reported profits will rise.

Thus, in the context of Kim and Verrecchia's (1997) model, when earnings are announced the K_i's can then be used to partially correct for the error in an earnings report.

Insight

Intuitively, advanced agents would trade in the wake of an earnings announcement not just because of the information contained in the announcement itself, but also because their private event-period information leads them to interpret the reported amounts differently than others who lack this information. In fact, event-period private information is often defined as "uniquely privately inferred information about future earnings." (Barron, Harris & Stanford, 2005. p.404)

According to Kim and Verrecchia (1997):

"All anticipated events or announcements motivate pre-announcement private information gathering. In addition, event-period private information is used in all announcements to provide a context or interpretation to the disclosure. Consequently, event-period information also seems a pervasive feature of disclosure" (p. 396).

[223] Concerning the earning reduction aspect of income smoothing, some firms will defer gains and recognize losses in these so called good years. While in an attempt to inflate earnings in bad years, some firms attempt to recognize gains and defer losses.

Essentially, a public release of information triggers agents with diverse processing capabilities, the ability to generate new idiosyncratic and perhaps more accurate information from the public announcement itself.

APPENDIX III

Chapter 4 closed with a survey of Kyle's (1985) Bayesian Nash equilibrium auction model. Regarding this model, the following caveat should be noted here: In Kyle's original (1985) formulation, the informed trader has private information about the risky asset, v, in the form of a signal $s = v + e$. However, in our review of the Kyle model, it was assumed that e the error in the informed traders signal was negligible (i.e., $e = 0$). The assumption that $e = 0$ has become somewhat of a conventional assumption in the literature (Foster & Viswanathan, 1996; Holden & Subrahmanyam, 1992; Vayanos, 1999). We follow the literature in this regard. This common approach is based on the following rationale: the structure of the informed trader's private information (e.g., whether or not the informed trader observes *v* perfectly or with noise) is not crucial. As articulated by Rochet and Vila (1994): "Given that all traders are assumed to be risk-neutral it is only needed that the information structures be nested i.e., that the informed trader knows more than the market" (p.132).

That being said, the more general case of Kyle's (1985) model i.e., where $e > 0$ is presented here. This exercise, although non fundamental, is attempted solely to enhance the comprehensiveness of this thesis.

Imperfect signal

Consider an asset with payoff $v \sim N(p_0, \sigma_v^2)$. The quantity traded by uninformed traders is denoted by $u \sim N(0, \sigma_u^2)$. Assume that the informed trader observes a signal $s = v + e$, where $e \sim N(0, \sigma_e^2)$. Conditioning on s the informed trader maximizes his expected profit by choosing x. Assume that v, u and e are independent of each other. There is a competitive risk-neutral market maker, who sets the asset price as $p = E[v|y]$ based on the batch order $y = x + u$.

Proposition C1. *There exists an equilibrium* (X, P)*, in which the informed trader's trading strategy* X *and the market maker's pricing rule* P *are linear functions of* s *and* y *respectively:*

$$X(s) = \beta(s - p_0), \tag{C1}$$

$$P(y) = p_0 + \lambda y, \tag{C2}$$

where

$$\beta = \frac{\sigma_u^2}{\sqrt{\sigma_v^2 + \sigma_e^2}} \tag{C3}$$

and:

$$\lambda = \frac{\sigma_v^2}{2\sigma_u\sqrt{\sigma_v^2 + \sigma_e^2}} \tag{C4}$$

Proof: Let $\pi = [v - p(y)]\, x$. Following the conjectured linear (C1 and C2) and conditioning on the signal s, the informed trader will choose a market order size x that will maximize his expected profit

$$\begin{aligned} E(\pi|s) &= E[(v - p_0 - \lambda y)\, x|s] \\ &= x\, E(v - p_0|s) - x\, \lambda\, E(x + u|s) \\ &= x\, E(v - p_0|s) - \lambda\, x^2, \end{aligned} \tag{C5}$$

where the projection theorem implies

$$E(v - p_0|s) = \frac{cov(v - p_0, s)}{var(s)}(s - Es)$$

$$= \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2}(s - p_0)$$

$$= \varphi(s - p_0), \qquad \text{(C6)}$$

with

$$\varphi = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2} \qquad \text{(C7)}$$

Notice that $\varphi = \sigma_v^2/(\sigma_v^2 + \sigma_e^2) < 1$ is the weight attached to the signal in forming expectations.

Maximizing $E(\pi|s)$ with respect to x leads to $-2\lambda x + E(v - p_0|s) = 0$, hence

$$x = \beta(v - p_0), \qquad \text{with} \qquad \beta = \frac{\varphi}{2\lambda} \qquad \text{(C8)}$$

Finally, the market maker sets prices such that the semi-strong efficiency condition (zero profit condition) obtains:

$$p = E[v|y] = p_0 + \lambda y$$

The net flow of orders that the market maker receives is denoted y. The solution is then:

$$\lambda = \frac{cov(v, y)}{var(y)}$$

$$= \frac{cov(v, \beta\,(v - p_0 + e) + u)}{vay(y)}$$

$$= \frac{\beta\sigma_v^2}{\beta^2(\sigma_v^2 + \sigma_e^2) + \sigma_u^2}$$

It is now straightforward to compute the equilibrium strategies that characterise the original model (i.e., where $e > 0$):

$$\beta = \frac{\sigma_u^2}{\sqrt{\sigma_v^2+\sigma_e^2}} = \frac{\sigma_u}{\sigma_v}\sqrt{\varphi} \qquad \text{(C9)}$$

$$\lambda = \frac{\sigma_v^2}{2\sigma_u\sqrt{\sigma_v^2+\sigma_e^2}} = \frac{\sigma_v}{2\sigma_u}\sqrt{\varphi} \qquad \text{(C10)}$$

Note that $1/\sigma_e^2$ indicates the precision of the informed trader's signal. Invariably, as the precision of the signal $1/\sigma_e^2$ increases, φ is closer to 1; β increases and the informed trader trades more aggressively.

It is now fairly easy to compute the degree to which the informed trader's private information is revealed by the equilibrium price, which is defined as the inverse of the conditional variance of the true value *v*, given the price *p*. From Bayes' rule and the projection theorem we know that:

$$var[v|p] = var[v] - \frac{cov[v,p]^2}{var[p]} \qquad \text{(C11)}$$

It immediately follows that:

$$var[v|p] = \sigma_v^2(1 - \tfrac{1}{2}\varphi) \qquad \text{(C12)}$$

Notice that with perfect information, $\varphi = 1$, we are left with $0.5\,\sigma_v^2$, implying that the market maker is able to infer half of the private information initially held by the informed investor.

The elegance of the results obtained above emanates, not only from the fact that they confirm Kyle's original formulation, but that they also validate the equilibrium derived in our own study.

This concludes Appendix III.

APPENDIX IV

Academic Definitions: Algorithmic Trading

Authors	Title	Def. Algorithmic Trading
Brownlees, Cipollini and Gallo (2011)	Intra-daily Volume Modeling and Prediction for Algorithmic Trading	The last few years have witnessed a widespread development of automated order execution systems, typically known in the financial industry as algorithmic (or algo) trading. Such algorithms aim at enhancing order execution by strategically submitting orders: computer-based pattern recognition allows for instantaneous information processing and for subsequent action taken with limited (if any) human judgment and intervention.
Chaboud, Chiquoine and Vega (2009)	Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market	[…] algorithmic trading, where computer algorithms directly manage the trading process at high frequency. In algorithmic trading (AT), computers directly interface with trading platforms, placing orders without immediate human intervention. The computers observe market data and possibly other information at very high frequency, and, based on a built-in algorithm, send back trading instructions, often within milliseconds. A variety of algorithms are used: for example, some look for arbitrage opportunities, including small discrepancies in the exchange rates between three currencies; some seek optimal execution of large orders at the minimum cost; and some seek to implement longer-term trading strategies in search of profits. Among the most recent developments in algorithmic trading, some algorithms now automaticallyread and interpret economic data releases, generating trading orders before economists have begun to read the first line.
Domowitz and Yegerman (2006)	The Cost of Algorithmic Trading: A First Look at Comparative Performance	Like Grossman [2005], we generally define algorithmic trading as the automated, computer-based execution of equity orders via direct market-access channels, usually with the goal of meeting a particular benchmark.
Hendershott, Jones and Menkveld (2009)	Does Algorithmic Trading Improve Liquidity?	Many market participants now employ AT, commonly defined as the use of computer algorithms to automatically make certain trading decisions, submit orders, and manage those orders after submission.

Jarnecic and Snape (2010)	An analysis of trades by high frequency participants on the London Stock Exchange	Algorithmic trading is the use of computer algorithms to execute human generated, pre-designated trading decisions.
Prix, Loistl and Huetl (2007)	Algorithmic Trading Patterns in Xetra Orders	Computerized trading controlled by algorithms.

Regulatory Definitions: Algorithmic Trading

Regulator	Document	Def. Algorithmic Trading
AFM (2010) Authority For the Financial Markets	High frequency trading: The application of advanced trading technology in the European marketplace.	Algorithm trading is the collective term for all strategies whereby orders are given according to a pre-programmed set of rules (algorithms).
ASIC (2010a) Australian Securities & Investment Commission	REPORT 215: Australian equity market structure.	We have characterized it in this report as electronic trading whose parameters are determined by strict adherence to a predetermined set of rules aimed at delivering specific execution outcomes. These parameters may include any one or more of volume, price, instrument, market, type, timing and news.
CESR (2010a) Committee of European Securities Regulators	CALL FOR EVIDENCE Micro-structural issues of the European equity markets.	Algorithmic trading or black-box trading, [is] based on the use of computer programs for entering orders with the computer algorithm deciding on individual parameters of the order such as the timing, price, or quantity of the order.
European Commission (2010)	Public consultation: Review of the Markets in Financial Instruments Directive (MiFID)	Automated trading also known as algorithmic trading can be defined as the use of computer programs to enter trading orders where the computer algorithm decides on aspects of execution of the order such as the timing, quantity and price of the order. This form of trading is used by an increasingly wide range of market users.

APPENDIX V

TABLE V. Kyle, Benchmark and Final Model: A Comparison

Table V (presented on the next page) provides a succinct comparison of the models of this thesis, noting the limiting cases.1

Differentiating Between Informed Traders: Kyle, Benchmark and Final

APPENDIX V: Table V

Kyle (static) Model

Informed Trader (s)	*Private Information*
(1) Informed (human) trader	In Kyle, the informed trader observes the fundamental value v but with some residual noise e_s. This single signal satisfies: $s = v + e_s$

Benchmark Model

Informed Trader (s)	*Private Information*
(1) Informed (System 2 algorithmic) trader	
Kyle's informed trader is replaced with a System 2 algorithmic trader. Note that for explanatory convenience we have used the term 'replace' instead of modify. We could have easily have stated that we modify Kyle's informed trader to align his information structure with the key assumptions made with regards to System 2 algorithmic trading in our taxonomy.	Like Kyle (1985) the informed agent - the System 2 algorithmic trader in our case - receives private information in the form of a noisy signal about *v* that satisfies $s = v + e_s$. Here e_s represents the noise in the signal However, unlike Kyle, the System 2 algorithmic trader can utilize advanced K information, gathered in the previous period, to correct for the noise in the signal; the System 2 algorithmic trader combines $s = v + e_s$ with $K_0 = s - e_s$, to form a perfect forecast of the firm's end of period fundamental value *v*. Most notably, our assumption that the informed trader's private information in equilibrium is perfect (i.e., can forecast without noise) is a stronger assumption than that made in Kyle (1985), where the informed trader has a weaker/less perfect signal of v

Limiting Case

If we nullify advanced K information (rendering it completely useless i.e.,$k = 0$), then the benchmark model collapses to Kyle's original (1985) model. The informed trader is left with the original signal: $s = v + e_s$

Final Model

Informed Trader (s)

Private Information

(1) Informed (System 2 algorithmic) trader

(2)Informed (System 1 algorithmic) trader

The informed System 2 algorithmic trader is analogous with the System 2 algorithmic trader in the benchmark model (with an identical signal structure). What's new in the final model is the addition of the System 1 algorithmic trader.

In the single trading round, the System 1 algorithmic trader obtains a signal I_y about the aggregate incoming order flow $y = x + u$. They rapidly trade *twice* in *one* trading round and do not carry inventory when the trading round ends. The quality of the signal I_y is represented by ρ, the squared correlation between I_y and y, i.e., $\rho \equiv Corr^2(y, I_y) \in (0,1]$. A more informative signal I_y has a higher ρ. If I_y reveals y precisely, $\rho = 1$; if I_y is almost all noise, $\rho \to 0$.

Limiting Case

In the limiting case, when the System 1 algorithmic trader has no information ($\rho = 0$), the equilibrium reduces to the benchmark model