Financial decision-making using data

A dissertation submitted for the Degree of Doctor of Philosophy

Martin Hauptfleisch

Discipline of Finance

University of Technology Sydney

November 3, 2019

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Martin Christiaan Hauptfleisch declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Finance Discipline Group at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: 16/08/2019

Acknowledgements

Firstly, I would like to thank my supervisors Dave Michayluk and Tālis Putniņš for their support and advice throughout by PhD. I would also like to thank my co-authors Brian Lucey, Harry Scheule, and Vinay Patel for their help with research direction. Furthermore, I would like to thank Christina Nikitopoulos and Tom Smith for their timely feedback and insightful comments throughout my candidature.

I thank the Capital Markets Cooperative Research Centre (CMCRC) for their financial assistance and resources. I am thankful to the American Finance Association (AFA) for their financial assistance in attending the AFA meeting in 2016. I am also thankful for the financial assistance afforded by The University for Technology Sydney to travel to the Asian Finance Association Conference 2016 and the KLU Finance Meeting on Gold in 2015. Data provided by a major Australian bank is also appreciated.

I greatly appreciate helpful comments from Asian Finance Association Conference 2016 (Bangkok) and the KLU Finance Meeting on Gold 2015 (Hamburg), as well as internal research meetings at the University of Technology Sydney.

Lastly, I would like to thank my wife for her continuous support and love, as well as that of my parents and the rest of my family. None of my success or achievements would have been possible without their support.

Abstract

The growth in data logging and availability allow for decisions which are increasingly informed through empirical study. The velocity and veracity of data, and the relative efficiency of financial markets enables decisions to be made with unprecedented speed and accuracy. In this thesis we outline three ways in which new data can be used to make decisions across three spheres of finance, namely in financial markets, consumer finance and policy decision making. These three areas show how the depth and breadth that new data and empirical methods can contribute to functional and timely decision making. We also explore information theory and financial markets. We discuss information transmission and processing, asymmetric information, individual behavior, and adverse selection.

In Chapter 2 we use high frequency tick data to ascertain the optimal market structure for information discovery and transmission in world gold markets. We investigate which of the two main centers of gold trading—the London spot market and the New York futures market—plays a more important role in setting the price of gold. Using intraday data during a 17-year period we find that although both markets contribute to price discovery, the New York futures play a larger role on average. This is striking given the volume of gold traded in New York is less than a tenth of the London spot volume and illustrates the importance of market structure on the process of price discovery. We find considerable variation in price discovery shares both intraday and across years. The variation is related to the structure and liquidity of the markets, daylight hours, and macroeconomic announcements that affect the price of gold. We find that a major upgrade in the New York trading platform reduces the relative amount of noise in New York futures prices, reduces the impact of daylight hours on the location of price discovery, but does not greatly increase the speed with which information is reflected in prices.

Chapter 3 uses the information contained in the bank transactions of borrowers to infer their time varying credit risk. We analyze the informativeness of non-mortgage bank transactions on mortgage default for a major retail bank. We find that short-term interest coverage, income changes, home maintenance expenses, and cash withdrawals are strong predictors of future default. Using transaction data improves model prediction by 8.74% from a standard baseline model, allowing for earlier intervention that can assist consumers and lower bank losses. We also develop new model fit measures which allow for the comparison of competing statistical models. These new measures compare already adequate models to clearly distinguish which model provides a

better fit considering the magnitude of the improvement (IMPMAG) and the reduction in improvable area (IMPRED).

The fourth chapter uses data from stock and options markets to ascertain the effect a policy announcement can have on the value of firms, while measuring and controlling for the unresolved uncertainty present when the announcement does not guarantee implementation. We examine the wealth effects of the solar and washing machine (SOLAR), and steel and aluminum (STEEL) tariffs enacted by President Trump in January and March 2018. Using a new approach formalized by Barraclough *et al.* (2013), Ball and Brown (1968), and Han *et al.* (2019), we estimate the effect of these announcement controlling for unresolved uncertainty. We find a negative value effect of approximately 546 and 358 billion dollars in the case of the solar and steel announcement respectively, with a statistically insignificant news effect in line with the market anticipating the announcements. This study outlines implications on the protectionist tariffs.

This dissertation illustrates how modern data and methods can be used to inform decisions regarding the structure of markets, the credit risk of a borrower and the effect of a potential policy on firm value.

Keywords

Price Discovery • Gold • Gold Futures • Market Microstructure • Big Data • Payment Transactions • Credit Risk • Model Performance • Mortgage • Probability of Default • Event study • Value effect • News Effect • Unresolved Uncertainty • Probability of Tariff Outcome

JEL Classification

G13 • G14 • G15 • G18 • G21 • G28 • C19 • H2 • H6

Contents

Financial decision-making using data 0

CERTIFICATE OF ORIGINAL AUTHORSHIP 1

Acknowledgements 2

Abstract 3

Chapter 1: Introduction 7

Chapter 2: Who sets the price of gold? London or New York 12

2.1. Introduction 12

2.2. Market structure 16

2.3. Data 18

2.4. Method 22

2.5. Results 23

2.5.1 Daily results 23

2.5.2 Intraday results 28

2.5.3 Determinants of price discovery 29

2.5.4 Robustness tests 35

2.6. Conclusion 39

APPENDIX - Chapter 2: 40

Chapter 3: Borrower payments and mortgage risk 42

3.1. Motivation 42

3.2. Literature review 44

3.2.1 Borrower payment transactions 44

3.2.2 Mortgage default 45

3.3 Data 47

3.3.1 Data pre-processing of transaction data 51

3.3.2 Final sample descriptive statistics 52

3.4. Methodological framework 55

3.4.1 Two stage regression model 55

3.4.2 Model fit measures 58

3.5. Results 60

3.5.1 Base model 60

3.5.2. Identification of transactional variables 63
3.5.3 Factor transform and relation to default 68
3.5.4 Full model 72
3.5.5 Timeliness of estimates 78
3.5.6 Scoring at origination 79
3.6. Robustness tests 80
3.6.1 Out-of-sample testing 80
3.6.2 Robustness over different time horizons 82
3.6.3 Alternate sampling and factor design 84
3.7. Conclusion 85
Internet appendix for Chapter 3 87
Appendix 3: Pre-processing of transactional data 91
Chapter 4: Announcement effect of Trump Tariffs 98
4.1. Motivation 98
4.2. Literature review 99
Tariffs 99
Impact of China on the US 100
Announcement effect with uncertainty 101
4.2.1. Timeline of the Trump tariffs 101
4.3. Data and summary statistics 103
4.3.1 Sample attributes 104
4.4. Methodological framework 105
4.5. Results 106
4.5.1 Solar and washing machine announcement 106
4.5.2 Steel and aluminum announcement 108
4.5.3 Robustness check: excluding confounding events 112
4.5.4 Alternative methodology 112
4.6. Conclusion 114
4.7. Appendix 115
Chapter 5: Concluding remarks 118
References: 121