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.

Production Note:
Signature: 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

Chapter 1: Introduction

Over the last several decades, the increasing power and prevalence of computers have allowed for automated systematic measurement of data of all types. In the 1950s, Bank of America introduced the first banking computer to automate banking bookkeeping, subsequently becoming the largest bank in the world by the 1970s. This automation was highly effective, allowing Bank of America to become the first bank to offer credit cards linked to a customer's bank account. In the 1960s, the introduction of Quotron to find the last traded price of a stock formed the basis of a broker's ability to quote the last traded price to their customers. Subsequent improvements in data recording, storage and retrieval allowed for related products such as stocks and options markets to leave the domain of professionals and enter the minds of common folk. As it currently stands, any person can see any price at any time, and not only current data but historical data as well.

The rapid decrease in data storage cost has led to the recording and storage of all types data, regardless of its apparent current value. The parallel increase in computing power caused a surge in the amount, speed, and variety of data generated. Individuals able to draw value from large amounts of apparently meaningless data, labelled "data scientists", have become some of the most well paid and sought-after professionals.

Drawing value from large volumes of data requires new tools and methods, three of which are discussed in this paper. Each paper uses a distinctly unique data source and cutting-edge diverse methodologies consistent with the defined problem, answering economically valuable questions. We investigate where new information is incorporated using high frequency trade data of world gold markets. We further use high frequency financial transactions to investigate how individual behavior impacts the mortgage default probability, and finally we use financial market data to measure the expected economic effect of the potential passage of a new tariff. Each of the subsequent chapters investigates one of these ideas.

This thesis further contributes to research in the area of information theory and financial markets. This thesis discusses information transmission and processing, asymmetric information, individual behavior, and adverse selection. Chapter 2 is about information discovery and transmission in world gold markets, Chapter 3 uses information to more accurately forecast mortgage delinquency and Chapter 4 measures the economic impact of new tariff information.

Information is becoming increasingly accessible, creating opportunities as illustrated in this thesis to investigate and understand our society and its development.

Chapter 2 investigates the efficiency and the speed of information incorporation into the price of gold for the London over the counter (LOTC) spot market and the New York Mercantile Exchange Futures Market (COMEX). In 2011, the estimated daily turnover in the international gold market was 4,000 metric tons, equivalent to over $240 billion. This is approximately the same as the daily dollar volume of trade on all the world's stock exchanges combined.[1] LOTC and COMEX account for approximately 78.0% and 7.7% of the total gold turnover, respectively. Despite the enormous size of the international gold market, relatively little is known about how information is incorporated into gold prices. This chapter aims to fill this void.

Gold trade internationally is largely decentralized, with price related information generated in many parts of the world, compared to firms confined to a particular area.[2] Understanding the price formation process and where/how information about the value of gold is impounded into its price is paramount to investors and regulators due to the economic significance of gold. The introduction of new capital requirements for banks have brought attention to the pivotal role that liquid assets play in bank risk management, and to the role that gold can play in diversifying a firm's liquid assets.

Using high frequency intraday trade data from LOTC and COMEX we answer the following questions; where do innovations in the price of gold originate, and how has this changed over time? We aggregate tick data to one second increments, and use cutting edge price discovery measures to compare the two gold markets.

We find that the US futures market contributes more to price discovery compared to the London OTC spot market, despite being ten times smaller. Our second key finding reinforces the importance of market structure on the process of price discovery. Our third key finding is that several factors affect the location of price discovery. Price discovery shares vary substantially at both daily and intraday levels. A platform change within our sample for COMEX allows for a

---

[1] According to the World Federation of Exchanges 2011 Annual Report (available at http://www.world-exchanges.org/files/statistics/pdf/2011_WFE_AR.pdf), the total value of all equities traded in electronic order books (stock exchanges) around the world in 2011 is $63 trillion, which, assuming 220 trading days per year, is a daily turnover of around $287 billion.

[2] The emergence of the gold market is described in O'Callaghan (1991).

clean event study of the impact of technology and processing speed. We find that increases in data speed and availability in an international context is a paramount driver of market efficiency.

Our findings carry important implications for market design. We show that market structure is of greater importance to price discovery than market size and liquidity. Market designers can improve the efficiency of their markets by carefully considering the structures they implement, as this can have greater impact than methods focused simply on increasing turnover and participation.

In Chapter 3 we investigate the effectiveness of using individual banking customer credit and debit transactions to predict a customer's credit default risk. We show that the use of this new data source in assessing post origination default probability improves default prediction by over 150%. We contribute to the current literature by developing new model comparison techniques for binary classification.

Banks collect most borrower information when making a lending decision. Post origination, banks rely on changes in house prices through re-valuations and macro indicators. However, borrower conditions can change substantially for many reasons, including changes in employment, income, lifestyle and risk preference. This paper analyzes to which degree banks may be informed through transactional accounts.

We find short-term interest coverage, income changes, home maintenance expenses, clothing expenses and cash withdrawals all contribute to the prediction of future default. The timeliness and accuracy of these variables allow for separation of risky and safe borrowers across the loans' life, with changes in factors allowing for time varying risk estimates. This is a large improvement from traditional models that depend in large parts on invariant origination variables.

The data is proprietary and is collected from a top four Australian retail bank with over 1,000 branches. Consumers are encouraged to package credit cards and transaction accounts with their mortgages. This feature allows for the link between transactions and mortgages. While this data is available within banks, to our knowledge it has not been used in this capacity.

We make several methodological contributions. Transaction data has the distinct feature that transactions may or may not be observed in a period, while traditional credit risk data is usually complete, with lending, borrower and collateral characteristics obtained at the time of credit granting. We devise a methodology to include these transactions and demonstrate that missing values are also informative to mortgage default. Our methodology includes a spline that controls for missing values for each dependent variable.

We propose novel measures for the marginal improvement of competing models. Namely, we derive and calculate transformations of Pseudo R-Square and the Area Under the Receiving Operating Characteristic (AUROC) curve by investigating the improvement magnitude (IMPMAG) and the reduction in the improvable area (IMPRED).

A more accurate and timely measure of risk can allow for more precise loan pricing, capital allocation and loan loss provisioning in financial institutions. For example, under IFRS 9 loan loss provisioning, banks are required to identify loans with a significant increase in credit risk. Our analysis shows that such increases can be much better identified by timely transactional variables that reflect variation in credit risk. Regulators may benefit from a greater understanding of the link between spending and credit default, allowing for more effective policymaking. Households may benefit as banks can provide more efficient pricing based on their payment details which may lead to greater housing affordability.

Chapter 4 includes a study with repercussions for policy decisions, and specifically deals with two of the tariffs announced by President Trump near the beginning of 2018. The final impact of the trade tariffs imposed by US president Donald Trump is unknown, with Trump supporters claiming a benefit and Trump opposition claiming a detriment to the US economy and its relationship with neighbors. Measuring the impact of the tariff announcements carries unresolved uncertainty concerning its final shape and execution, with traditional event studies not able to resolve the ambiguity in the announcement.

In this chapter we quantify the announcement effect of the solar and washing machine tariff (SOLAR), and the steel and aluminum tariff (STEEL), accounting for unresolved uncertainty while measuring the value and news effects. We accomplish this through a novel options-based methodology developed by Barraclough *et al.* (2013) and modified by Han *et al.* (2019).

We find that both the solar tariff and steel tariff demonstrate a negative and significant value effect, indicating a large reduction in the value in the US economy. The news effect is not significant, consistent with tariffs being part of Trump's election promises. In this case the news of impending tariffs was already priced, but the final value effect of the execution was not known.

We contribute to the literature of the use of tariffs by quantifying the value and wealth effects of the announcement, controlling for the possibility that the announcement does not come into effect. We show the effect of protectionist tariffs on growth can be negative (Osang & Pereira

1996), and show that tariffs can generate large welfare losses in the medium and long term (Akcigit *et al.* 2018).

We also contribute to the literature on announcements with unresolved uncertainty. The model used in this chapter allows for a separation of the different effects present in any announcement. In the case of this paper, using traditional event study methodologies when the final implementation of the tariffs is uncertain can cause an underestimation of the announcement effect. Our methodology captures this uncertainty and assigns a probability to a successful outcome. Our methodology also separates the new information entering the market from its effect on the economic value of firms. We find the announcements released an insignificant amount of news and caused a significant effect on the value of firms.

This paper has implications for future tariff decisions. Trump's aim with the tariffs is to increase the value of the US market by reducing the trade deficit to other countries. Although this may eventually be the case, the increase in the cost of goods created by the tariffs results in a large value drain.

This dissertation proceeds as follows: Chapter 2 examines the price discovery shares between the LOTC and COMEX gold markets. Chapter 3 investigates the additional information contained in mortgage customers non-mortgage transactions. Chapter 4 disentangles the news and value effects of the Trump tariff announcements in the presence of unresolved uncertainty. Chapter 5 concludes.

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

2.1. Introduction

Gold is one of the most traded assets worldwide. In 2011, the estimated daily turnover in the international gold market was 4,000 metric tons, equivalent to over $240 billion. This is approximately the same as the daily dollar volume of trade on all of the world's stock exchanges combined.[3] The turnover in the gold market exceeds turnover in all but four currency pairs.[4] The two major centers for gold trading, the London over-the-counter (LOTC) spot market and the New York Mercantile Exchange Futures Market (COMEX), account for approximately 78.0% and 7.7% of the total gold turnover, respectively. Although gold futures account for a smaller proportion of total turnover, several studies of other markets show that futures play an important role in price discovery (Rosenberg & Traub 2009; Bohl *et al.* 2011; Boyd & Locke 2014; Dolatabadi *et al.* 2015), although not in all settings (e.g., Cabrera *et al.* (2009)). Despite the enormous size of the international gold market, somewhat surprisingly, relatively little is known about how information is incorporated into gold prices. This paper aims to fill this void.

Gold trade internationally is largely decentralized, with physical gold traded in OTC markets and financial securities linked to gold (including futures, options and other derivatives) traded on organized exchanges and trading platforms worldwide.[5] Understanding the price formation process and where/how information about the value of gold is impounded into its price is paramount to investors and regulators due to the economic significance of gold. The introduction of new capital requirements for banks have brought attention to the pivotal role that liquid assets play in bank risk management, and in particular to the role that gold can play in diversifying a firm's liquid assets.

In this paper we answer two main questions: where do innovations in the price of gold originate, and how has this changed over time? The demand for gold arises from a variety of sectors. Gold retains a significant industrial and dental use pattern, is popular as an adornment, is

[3] According to the World Federation of Exchanges 2011 Annual Report (available at http://www.world-exchanges.org/files/statistics/pdf/2011_WFE_AR.pdf), the total value of all equities traded in electronic order books (stock exchanges) around the world in 2011 is $63 trillion, which, assuming 220 trading days per year, is a daily turnover of around $287 billion.

[4] The four currency pairs include USD/EUR, USD/YEN, USD/GBP, and USD/AUD with turnover figures of $1,101 billion, $568 billion, $360 billion, and $249 billion respectively (Report on global foreign exchange market activity in 2010).

[5] The emergence of the gold market is described in O'Callaghan (1991).

held as a quasi-reserve currency by official sectors, and is a popular investment vehicle. Supply also is disaggregated, with new gold coming from mines, recycled gold coming from scrap and reuse, and releases of investment and official sector holdings. As a consequence of this, the influences on the gold price are many, as are the sources of market moving information. Given the geographic dispersion in new information about the value of gold, no trading venue will have a clear locational advantage, which is an important consideration in equity price discovery (Anand *et al.* 2011). Furthermore, the large distances between the different trading locations for gold give rise to relatively high latency in information transmission and limit high-frequency trading (Frino *et al.* 2014).

We use intraday data on gold prices in the LOTC and the US futures market during the period 1997–2014. This allows us to examine variation in price discovery through the course of several years, examine the effects of market structure changes during our sample, and analyze the intraday patterns in price discovery and the process by which specific news announcements are impounded into prices.

A striking result of our analysis is that although the volume of gold traded in the LOTC spot market is approximately ten times higher than that of the US futures market (78.0% market share compared to 7.7%), the futures market tends to lead in incorporating new information about the value of gold. This result highlights the importance of market structure and instrument type. Our results support the notion that the centralization and relative transparency of the futures market contribute to its disproportionately large role in price discovery. It is also likely that the low transaction costs, inbuilt leverage and ability to avoid dealing with the underlying asset, make futures contracts an attractive option for those that trade gold as a financial asset, and such trades contribute disproportionately to price discovery.

Our second key finding reinforces the importance of market structure on the process of price discovery. During our sample period, at the end of 2006, the US futures market changed from an open outcry floor-based system to the fully electronic, nearly 24-hour low latency GLOBEX platform. We find that this change notably decreased the amount of noise in US futures prices relative to the UK spot prices, but did not have a large impact on the speed with which the futures market reflects new information about the price of gold.

Our third key finding is that although the US futures market leads with respect to price discovery overall, several factors affect the extent to which it leads. Our results indicate that price discovery shares vary substantially at both the daily and intraday levels, with the two markets changing their relative importance throughout the day and from day to day. Prior to the US futures market's introduction of the fully electronic GLOBEX platform, the price discovery shares of the two major trading centers are largely dictated by daylight hours within each market locale—the LOTC market plays a more important role during UK daylight hours and the US futures market plays a more important role during US daylight hours. With the upgrade of the futures market to the GLOBEX platform and introduction of incentives for international participants, intraday variation in price discovery declines and the futures market takes a consistent lead throughout the day irrespective of market hours. This finding illustrates how highly accessible electronic platforms can decrease the importance of geographic location and create a more integrated global market.[6]

We hypothesize that some of the variation in daily and intraday price discovery shares is related to specific news events that affect the value of gold. The importance of gold to central banks and governments suggests that macroeconomic announcements and central bank announcements may affect the price of gold and the location of price discovery (Batten *et al.* 2010; Hautsch *et al.* 2011; Elder *et al.* 2012). The information contained in specific announcements may have a tendency to be incorporated into specific markets, generating variation in price discovery shares. Furthermore, there is evidence that there may be information leakage from the UK gold price fix, in that the market reacts to the fix while the fix is being undertaken (see Caminschi and Heaney (2014)). However, to what extent the market reaction represents the particular structure of the fix at the time and to what extent it reflects actual price sensitive information leakage is less clear. Leakage from the fix could influence the location of price discovery. To examine these possibilities, we regress price discovery shares on dummy variables for gold fixing times and various major macroeconomic announcements. The LOTC market is relatively faster at reflecting new information around the AM gold price fixing. Also, our results indicate that US GDP announcements are associated with an increase in the US futures market's share of price discovery, whereas US employment announcements including Non-Farm Payroll are associated with an increase in noise in futures prices. UK announcements in general have no effect on the location of price discovery, although some are associated with an increase in noise in the LOTC.

[6] The efficacy of floor and electronic trading is investigated by Ates and Wang (2005)

Our findings contribute to the literature on how different market structures and securities affect the nature of price discovery. It is generally accepted that futures contracts lead their respective underlying assets in price discovery (Rosenberg & Traub 2009; Bohl *et al.* 2011), yet this relationship has not been confirmed in the gold market despite its immense size and economic impact. Cabrera *et al.* (2009) find that foreign exchange spot quotes consistently lead foreign exchange futures prices. The gold and foreign exchange markets are similar in that the spot market accounts for a substantial share of trading activity, and therefore the findings of Cabrera *et al.* (2009) suggest that ex-ante it is not obvious that gold futures would lead the spot market in price discovery.

Our findings are also related to other studies of gold price discovery, including comparisons between COMEX Futures and Tokyo Commodities Exchange Futures (Xu & Fung 2005; Lin *et al.* 2008) and Indian gold futures (Fuangkasem *et al.* 2014). These studies conclude that COMEX dominates price discovery, which in addition to its considerable volume, is why we choose to compare gold prices from COMEX with those in the UK spot market. The papers above do not examine the LOTC spot market, which accounts for 78.0% of global trade. Lucey *et al.* (2013) compare the LOTC market and COMEX using daily data and find that price leadership shifts between the two markets. Our use of intraday data allows for more accurate measurement of price discovery and allows us to characterize intraday variation in the location of price discovery, which is important in this 24-hour global market.[7] Furthermore, our detailed analysis of the determinants of gold price discovery is novel in this literature. A final and non-trivial contribution of our paper to the literature on gold price discovery is in disaggregating gold price discovery into two distinct components—the relative speed at which information is reflected in prices and the relative noise in prices. The disaggregation is important as we show that several of the factors that influence price discovery have opposite effects on the two components of price discovery. Without the disaggregation, the effects of several of the factors would be obfuscated.

Our study also contributes to the broader literature on gold. Prior research in this area is concerned with the hedging value of gold and whether it can be considered a safe haven (Capie *et al.* 2005; Baur & Lucey 2010; Baur & McDermott 2010), the relationship between gold and other

[7] With daily observations, the lead-lag relations between markets that exist at intra-day horizons appear as contemporaneous correlation.

precious metals (Batten *et al.* 2010), psychological price barriers in gold (Aggarwal & Lucey 2007) and its investment value (Sherman 1982; Hillier *et al.* 2006).

Our comparison of physical gold and its derivative securities contributes to the study of physical and financial assets. Physical gold and futures contracts on gold are by nature very different, even though both share the same underlying asset. The former is used by banks, jewelers and manufacturers of many kinds, whereas the latter is primarily used by hedgers and speculators. Our comparison provides an insight into the price setting power of each group, with the pricing lead of the futures market indicating that hedgers and speculators are more sensitive to new information. Such a conclusion is intuitive as this group's main focus is the price of gold, as opposed to the use of gold. Our result is consistent with the current declining physical and increasing financial demand for gold.

This paper proceeds as follows. Section 2 discusses the structures of the LOTC and COMEX. Sections 3, 4 and 5 describe the data, method, and results, respectively, and Section 6 concludes.

2.2. Market structure

COMEX and the LOTC markets are structurally different. COMEX is a centralized exchange in which all orders are routed through one system. The LOTC market on the other hand is a decentralized over-the-counter market in which a number of dealers each quote bid and ask prices. In this section we explain in greater detail the structures of these two markets, and more specifically, the structure of the securities that are of interest to this paper.

Participants in the LOTC comprise market-making members and ordinary members. Major international banks make up 12 of the 13 market-making members and are required by the London Bullion Market Association (LBMA) to provide two-way quotes during London market hours, and whenever the New York market is closed. Ordinary members comprise companies that are operational in areas that are closely related to the physical gold itself, including trading, broking, shipping and storage, mining, refining, inspection and assaying and research. Trading occurs between members of both types. This membership restriction leads to a market with few highly specialized participants representing clients internationally. In 2011, there were 56 full members in this market (Murray 2011). COMEX participants can trade on the LOTC only through LOTC members and cannot trade using their own account.

The usual minimum transaction size is 2,000 fine troy ounces for gold (LBMA 2014), and typical transactions between market makers are 5,000 ounces. Quotes are in US dollars per fine troy ounce with a minimum tick size of one cent. Fine gold content represents the true quantity of gold in a bar, which may be less than the total bar weight due to impurities.

Transparency in the LOTC market is low. There is no public record of trade volumes or prices, only the quotes are observable. The lack of transparency is the major motivation of the Loco London Liquidity Survey (Murray 2011) which endeavors to show that gold is a "high quality liquid asset". The only figures published on a regular basis by the LOTC market are monthly clearing statistics, based on returns from the six clearing members that form the London Precious Metals Clearing Company. The LOTC does not require its members to report turnover and other related statistics. There are no major structural changes to the LOTC market during our sample period that would have a large impact on price discovery.

COMEX is a futures exchange that trades many commodities. Individuals and firms can trade and membership requirements are less stringent than the LOTC. Requirements for individual membership include good moral character and business integrity (CME Group 2014d). Corporate membership is open to various company types and requires some ownership stake in the equity of the exchange. LOTC participants can trade directly on COMEX using their own account.

Each gold futures contract on COMEX represents 100 troy ounces (CME Group 2014a) and is quoted in US dollars per troy ounce. Delivery takes place on any business day within the delivery month, but not later than the last business day of the delivery month. Gold delivered under this contract needs to meet the 995 minimum fineness requirements. Minimum tick size on these contracts is 10 cents per troy ounce.

Most volume in the COMEX gold futures market is in contracts that are marked Trading At Settlement (TAS). TAS allows traders to commit to a trade without knowing the price at which it will settle. A trader submits an order at any time, with this order matched to a countering order. The trade is finally settled at the settlement price which is determined by the exchange at 13:30 Eastern Time. This method essentially allows for a trade to occur at a price that is determined in the future. On the one hand, this added uncertainty allows uninformed liquidity traders to trade on a more equal level with the informed, as neither party should know what the final settlement price will be. Informed traders can use TAS in order to arbitrage any intraday price deviations. With sufficient informed trade, the price at any time should be indicative of the future settlement price.

TAS orders are only available on five contract months, deemed Active Months (CME Group 2014b). These months are February, April, June, August and December, representing every two months except for a break in October. In any given month, the most actively traded contract is the one closest to expiry.

Transparency in the COMEX futures market is much greater than in the LOTC. Although traders are anonymous, bid and ask depth is available for ten price levels. Partially hidden iceberg orders are also available, with each addition to the visible part of the order being placed at the bottom of the queue in the order book (CME Group 2014e).

Although there have been many structural changes to COMEX, only a small number have changed the way gold futures trade. From the start of our sample period in 1997 to December 3, 2006, gold futures on COMEX traded on the floor during floor hours only and on the NYMEX electronic trading platform outside of floor hours only. From December 3, 2006, these contracts commenced trading on the GLOBEX platform, which has lower latency and more accessibility to international traders. Along with this shift, electronic trading hours were extended to overlap floor hours, allowing concurrent trading on the floor and GLOBEX platform. We treat this change as a structural break and analyze how it affects the location of price discovery. Although there are a few other latency-related changes during our sample period, these do not have a meaningful impact on the way trade occurs in this market.

Overall, the wholesale nature of the LOTC market is very different to the open retail exchange system available through COMEX. With major differences in products, trade sizes, centrality and participants, we expect that the contributions to price discovery of the New York futures and London spot markets come from different sources. These two markets constitute an interesting cross-border environment in which to study the price discovery of gold.

2.3. Data

This study compares futures contracts traded on COMEX and gold spot quotes from LOTC. Since accurate estimation of price discovery shares can only be achieved through the use of very high frequency data, we use trade and quote data sampled at a one-second frequency. Our sample period extends from January 1, 1997 to November 30, 2014. In total this includes 3,872 trading days and 51,702,414 one-second observations.

TABLE 1

Most liquid gold futures contracts

Month of year	Expiry month
January#	February
February	April
March#	April
April	June
May#	June
June	August
July#	August
August	December
September	December
October	December
November#	December
December	February

Note: This table illustrates the most actively traded gold futures contracts on COMEX in each month of the year. Month of year is the calendar month of any respective trading day. Expiry month indicates the most actively traded futures contract expiry month for each calendar month. # Volume shifts to next active contract on the last two business days of this month.

We obtain intraday and trade and quote data for the futures contracts and the LOTC market from *Thomson Reuters Tick History*. Our data include the best (inside) bid and ask quotes in each market, time-stamped to the millisecond. From the best bid and ask quotes, we calculate the midquote (the simple average of the bid and ask quotes at that point in time), which reduces the effects of bid-ask bounce. Using these data, we identify the most actively traded futures contract for any given day, illustrated in Table 1[8]. The contracts are deliverable on any day within their expiry month. Consequently, contracts are not very actively traded within their delivery period. For example, because a February contract is deliverable on any business day in February, it is no longer actively traded during February. Intuitively, contract holders would like to take delivery as early as possible to collect their returns. We find that volume shifts to the next active contract two days before the current contract becomes deliverable. For example, the most traded contract in January is the one closest to expiry (the February contract); however, on the last two days of January, volume shifts to the next active contract (the April contract). On this basis, we create a futures price series that uses the most actively traded contract at every point in time. This price

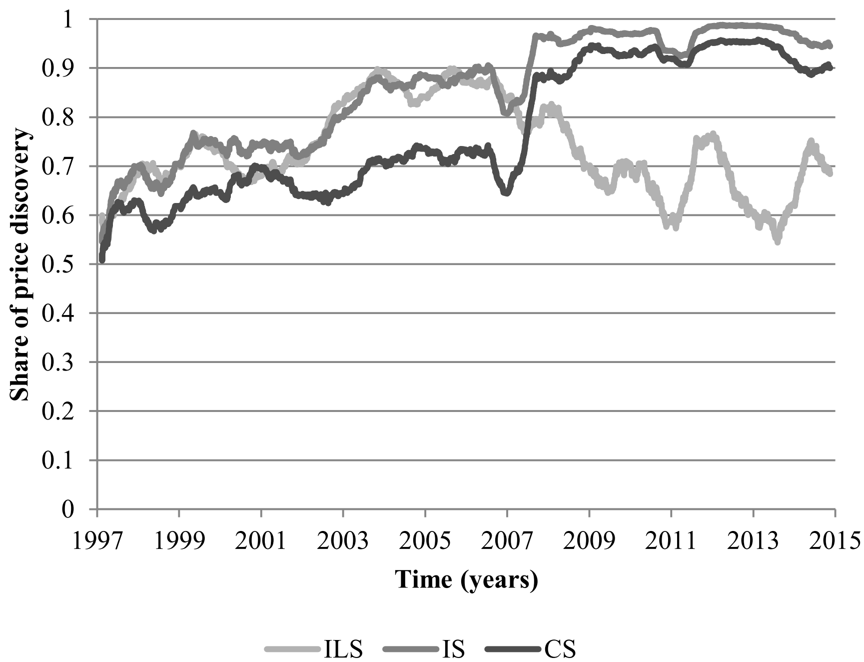
[8] Using futures contracts which are less traded can cause spurious results. New information in secondary futures contracts will lag the most actively traded contracts. Therefore, we only use the most actively traded contract in this thesis.

series is converted into a one-second sampled time series. Similarly, we convert the intraday quotes for the LOTC gold spot market to one-second increments. The two midquote series are merged by date and time, resulting in one time series of two prices.

Figure 1

Futures market price discovery shares through time

Note: This figure plots the futures market (COMEX) price discovery shares through time. Each line is a 180-day moving average of the daily price discovery estimates. The daily estimates of Information Shares (*IS*), Component Shares (*CS*), and Information Leadership Shares (*ILS*) are calculated using intraday data with a sampling frequency of one second.



Because the LOTC gold spot market and the COMEX futures market both trade almost 24 hours per day, there is ample overlap between their trading times. COMEX gold futures trade in an electronic exchange setting from Sunday to Friday, with a break from 17:15 to 18:00 New York time. The LOTC on the other hand allows continuous trading through their inter-office telephone service in an OTC setting, with indicative quotes from members available at all times. Summer and winter time changes in the US change the concurrent trading time of the two markets. New York time shifts between -4h GMT and -5h GMT. Converting the COMEX break period to GMT yields 21:15 to 22:00 in summer and 22:15 to 23:00 in winter. To simplify calculation and ensure both markets trade concurrently for the entire sample, we eliminate all trades and quotes after

20:00 GMT. This also eliminates a short period of time around the opening and closing of COMEX, consistent with the approach taken in other intraday microstructure studies.

We extract data on macroeconomic announcements for both the UK and US from Bloomberg. The data include the announcement content and announcement timestamp for major economic announcements including GDP, central bank target rate, employment figures (including Non-Farm Payroll for US), PPI, and CPI.

The gold spot market is large, with most trades occurring in LOTC market. Due to the nature of OTC markets, trade volumes are not reported making it difficult to measure the size of the gold spot market. Table 2 illustrates the trade in gold in each of the six major gold trading countries. These estimates are sourced from (Lucey *et al.* 2013) (who use Murray (2011) and GFMS Ltd. (2012) data) and reflect trading in all gold-related instruments (including spot and gold derivatives).

TABLE 2

Global gold turnover during 2011

	'000 Ounces	Share of total (%)
United Kingdom	43,775,704	86.75%
United States	4,991,604	9.89%
China	697,002	1.38%
India	494,547	0.98%
Japan	488,502	0.97%
Dubai	12,507	0.02%
Total Volume	50,459,866	

Note: This table reports the estimated trading volume and proportion of volume traded in each of the six major gold trading countries for all gold-related instruments, including spot and derivatives. UK data are from Murray (2011), all other data are from GFMS Ltd. (2012). This table is originally compiled by Lucey *et al.* (2013).

Table 2 illustrates that the UK accounts for approximately 87% of the combined volume of gold trading in the six main gold trading countries, with 90% of this being in the spot market. According to the Loco London Liquidity Survey (Murray 2011), the daily turnover on the London gold spot market alone is in excess of $216 billion, which is comparable in value to US-Australian and US-Canadian dollar foreign exchange settlements (based on 2010 data in Bank for International Settlements (2011)), as well as the daily turnover of all stock exchanges in the world combined. The average daily dollar volume of our selected futures contract over the same period is approximately $22 billion, illustrating the size disparity between our markets.

2.4. Method

Our aim in this paper is to analyze where information enters the gold market, and how this has changed over time. We begin with two measures that are widely used in the price discovery literature, namely the Hasbrouck (1995) Information Share (*IS*) and the Gonzalo and Granger (1995) Component Share (*CS*). Fundamentally, both *IS* and *CS* decompose price innovations into permanent and temporary components. They are estimated using a Vector Error Correction Model (VECM):

$$\boldsymbol{\Delta p_{1,t} = \alpha_1(p_{1,t-1} - p_{2,t-1}) + \sum_{i=1}^{200} \gamma_i \Delta p_{1,t-i} + \sum_{j=1}^{200} \delta_j \Delta p_{2,t-j} + \varepsilon_{1,t}} \tag{1}$$

$$\Delta p_{2,t} = \alpha_2(p_{1,t-1} - p_{2,t-1}) + \sum_{k=1}^{200} \varphi_k \Delta p_{1,t-k} + \sum_{m=1}^{200} \phi_m \Delta p_{2,t-m} + \varepsilon_{2,t} \tag{2}$$

where $\Delta p_{i,t}$ is the change in the log price ($p_{i,t}$) of the asset traded in market i at time t. Appendix A outlines the calculation of *IS* and *CS* from the VECM model above. We estimate the VECM and the price discovery measures separately for each day in the sample period (avoiding the need to deal with periods of market closure) and each hourly interval in the sample.

Recent studies of price discovery measures show that *IS* and *CS* both are sensitive to the relative level of noise between two markets—they measure a combination of leadership in impounding new information (what price discovery metrics aim to measure), and the relative level of noise in the price series (Yan & Zivot 2010; Putniņš 2013). Consequently, *IS* and *CS* tend to overstate the price discovery contribution of the less noisy market. Of the two, *IS* places greater weight on the speed at which a price series impounds new information, compared to the *CS* metric which is a measure of the relative levels of noise. It is likely that the levels of noise in the prices of the two markets examined in this paper are vastly different considering their differences in liquidity, market structure and instrument types. Therefore, it is important to keep in mind the sensitivity of *IS* and *CS* to differences in microstructure noise when interpreting the results.

An important insight of the recent price discovery literature is that a combination of *IS* and *CS* is able to correctly attribute contributions to price discovery without being influenced by differences in noise levels. *IS* and *CS* can be combined in such a way that their dependence on noise cancels out. This measure, known as the Information Leadership Share (*ILS*), developed in Yan and Zivot (2010) and Putniņš (2013), is calculated as follows:

$$ILS_1 = \frac{\left|\frac{IS_1 CS_2}{IS_2 CS_1}\right|}{\left|\frac{IS_1 CS_2}{IS_2 CS_1}\right| + \left|\frac{IS_2 CS_1}{IS_1 CS_2}\right|},\ ILS_2 = \frac{\left|\frac{IS_2 CS_1}{IS_1 CS_2}\right|}{\left|\frac{IS_1 CS_2}{IS_2 CS_1}\right| + \left|\frac{IS_2 CS_1}{IS_1 CS_2}\right|} \quad (3)$$

We estimate all three price discovery metrics, noting that they measure different aspects of price discovery. *ILS* measures the relative speed at which a market reflects new information and therefore is informative about where information first enters the market; higher values of *ILS* for a market indicate the market is more often the first to reflect new information. In contrast, *CS* and *IS* contain incremental information about the relative amount of noise in one market's prices compared to those of the other market; higher *IS* and *CS* for a market indicate its prices are relatively less noisy.

2.5. Results

2.5.1 Daily results

We estimate the price discovery measures for each trading day in our sample period using a one-second sampling frequency. Table 3 reports the annual averages of the daily price discovery shares for the futures market (the spot market price discovery shares are simply one minus the futures market share), and Figure 1 illustrates the trends through time using a 180-day moving average.[9] The *IS* and *CS* measures at the start of the sample are only slightly above 50% (67% and 61%, respectively). *CS* rises steadily until 2006, after which it remains consistently above 90%. *IS* also increases sharply after 2006 and remains very high for the remainder of the sample period.

ILS tells a slightly different story. The futures market in the first year of our sample has an *ILS* of 66%, which rises above 80% for the years 2002 to 2007, after which it falls slightly and remains stable around 66%. Due to its insensitivity to differences in noise, the *ILS* estimates paint the clearest picture of trends in impounding new information. It indicates that in each of the past

[9] The moving average smooths the day-to-day variation in price discovery shares, allowing one to see the differences in the mean levels of price discovery. In Table III, we show that the price discovery shares are statistically different to 0.50, indicating that the day-to-day variance does not impede our ability to identify where price discovery occurs on average.

17 years, the futures market has been a more important source of gold price discovery than the UK spot market and that its contribution has fluctuated since the start of the sample. The fact that *IS* and *CS* tend to increase sharply after 2006 and this increase is not reflected in *ILS* suggests that after 2006 the futures market experienced a substantial decline in the relative amount of noise in its prices compared to the spot market. The Global Financial Crises does not appear to have a discernible impact on the any of the price discovery measures.

TABLE 3

Futures market share of gold price discovery by year

Year	*IS*	*CS*	*ILS*
1997	0.6704***	0.6079***	0.6584***
1998	0.6916***	0.6020***	0.7013***
1999	0.7495***	0.6490***	0.7434***
2000	0.7339***	0.6841***	0.6705***
2001	0.7241***	0.6461***	0.7082***
2002	0.7913***	0.6444***	0.8104***
2003	0.8727***	0.7093***	0.8867***
2004	0.8724***	0.7288***	0.8397***
2005	0.8859***	0.7288***	0.8731***
2006	0.8239***	0.6543***	0.8449***
2007	0.9618***	0.8749***	0.8186***
2008	0.9725***	0.9238***	0.7038***
2009	0.9694***	0.9287***	0.6893***
2010	0.9444***	0.9208***	0.6042***
2011	0.9777***	0.9441***	0.7483***
2012	0.9871***	0.9552***	0.6139***
2013	0.9740***	0.9241***	0.6118***
2014	0.9400***	0.8915***	0.7094***

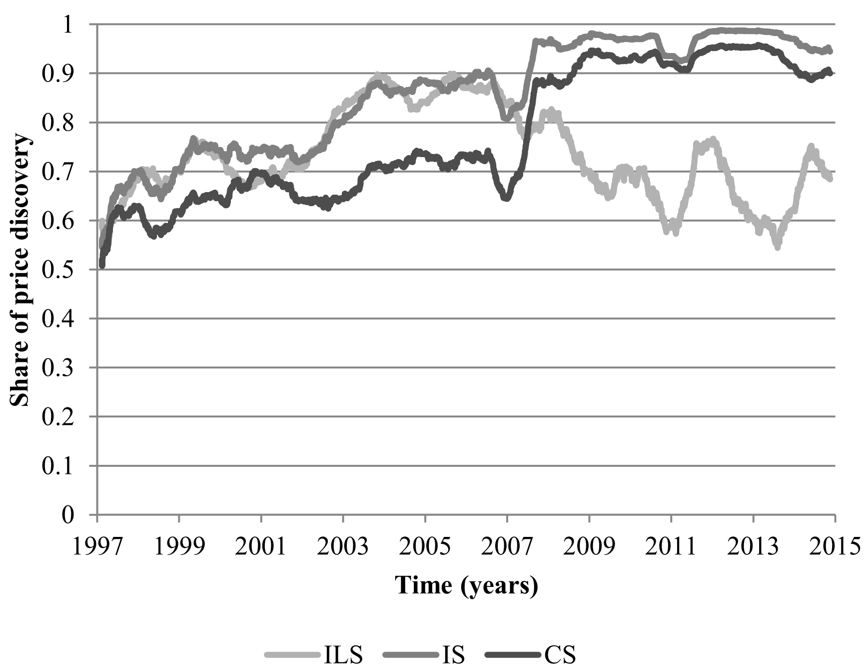
Note: This table reports annual averages of daily gold price discovery shares (estimated form one-second intraday observations) for the New York (COMEX) futures market. The futures market price discovery shares, which are estimated relative to the London spot market, are: Information Shares (*IS*), Component Shares (*CS*), and Information Leadership Shares (*ILS*). *** denotes an estimate is significantly different from 0.50 at the 1% level.

Such a distinct change in the price discovery metrics begs further explanation. The change occurs at the end of 2006 and is statistically significant controlling for other factors (regressions results reported in Section 5.3). The likely cause is a substantial change in the structure of trading at COMEX during the last two months of 2006. At this time COMEX opened electronic trading alongside floor trading. Before the change, the bulk of the trading volume on COMEX was generated by the trading floor in an open outcry system. Outside of the floor hours, COMEX used NYMEX's Access electronic trading platform with floor and electronic hours not overlapping (Morrison 2006). The extended trading hours of the electronic platform were primarily to stay competitive with other exchanges that were providing electronic trading at this time (Goodman 2011). At the same time, COMEX adopted the GLOBEX platform, providing near 24-hour electronic trading, internationally, with low latency. The move to an international electronic exchange was complemented by an international incentive program that allowed traders outside the US to trade at lower costs (CME Group 2014c).

Figure 1

Futures market price discovery shares through time

Note: This figure plots the futures market (COMEX) price discovery shares through time. Each line is a 180-day moving average of the daily price discovery estimates. The daily estimates of Information Shares (*IS*), Component Shares (*CS*), and Information Leadership Shares (*ILS*) are calculated using intraday data with a sampling frequency of one second.



Beginning on December 3, 2006, COMEX also expanded its metals electronic trading to include side-by-side trading of Asian and London metals futures contracts (CME Group Media

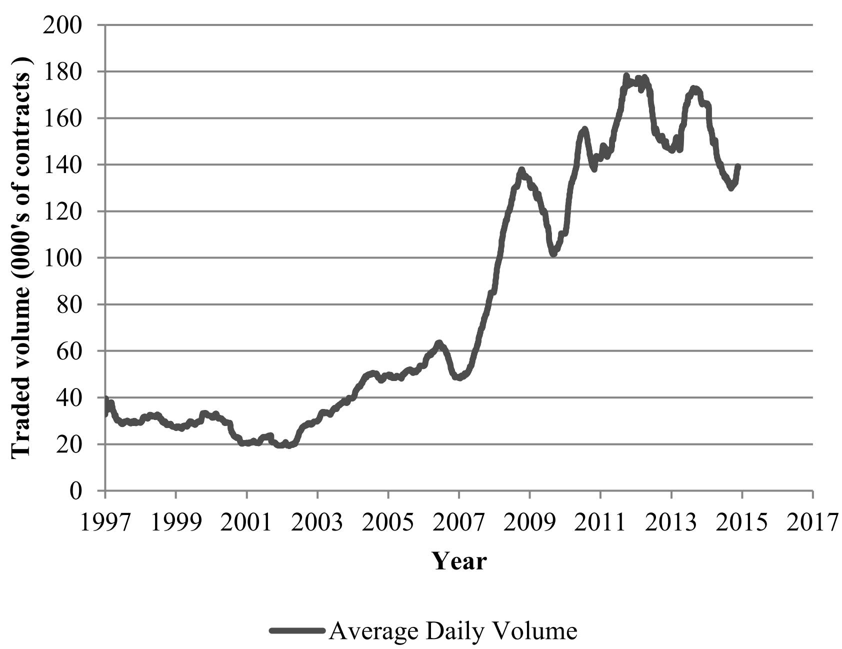
Room 2014). Parallel trading of international futures contracts is likely to have enhanced the attractiveness of COMEX.

The changes in the structure of the COMEX at the end of 2006 flow through to the volume of contracts traded. Figure 2 illustrates the daily average traded volume for the most active contract at every point in time. [10] There is a clear increase in the traded volume on COMEX per contract during the sample period, and in particular, a stark increase after the changes at the end of 2006. Intuitively, greater volume leads to greater liquidity and lower spreads, resulting in less noisy prices and higher values of *IS* and *CS*, as is evident in our results. The Global Financial Crises caused an increase in gold trading volumes as market participants purchase gold as a safe haven.

Figure 2

Futures market volume through time

Note: This figure plots the traded volume of gold futures contracts (the most active contract at every point in time, as specified in Table 1) through time. The line is a 180-day moving average of daily traded volume.



The increase in futures market volume is also likely to affect the average bid-ask spread. Figure 3 illustrates a moving average of the spreads in each of the two markets during our sample period. There is a distinct reduction in the spread of COMEX at the end of 2006. Interestingly, at

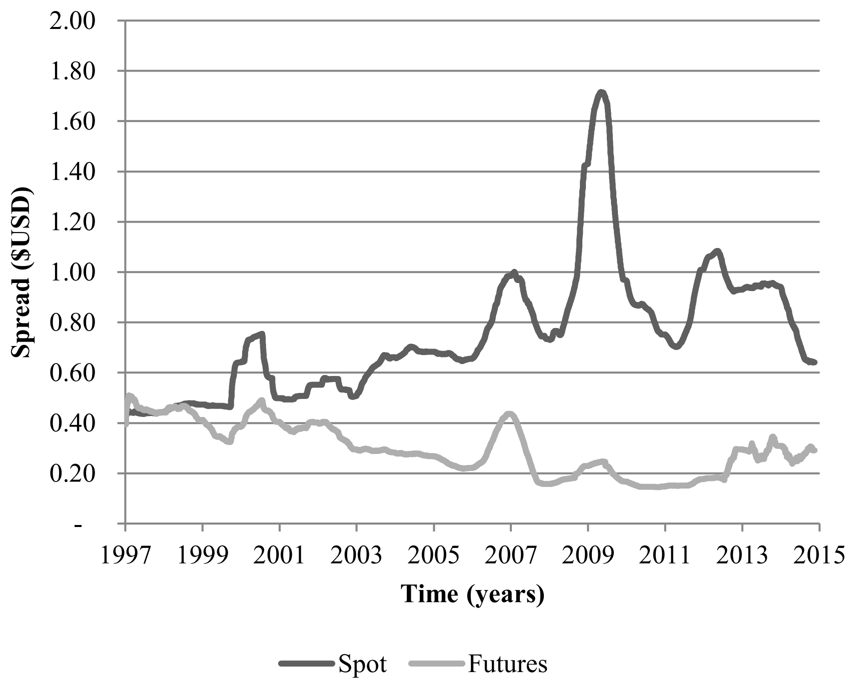
[10] Our data from *Thomson Reuters Tick History* are unable to separate COMEX trades into those executed on the floor versus those executed in their electronic system.

the same time, there is a corresponding increase in the spread of LOTC, suggesting that perhaps some of the volume from LOTC migrated to COMEX after the upgrade of the COMEX platform and implementation of international incentive programs. Unfortunately, due to the absence of a time series of LOTC volumes we are unable to further investigate this conjecture.[11] At many times the futures market spread is constrained by its minimum tick size of ten cents. The LMBE however is not limited by its minimum tick size of one cent. The spike in the spot spread in 2009 is due to the Global Financial Crises causing increased uncertainty, leading to larger spreads in the spot gold market.

Figure 3

Bid-ask spreads through time

Note: This figure plots the 180-day moving average bid-ask spread of the London OTC gold spot quotes and the New York COMEX gold futures quotes. Each quote is for one fine troy ounce and is measured in US Dollars.



The evidence in this section shows that the New York COMEX futures market provides the greater share of gold price discovery throughout our sample period. The change in the market structure of COMEX at the end of 2006 led to significant changes in the global gold market. A more accessible, electronic 24-hour market with low costs and fast execution increased the volume

[11] The absence of data on LOTC volume also means that we are unable to calculate liquidity measures for this market.

and liquidity of COMEX, decreasing the relative amount of noise in COMEX prices relative to those of the LOTC.

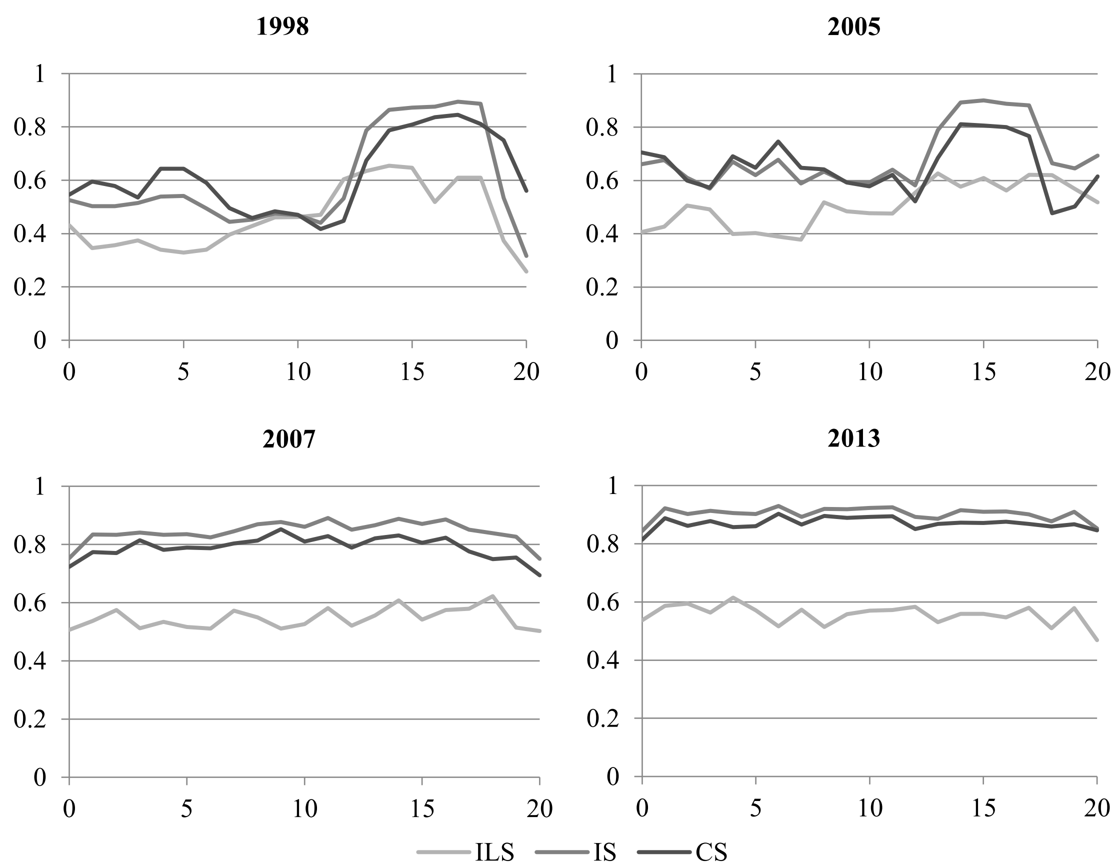
2.5.2 Intraday results

In this section we examine intraday patterns in price discovery by estimating the three price discovery metrics in each hour of each day. Hourly observations allow us to examine the effect of time zone on price discovery. London is either four or five hours ahead of New York, depending on the time of year. Consequently, for up to five hours of the London working day, New York may not yet have started work, and at the end of the day, London would finish work five hours before New York.

Figure 4

Intraday patterns in futures market gold price discovery shares

Note: This figure plots the futures market (COMEX) gold price discovery share intraday averages for four indicative years. Each line represents an average of the hourly price discovery estimates (Information Shares (*IS*), Component Shares (*CS*), and Information Leadership Shares (*ILS*)), which are estimated using intraday data sampled at a one-second frequency. The title of each graph is the year during which the intraday averages are estimated and the horizontal axis measures intraday time in GMT.



The difference in London and New York time zones raises the interesting question of whether the US futures market leads price discovery even when most of the local population is asleep? In other words, is there a tendency for informed investors to be more active in their domestic market during daylight hours? Our results indicate that this is the case in the earlier years of our sample, but not in more recent years. Figure 4 illustrates the average intraday price discovery measures each hour for a few indicative years. The years 1998 and 2005 are representative of the intraday trends for all other years in between—this is the period before the trading platform upgrades in COMEX. Similarly, 2007 and 2013 are representative of the intraday trends from 2007 onwards, after the market structure change.

Figure 4 shows that at the beginning of the sample, there is a distinct pattern in price discovery throughout the day. The opening of floor trading at COMEX around 13:20 GMT (12:20 GMT) is associated with a substantial increase in the price discovery share of the futures market. This increase reverses when floor trading ends at 18:30 GMT (17:30 GMT). In other words, the intraday period during which the COMEX floor is open is associated with substantially more price discovery occurring in the US.

From 2007 onwards, after COMEX introduced the near 24-hour electronic GLOBEX platform, the intraday patterns are substantially different. There are no longer clear intraday patterns in the price discovery shares and instead the price discovery shares remain relatively stable throughout the day. The impact of the floor opening hours is no longer present from 2007. Daylight or working hours no longer affect the location of price discovery.

2.5.3 Determinants of price discovery

In this section we test various determinants of gold price discovery shares, adding multivariate statistical evidence to support the casual observations made in previous sections. We also examine whether different types of macroeconomic news tend to be impounded in one or the other market. To do this, we estimate time-series regressions in which the dependent variable is the futures market's (COMEX) share of gold price discovery, measured each hour. A positive coefficient for an independent variable signifies that variable increases the future market's share of price discovery. Using the high frequency (hourly) estimates of price discovery allows us to investigate intraday effects such as time zones, gold price fixings and news announcements.

Our choice of variables is driven by the prior literature as well as our reasoning about factors that could influence the location of price discovery. The London Gold Fixing is an important determinant of the gold price (Caminschi & Heaney 2014). The AM and PM fixings occur daily at 10:30am and 3:00pm UK time, respectively.[12] We include two dummy variables, AMFIX and PMFIX, that equal one during the hour of the AM and PM London gold price fixing and zero otherwise.[13] We include dummy variables for a set of announcements shown in previous studies to affect the price of gold and other precious metals (Christie–David *et al.* 2000; Cai *et al.* 2001; Batten *et al.* 2010; Elder *et al.* 2012). The announcements include central bank rate announcements and major macroeconomic news such as GDP, employment, PPI and CPI. The dummy variables UKRATE and USRATE equal one during hourly periods that contain announcements from UK and US central banks regarding target interest rates. Similarly, UKEMPLOY and USEMPLOY are dummy variables for national employment announcements in the UK and US. UKGDP, USGDP, UKPPI, USPPI, UKCPI, and USCPI are dummy variables for GDP, PPI and CPI announcements made by the UK and US. With all of the dummy variables, if the announcement occurs on the hour the dummy variable is equal to one for the hourly interval that starts with the announcement, consistent with our approach for the PMFIX variable.

To formally test the effects of the market structure changes in COMEX at the end of 2006, we include a dummy variable, POST, which is equal to one after the introduction of the GLOBEX trading platform. To account for the intraday trends reported in the previous section, we include three dummy variables, ASIA, USA and UK, which equal to one when the Hong Kong Stock Exchange, New York Stock Exchange, and London Stock Exchange are trading, respectively. These dummy variables are proxies for the business hours in the region; we do not expect stock market activity per se to influence gold price discovery. Finally, we include variables that measure the liquidity in each of the markets: FTRSPREAD and SPOTSPREAD are the time-weighted average bid-ask spreads in the COMEX and LOTC markets, respectively.

[12] Our data end before the commencement of the new benchmark price mechanism set by the LBMA as an improvement on the traditional fix. See http://www.lbma.org.uk/lbma-gold-price for an explanation of the new price setting approach.

[13] Our units of observation are hourly intervals, so AMFIX=1 for the hourly interval 10am-11am and PMFIX=1 for the hourly interval 3pm-4pm.

TABLE 4

Regressions of hourly futures market gold price discovery shares

	CS			*IS*			*ILS*		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
INTERCEPT	0.70***	0.68***	0.69***	0.66***	0.63***	0.61***	0.45***	0.43***	0.40***
	(165.24)	(120.95)	(99.71)	(154.09)	(112.47)	(86.71)	(76.26)	(61.33)	(46.89)
POST	0.16***	0.17***	0.15***	0.17***	0.20***	0.22***	0.00	0.02**	0.07***
	(71.96)	(29.86)	(17.67)	(76.88)	(35.89)	(26.57)	(0.86)	(2.38)	(5.57)
ASIA	0.02***	0.02***	0.01**	0.01**	-0.01**	-0.02***	-0.01	-0.04***	-0.05***
	(5.33)	(3.77)	(2.37)	(2.41)	(-2.40)	(-3.45)	(-1.33)	(-7.66)	(-7.70)
ASIA*POST		-0.01*	0.00		0.03***	0.04***		0.07***	0.08***
		(-1.94)	(0.38)		(5.41)	(7.21)		(9.09)	(9.05)
USA	0.04***	0.09***	0.08***	0.06***	0.13***	0.12***	0.03***	0.06***	0.06***
	(12.35)	(17.11)	(14.18)	(18.65)	(25.13)	(21.07)	(5.97)	(9.25)	(7.76)
USA*POST		-0.09***	-0.06***		-0.12***	-0.09***		-0.05***	-0.04***
		(-14.69)	(-9.47)		(-20.80)	(-13.77)		(-5.65)	(-3.56)
UK	0.02***	-0.01***	-0.02***	0.04***	0.03***	0.03***	0.04***	0.06***	0.06***
	(6.14)	(-3.31)	(-4.35)	(16.35)	(8.31)	(6.67)	(11.12)	(12.62)	(11.89)
UK*POST		0.06***	0.07***		0.02***	0.03***		-0.04***	-0.04***
		(13.17)	(13.94)		(3.76)	(5.25)		(-6.51)	(-5.68)
FTRSPREAD	-0.35***	-0.28***	-0.28***	-0.41***	-0.30***	-0.30***	-0.12***	-0.04***	-0.04***
	(-39.13)	(-26.71)	(-25.19)	(-41.03)	(-27.07)	(-25.91)	(-10.46)	(-3.28)	(-3.52)
FTRSPREAD*POST			-0.03			0.06*			0.10**
			(-0.92)			(1.71)			(2.23)
SPOTSPREAD	0.02***	0.02***	0.01**	0.06***	0.05***	0.10***	0.06***	0.05***	0.11***
	(6.95)	(5.84)	(2.03)	(18.36)	(15.84)	(13.86)	(12.48)	(10.78)	(13.18)
SPOTSPREAD*POST			0.01			-0.07***			-0.10***
			(1.26)			(-8.48)			(-8.71)
AMFIX	0.02***	0.02***	0.01	0.01**	0.01	-0.01	-0.01	-0.01	-0.03***
	(4.29)	(3.73)	(1.23)	(2.40)	(1.50)	(-1.25)	(-1.19)	(-1.59)	(-2.86)
AMFIX*POST			0.02			0.04***			0.04**
			(1.60)			(4.05)			(2.47)
PMFIX	0.02***	0.02***	0.05***	0.02***	0.02***	0.05***	-0.01	-0.01	0.00
	(4.46)	(4.28)	(7.67)	(6.06)	(5.86)	(9.07)	(-0.94)	(-1.05)	(0.07)
PMFIX*POST			-0.07***			-0.07***			-0.02
			(-8.43)			(-9.21)			(-1.23)
UKRATE	0.00	-0.01	-0.02	-0.01	-0.01	-0.04	-0.01	-0.01	-0.04
	(-0.22)	(-0.32)	(-0.48)	(-0.33)	(-0.43)	(-0.99)	(-0.38)	(-0.38)	(-0.91)
UKRATE*POST			0.02			0.06			0.05
			(0.56)			(1.33)			(0.83)
USRATE	-0.04	-0.05	0.00	-0.05*	-0.06**	-0.02	-0.02	-0.03	0.00
	(-1.38)	(-1.59)	(-0.07)	(-1.86)	(-2.16)	(-0.61)	(-0.60)	(-0.68)	(-0.08)
USRATE*POST			-0.09			-0.07			-0.04
			(-1.56)			(-1.32)			(-0.58)
UKEMPLOY	-0.03	-0.03	-0.07**	-0.05**	-0.05**	-0.10***	-0.01	-0.01	-0.02
	(-1.52)	(-1.63)	(-2.13)	(-2.20)	(-2.31)	(-2.67)	(-0.37)	(-0.40)	(-0.43)
UKEMPLOY*POST			0.09**			0.11***			0.01
			(2.28)			(2.69)			(0.21)

TABLE 4 CONT.

Regressions of hourly futures market gold price discovery shares

	CS			*IS*			*ILS*		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
USEMPLOY	-0.09***	-0.09***	-0.13***	-0.06***	-0.05**	-0.07*	0.02	0.02	0.01
	(-3.61)	(-3.45)	(-3.13)	(-2.65)	(-2.28)	(-1.94)	(0.53)	(0.70)	(0.18)
USEMPLOY*POST			0.08*			0.04			0.03
			(1.73)			(0.89)			(0.49)
UKGDP	0.00	0.00	-0.01	0.00	-0.01	-0.03	0.00	0.00	-0.04
	(0.05)	(-0.10)	(-0.17)	(-0.19)	(-0.31)	(-0.86)	(-0.03)	(-0.03)	(-1.08)
UKGDP*POST			0.01			0.05			0.08
			(0.21)			(1.25)			(1.35)
USGDP	-0.12***	-0.11***	-0.21***	-0.05***	-0.05**	-0.09**	0.08***	0.09***	0.16***
	(-5.06)	(-5.09)	(-5.96)	(-2.61)	(-2.27)	(-2.45)	(2.73)	(2.95)	(4.63)
USGDP*POST			0.20***			0.08**			-0.16***
			(5.06)			(2.13)			(-2.82)
UKPPI	-0.02	-0.02	-0.03	-0.02	-0.02	-0.02	0.01	0.00	0.00
	(-1.13)	(-1.19)	(-0.99)	(-0.80)	(-0.93)	(-0.61)	(0.23)	(0.15)	(-0.03)
UKPPI*POST			0.03			0.01			0.02
			(0.78)			(0.36)			(0.36)
USPPI	-0.05**	-0.05**	-0.09**	-0.02	-0.02	-0.04	0.06*	0.06**	0.05
	(-2.55)	(-2.48)	(-2.53)	(-1.09)	(-0.81)	(-1.19)	(1.81)	(1.99)	(1.14)
USPPI*POST			0.08*			0.05			0.03
			(1.88)			(1.36)			(0.43)
UKCPI	0.07***	0.05***	0.13***	0.05***	0.04**	0.12**	-0.03	-0.03	0.07
	(3.88)	(3.10)	(2.78)	(3.01)	(2.39)	(2.38)	(-0.69)	(-0.68)	(0.96)
UKCPI*POST			-0.11**			-0.12**			-0.13(-
			(-2.25)			(-2.18)			1.60)
USCPI	-0.05**	-0.05**	-0.12***	-0.03	-0.02	-0.06*	0.02	0.03	0.06*
	(-2.40)	(-2.42)	(-3.25)	(-1.30)	(-1.12)	(-1.75)	(0.78)	(0.93)	(1.66)
USCPI*POST			0.14***			0.08**			-0.08
			(3.59)			(2.18)			(-1.28)

Note: This table reports regression results in which the dependent variables are hourly gold futures price discovery shares (Component Share, CS; Information Share, IS; Information Leadership Share, ILS). POST is a dummy variable that takes the value of one after COMEX introduces the GLOBEX platform. ASIA, USA, and UK are dummy variables equal to one when the Hong Kong Stock Exchange, New York Stock Exchange and London Stock Exchange are trading, respectively. FTRSPREAD and SPOTSPREAD are the average spreads for each hour on COMEX and LOTC, respectively. AMFIX and PMFIX are dummy variables equal to one in the hour that the London gold fixing occurs. UKRATE and USRATE are dummy variables equal to one in the hour that the UK and US central banks announce a new target interest rate, respectively. UKEMPLOY and USEMPLOY are dummy variables equal to one in the hour that the UK and US governments announce employment figures, respectively. UKGDP and USGDP are dummy variables equal to one in the hour that the UK and US governments announce GDP figures, respectively. UKPPI and USPPI are dummy variables equal to one in the hour that the UK and US governments announce PPI figures, respectively. UKCPI and USCPI are dummy variables equal to one in the hour that the UK and US governments announce CPI figures, respectively. Newey-West *t*-statistics are reported in parenthesis. ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively.

Table 4 reports the regression results. The dependent variables are hourly price discovery estimates (*CS*, *IS*, and *ILS*) for the futures market and the independent variables are the determinants of price discovery described above.[14] We calculate *t*-statistics using Newey-West corrected standard errors. Model 1includes all the determinants of price discovery, Model 2 adds interaction of POST and the time zone dummy variables, and Model 3 adds all interactions between POST and the determinants of price discovery.

The COMEX trading platform upgrade (POST) has a highly significant, positive effect on all price discovery measures (with exception of the *ILS* in Model 1). According to Model 1, *CS*, *IS*, and *ILS* increase by 16%, 17%, and 0%, respectively after the change, holding other variables constant. These multivariate results support our earlier observations that the platform upgrade increased the relative liquidity of COMEX, thereby decreasing the amount of noise in COMEX prices (hence the increase in *CS* and *IS*), but did not have a large impact on the relative speed at which the futures market reflects new information (no significant change in *ILS*). The other regressions support this result with the POST coefficient for the *ILS* regressions remaining relatively small compared to the *CS* and *IS* regressions.

The variables measuring business hours in the three regions, ASIA, USA and UK, are statistically significant determinants of the intraday price discovery shares. The coefficients for USA and UK are positive, indicating that on average across the whole sample period the futures market increases its contribution to price discovery during US and UK business hours. For example, Model 1 implies that *ILS* is on average (across the full sample) 3% and 4% higher during US and UK business hours, ceteris paribus. The intraday patterns in price discovery change significantly after the upgrade and internationalization of COMEX (as we noted earlier in Figure 4). For the *ILS* regressions, the coefficient on the interaction of POST with the USA and UK dummy variables is negative and significant, largely negating the intraday pattern that existed before the upgrade of COMEX. For example, Model 2 indicates that before the upgrade, *ILS* is 6% higher during US business hours and 6% higher during UK business hours, whereas after the upgrade, it is only 2% (6% minus 4%) higher during US business hours and 1% (6% minus 5%) higher during UK business hours. For the *ILS* regressions, the coefficients associated with ASIA indicate that during Asian business hours, relatively less price discovery occurs in the futures market before the upgrade of COMEX, and that the upgrade attenuates this intraday pattern.

[14] Unlike Figure 1, the hourly price discovery estimates used in the regressions do not use a moving average.

Overall, these results show a general movement toward a more globalized gold market after the microstructure change, with diminishing importance of regional time zones.

The liquidity of both COMEX and LOTC (measured by their bid-ask spreads) has a significant impact on where and how price discovery occurs. Both *IS* and *CS* tend to be higher for markets that have less noisy prices. This tendency can be seen in the regressions with the large negative coefficients for FTRSPREAD and positive coefficients for SPOTSPREAD, indicating that a decrease in the relative liquidity of COMEX compared to the spot market (either wider spreads on COMEX or narrower spreads in the spot market) is associated with lower futures market *IS* and *CS*. The *ILS* regressions indicate that more price discovery occurs in the futures market when it is relatively more liquid, i.e., when the futures market spread is narrower or the spot market spread is wider. Because *ILS* is unaffected by the level of noise, the results from the *ILS* regressions suggest that informed traders prefer to trade in the more liquid market. This tendency is driven by the period before the COMEX upgrade, after which *ILS* is no longer positively associated with liquidity. One possibility is that informed traders become less sensitive to relative liquidity once it is plentiful.

Although the UK gold fixing is known to affect the price of gold, its effect on the location of price discovery is relatively modest. In the *ILS* regressions, the AMFIX coefficient is significantly negative suggesting that around the AMFIX the LOTC market increases in the speed at which it reflects new information (a decrease in the relative speed of the futures market). This effect is attenuated by the upgrade to COMEX. The PMFIX coefficients for the *CS* and *IS* regressions are highly significant and positive, yet there is no significance in the *ILS* regressions, indicating that immediately after the PM fix the LOTC market becomes relatively noisier than the futures market.

The effect of macroeconomic announcements varies across the different announcement types. Arguably, the most important announcement in our sample is the US GDP announcement, which is reflected in the significance and size of the coefficients for this variable (across all price discovery measures). The negative coefficient of USGDP in the *IS* and *CS* regressions suggests an increase in the relative noise of futures market quotes around US GDP announcements, while the positive coefficient in the *ILS* regressions indicates that around US GDP announcements, the US futures market takes on an even more important role in gold price discovery. In other words, much of the trading on US GDP news is likely to occur in the futures market, making it both faster to reflect the new information and noisier around such announcements. Interestingly, the significant changes in the location of price discovery around US GDP announcements are largely attenuated after the upgrade and internationalization of COMEX.

Similar to the US GDP result, in *CS* regressions, the coefficient for USEMPLOY is large, negative and highly significant, indicating an increase in noise in the futures market. The last noteworthy result in Table 4 is the UKCPI variable. The increases in futures market *IS* and *CS* are both large and highly significant (with no corresponding increase in *ILS*), indicating UK CPI announcements tend to be associated with a relative decrease in futures market noise or a relative increase in spot market noise.

In these regressions, the intercept is not very meaningful, representing a very small subset of the data. It represents the time period between 0:00 GMT and 1:00 GMT when there are no announcements, before the COMEX upgrade, and with hypothetical zero bid-ask spreads in both markets.

Overall the results indicate that although many factors affect the location of price discovery in the gold market, the upgrade of COMEX to a 24-hour electronic low latency platform with incentives for international participation causes many of these effects to diminish. This finding reflects greater globalization of international gold trade, with the traders' choice of markets being driven by market structure, as opposed to market location.

2.5.4 Robustness tests

In this section we assess the robustness of the results. We test the sensitivity of the results to the choice of lag length as well as different forms of the dependent and independent variables used in our regressions. Overall the results are qualitative unchanged in the robustness tests and do not change our overall conclusions.

The baseline models use 200 lags in the VECM as this should be sufficient to allow markets to reach equilibrium after a price change in one market. The decision to use 200 lags is somewhat arbitrary, so for robustness we use the Akaike Information Criterion (AIC) to select the optimal lag structure. We estimate a VECM with various lag lengths and compare the AIC across individual days and hours for a small randomly selected group of days. Taking the median suggests that 460 lags is the optimal lag length. We re-estimate all of our results using 460 lags in the VECM. Table 5 reports results from the re-estimated regressions.

Overall our results are robust to the increased lag length, and although our conclusions remain the same there are a few minor changes to the coefficients in our regressions. First, the effect of POST in the *ILS* regressions is negative but remains small, confirming that the COMEX platform upgrade did not have a large impact on the futures market's contribution to impounding new information. Second, the coefficient on FTRSPREAD*POST in *ILS* regression Model 3 is larger, suggesting that an increase in the COMEX spread after the microstructure change increases this market's share of price discovery. Third, some of the coefficients for the London Gold Fixings change in significance; however our interpretation does not, with coefficients remaining small and suggesting only a modest impact on the location of price discovery.

Many macroeconomic announcements simply confirm analyst expectations and thus do not result in much, if any, surprise. To test whether unexpected announcements have a different effect to announcements in general, we re-estimate the Table 4 regressions using announcement surprise instead of simple announcement times. In this case our dummy variables are equal to one if the announced change in the macroeconomic variable is contrary to analyst expectations as measured by the Bloomberg Analyst Survey. We find that our original results are robust to this specification.[15]

The final robustness test that we conduct is with the dependent variables. Instead of using the level of price discovery, we use the change in price discovery over the previous hour. This addresses concerns about the stationarity of the price discovery shares. We find that results are similar to our original specification.

Overall our results are robust to lag lengths, whether the outcome of an announcement is a surprise, and using first differences of the dependent variable.

[15] For conciseness these results are not reported, but are available from the authors upon request.

TABLE 5

Regressions of hourly futures market gold price discovery shares with 460 lags

	CS			*IS*			*ILS*		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
INTERCEPT	0.62*** (120.59)	0.62*** (95.09)	0.65*** (81.46)	0.59*** (122.72)	0.57*** (96.03)	0.56*** (74.52)	0.44*** (73.67)	0.43*** (58.19)	0.40*** (44.12)
POST	0.13*** (46.10)	0.12*** (16.74)	0.08*** (6.95)	0.10*** (38.02)	0.10*** (15.05)	0.13*** (12.72)	-0.05*** (-14.06)	-0.03*** (-3.77)	0.00 (0.31)
ASIA	0.05*** (13.11)	0.06*** (10.51)	0.05*** (8.54)	0.02*** (6.56)	-0.01 (-0.98)	-0.01* (-1.95)	-0.03*** (-6.87)	-0.06*** (-9.86)	-0.06*** (-8.71)
ASIA*POST		-0.02*** (-2.88)	0.00 (-0.36)		0.05*** (7.89)	0.06*** (8.97)		0.06*** (7.09)	0.05*** (5.63)
USA	0.03*** (7.17)	0.04*** (6.78)	0.03*** (4.92)	0.06*** (15.75)	0.11*** (19.90)	0.10*** (17.25)	0.03*** (4.99)	0.06*** (8.08)	0.06*** (7.92)
USA*POST		-0.02*** (-3.07)	0.00 (-0.16)		-0.08*** (-11.97)	-0.06*** (-8.07)		-0.05*** (-5.54)	-0.05*** (-5.41)
UK	0.03*** (10.54)	0.01* (1.85)	0.00 (0.64)	0.04*** (13.74)	0.02*** (5.41)	0.02*** (4.60)	0.01*** (3.24)	0.03*** (5.53)	0.03*** (6.15)
UK*POST		0.05*** (9.25)	0.06*** (10.07)		0.04*** (7.07)	0.04*** (6.94)		-0.03*** (-5.02)	-0.04*** (-5.68)
FTRSPREAD	-0.24*** (-24.63)	-0.22*** (-19.76)	-0.22*** (-18.36)	-0.28*** (-29.66)	-0.18*** (-17.80)	-0.18*** (-16.85)	-0.03*** (-3.15)	0.03*** (2.68)	0.02 (1.46)
FTRSPREAD*POST			-0.07 (-1.58)			0.00 (0.01)			0.22*** (4.22)
SPOTSPREAD	-0.01** (-2.16)	-0.01* (-1.93)	-0.04*** (-5.28)	0.04*** (10.82)	0.03*** (8.96)	0.07*** (9.37)	0.06*** (12.70)	0.06*** (11.23)	0.10*** (10.76)
SPOTSPREAD*POST			0.05*** (5.50)			-0.06*** (-5.97)			-0.08*** (-6.93)
AMFIX	0.03*** (5.83)	0.03*** (5.72)	0.02* (1.84)	0.02*** (3.58)	0.02*** (2.91)	-0.01 (-0.71)	-0.02*** (-2.99)	-0.02*** (-3.31)	-0.03*** (-2.82)
AMFIX*POST			0.03*** (2.84)			0.05*** (4.30)			0.01 (0.79)
PMFIX	0.01** (2.20)	0.01** (2.20)	0.04*** (5.08)	0.01 (1.41)	0.01 (1.48)	0.02*** (3.18)	0.00 (0.18)	0.00 (0.20)	-0.02* (-1.94)
PMFIX*POST			-0.07*** (-5.66)			-0.03*** (-3.24)			0.05*** (3.08)
UKRATE	0.00 (-0.08)	0.00 (-0.12)	-0.01 (-0.36)	0.00 (0.01)	0.00 (-0.04)	-0.04 (-0.92)	-0.01 (-0.28)	-0.01 (-0.29)	-0.03 (-0.62)
UKRATE*POST			0.02 (0.48)			0.07 (1.46)			0.03 (0.55)
USRATE	0.03 (0.92)	0.03 (0.89)	0.08* (1.73)	-0.04 (-1.25)	-0.04 (-1.39)	-0.06 (-1.45)	-0.06 (-1.59)	-0.06 (-1.63)	-0.15*** (-2.67)
USRATE*POST			-0.10 (-1.61)			0.04 (0.66)			0.17** (2.22)
UKEMPLOY	-0.03 (-1.13)	-0.03 (-1.17)	-0.04 (-1.08)	-0.04 (-1.60)	-0.04* (-1.72)	-0.07** (-2.12)	0.02 (0.64)	0.02 (0.59)	0.00 (-0.04)
UKEMPLOY*POST			0.03 (0.58)			0.07* (1.72)			0.04 (0.66)

TABLE 5 CONT.

Regressions of hourly futures market gold price discovery shares with 460 lags

	CS			*IS*			*ILS*		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
USEMPLOY	-0.08***	-0.08***	-0.11***	-0.02	-0.01	0.00	0.05*	0.06*	0.12***
	(-3.03)	(-3.00)	(-2.71)	(-0.91)	(-0.60)	(-0.12)	(1.69)	(1.85)	(2.85)
USEMPLOY*POST			0.06			-0.02			-0.13**
			(1.04)			(-0.46)			(-2.19)
UKGDP	0.00	-0.01	0.00	-0.01	-0.01	-0.02	0.02	0.02	-0.02
	(-0.19)	(-0.25)	(0.10)	(-0.25)	(-0.40)	(-0.40)	(0.59)	(0.55)	(-0.51)
UKGDP*POST			-0.02			0.01			0.07
			(-0.37)			(0.27)			(1.21)
USGDP	-0.08***	-0.08***	-0.18***	-0.02	-0.01	-0.06	0.06*	0.06**	0.12***
	(-3.20)	(-3.27)	(-4.43)	(-0.81)	(-0.60)	(-1.64)	(1.85)	(2.01)	(2.82)
USGDP*POST			0.19***			0.10**			-0.12**
			(4.00)			(2.18)			(-2.00)
UKPPI	-0.02	-0.02	-0.01	-0.02	-0.02	-0.01	-0.02	-0.02	-0.03
	(-0.89)	(-0.87)	(-0.26)	(-0.93)	(-0.99)	(-0.17)	(-0.58)	(-0.65)	(-0.76)
UKPPI*POST			-0.02			-0.03			0.02
			(-0.38)			(-0.66)			(0.37)
USPPI	-0.07**	-0.07**	-0.14***	-0.02	-0.02	-0.05	0.04	0.04	0.08*
	(-2.27)	(-2.30)	(-3.16)	(-0.93)	(-0.74)	(-1.32)	(1.11)	(1.25)	(1.74)
USPPI*POST			0.15***			0.06			-0.08
			(2.65)			(1.24)			(-1.20)
UKCPI	0.08***	0.07**	0.15**	0.07***	0.06**	0.14**	-0.04	-0.04	-0.07
	(2.84)	(2.48)	(2.47)	(2.99)	(2.40)	(2.51)	(-1.05)	(-1.06)	(-0.94)
UKCPI*POST			-0.11			-0.10*			0.04
			(-1.58)			(-1.68)			(0.42)
USCPI	-0.07***	-0.07***	-0.11***	-0.04*	-0.04*	-0.04	0.02	0.02	0.07
	(-2.62)	(-2.72)	(-3.06)	(-1.77)	(-1.69)	(-1.27)	(0.57)	(0.70)	(1.59)
USCPI*POST			0.09*			0.01			-0.09
			(1.87)			(0.24)			(-1.54)

Note: This table reports regression results in which the dependent variables are hourly gold futures price discovery shares (Component Share, CS; Information Share, IS; Information Leadership Share, ILS). In contrast to Table 4, the price discovery metrics here are estimated using 460 lags of one-second intervals as suggested by AIC. POST is a dummy variable that takes the value of one after COMEX introduces the GLOBEX platform. ASIA, USA, and UK are dummy variables equal to one when the Hong Kong Stock Exchange, New York Stock Exchange and London Stock Exchange are trading, respectively. FTRSPREAD and SPOTSPREAD are the average spreads for each hour on COMEX and LOTC, respectively. AMFIX and PMFIX are dummy variables equal to one in the hour that the London gold fixing occurs. UKRATE and USRATE are dummy variables equal to one in the hour that the UK and US central banks announce a new target interest rate, respectively. UKEMPLOY and USEMPLOY are dummy variables equal to one in the hour that the UK and US governments announce employment figures, respectively. UKGDP and USGDP are dummy variables equal to one in the hour that the UK and US governments announce GDP figures, respectively. UKPPI and USPPI are dummy variables equal to one in the hour that the UK and US governments announce PPI figures, respectively. UKCPI and USCPI are dummy variables equal to one in the hour that the UK and US governments announce CPI figures, respectively. Newey-West *t*-statistics are reported in parenthesis. ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively.

2.6. Conclusion

This paper investigates price discovery in the global gold market using high-frequency data for the period 1997 to 2014. The gold market is of interest to both researchers and industry due to its immense size and economic importance.

Our first major finding is that the US futures market contributes more to price discovery compared to the London OTC spot market, despite being ten times smaller. This result indicates that market structure and instrument type is of greater importance than market size and liquidity. The finding supports the idea that the futures market's key features of centralization, relative transparency, inbuilt leverage, and ability to avoid dealing with the underlying asset, contribute to its disproportionately large role in price discovery.

Our second key finding reinforces the importance of market structure on the process of price discovery. Changes in the US gold futures market from a floor-based system to the nearly 24-hour, fully electronic low latency GLOBEX platform is associated with a considerable decrease in noisiness of prices in this market, but no increase in the relative speed with which the market reflects new information.

Our third key finding is that several factors affect the location of price discovery. Price discovery shares vary substantially at both daily and intraday levels. Prior to the futures market's upgrade to the GLOBEX platform, the location of price discovery is largely dictated by daylight hours within each locale, with the US futures market contributing more to price discovery during US daylight hours. After the move to the GLOBEX platform and introduction of international participation incentives in the futures market, intraday price discovery variation declines with US futures taking a consistent lead irrespective of market hours. This decline illustrates how a market structure that is more accessible internationally can negate the importance of geographic location and create a more integrated global market.

Our findings carry important implications for market design. We show that market structure is of greater importance to price discovery than market size and liquidity. Changes to market structure can have opposite effects on the speed at which prices reflect new information and accuracy with which they reflect the information (amount of noise), as is the case in the market structure change that we examine. Market designers can improve the efficiency of their markets by carefully considering the structures they implement, as this can have greater impact than methods focused simply on increasing turnover and participation.

There are many avenues left open for future research. Markets are implementing new structural changes on a continual basis; other structural changes in the gold market may have a different effect on price discovery. Further, gold mining volumes are small compared to holdings, meaning that gold is more decentralized than many other commodities that are consumed. This market allows for research into the centralization of markets. Although many markets are becoming more decentralized, the gold market seems to be moving in the opposite direction. The effect of the new benchmark process for the gold price (introduced in 2015) also remains unclear. If the new process increases transparency, as it is designed to do, it should result in a move back towards the LOTC market as a price setter. Finally, it will be interesting to see whether the conclusion of the COMEX International Incentive Program on December 31, 2015 will have any effect on the price discovery shares for the gold market. This program provides lower cost trading to traders outside the US.

These findings can also be confirmed on other commodity markets such as oil and silver, where the underlying usage of the commodity is slightly different to gold. Oil is a clear consumption good, whereas silver is placed between oil and gold.

APPENDIX - Chapter 2:
Calculation of Component Shares (CS) and Information Shares (IS)

We estimate the *IS* and *CS* metrics using the error correction parameters and variance-covariance of the error terms from equation (1) and (2) as in Baillie *et al.* (2002). The component shares are obtained from the normalized orthogonal to the vector of error correction coefficients, $\alpha_{\perp} = (\gamma_1, \gamma_2)'$, thus:

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}, \quad CS_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2} \tag{A.1}$$

Given the covariance matrix of the reduced form VECM error terms,

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \tag{A.2}$$

and its Cholesky factorization, $\Omega = MM'$, where

$$\mathrm{M} = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2(1-\rho^2)^{1/2} \end{pmatrix}, \tag{A.3}$$

we calculate the *IS* using:

$$IS_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, \quad IS_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}. \qquad \text{(A.4)}$$

Because *IS* is impacted by the order of the price series in the Cholesky factorization, we calculate *IS* under each of the potential orderings and take the simple average, as per Baillie *et al.* (2002).

Chapter 3: Borrower payments and mortgage risk

3.1. Motivation

Banks collect most borrower information when making a lending decision. Post origination, banks rely on changes in house prices through re-valuations and macro indicators. However, borrower conditions can change substantially for many reasons, including changes in employment, income, lifestyle and risk preference. This paper analyzes to which degree banks may be informed through transactional accounts.

Puri *et al.* (2017) demonstrate that bank relationships such as bank accounts are informative and help in reducing credit risk. Our paper follows this thought and analyzes the impact of over three billion transaction observations, which capture habitual and behavioral borrower aspects included in income and expense data from debit cards, credit cards and bank transactions on the performance of loans.

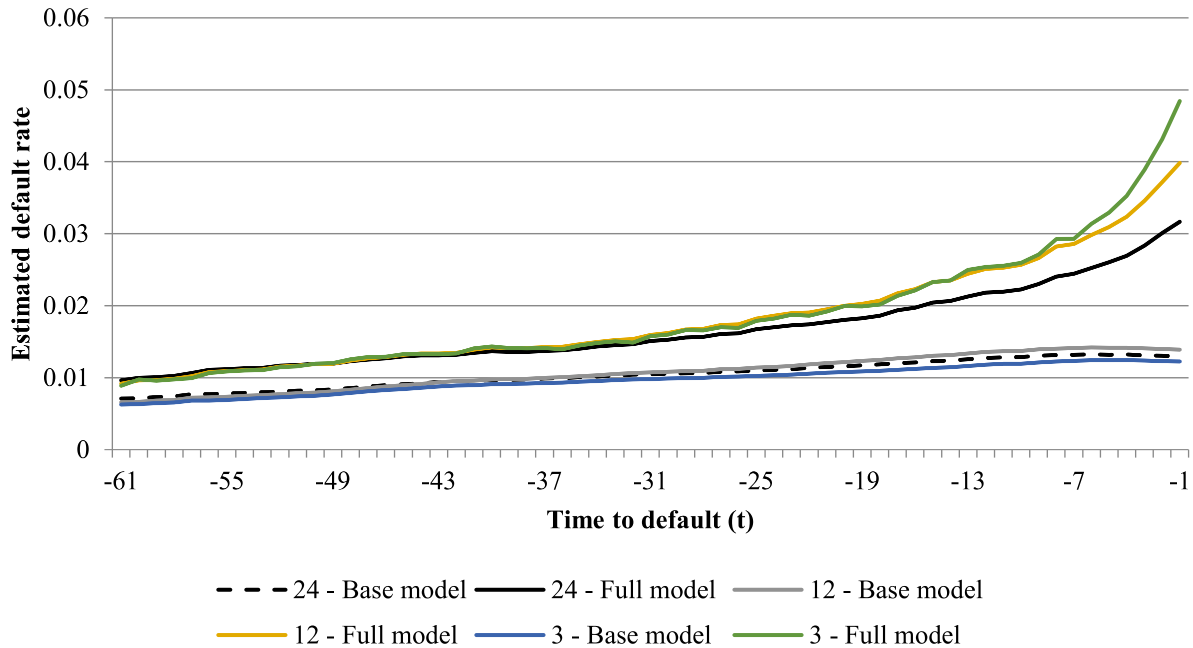
We find that short-term interest coverage, income changes, home maintenance expenses, clothing expenses and cash withdrawals all contribute to the prediction of future default. The timeliness and accuracy of these variables allow for separation of risky and safe borrowers across the loans' life, with changes in factors allowing for time varying risk estimates. This is a large improvement from traditional models that depend in large parts on invariant origination variables.

Our model improves upon a traditional credit risk model and accurately captures the time varying nature of credit risk within our sample. As a borrower nears default, their change in behavior leads to a change in their estimated credit risk as depicted in Figure 1. As a borrower becomes increasingly stressed, our model reflects this through increasing default risk estimates.

The data is collected from a top four Australian retail bank with over 1,000 branches. Consumers are encouraged to package credit cards and transaction accounts with their mortgages. This feature allows for the link between transactions and mortgages. This contrasts with US banks, where lending and other products are generally separate. While this data is available within banks, to our knowledge it has not been used in this capacity. Australian mortgages are nonrecourse, meaning the collections of the lender are limited to the asset value. This contrasts with the USA, where full recourse loans are more common.

Figure 1

This figure compares the use of different modelled default lags on the estimated probability of default for defaulting loans as they approach default. Default is at t = 0, and negative t represents months to default. The solid lines represent the average probability estimate for the full model and the dashed lines represent the base model. The full model estimates are higher than the base model estimates, particularly as default becomes imminent.



We make several methodological contributions. Transaction data has the distinct feature that transactions may or may not be observed in a period, while traditional credit risk data is usually complete, with lending, borrower and collateral characteristics obtained at the time of credit granting. We devise a methodology to include these transactions and demonstrate that missing values are also informative to mortgage default. Our methodology includes a spline that controls for missing values for each dependent variable.

We also propose novel measures for the marginal information of competing models. Namely, we derive and calculate transformations of Pseudo R-Square and the Area Under the Receiving Operating Characteristic (AUROC) curve by investigating the improvement magnitude (IMPMAG) and the reduction in the improvable area (IMPRED).

A major outcome of our research is the formation of an early warning system for loans which are increasing in risk. We show that a borrower's recent financial transactions are powerful predictors of future default and complement the traditional untimely origination variables. This increase in model accuracy can assist a financial institution's portfolio risk management efforts as well as permit earlier intervention to assist consumers.

A more accurate and timely measure of risk can allow for more precise loan pricing, capital allocation and loan loss provisioning in financial institutions. For example, under IFRS 9 loan loss provisioning, banks are required to identify loans with a significant increase in credit risk. Our analysis shows that such increases can be much better identified by timely transactional variables that reflect variation in credit risk. Regulators may benefit from a greater understanding of the link between spending and credit default, allowing for more effective policymaking. Households may benefit as banks can provide more efficient pricing based on their payment details which may lead to greater housing affordability.

The paper is structured as follows; Section 2 reviews the literature. In Section 3, we describe the data and present sample statistics. Section 4 outlines the testing methodology which includes a two-stage regression model with clustered standard errors to control for adverse selection in relation to mortgage payoffs. In Section 5 we estimate a baseline regression model based on the current state of the literature. Further, we develop risk factors based on transactional data and present a combined Full model that includes the baseline model plus factors derived from transactional data. Section 6 contains additional robustness tests and Section 7 concludes.

3.2. Literature review

3.2.1 Borrower payment transactions

Transaction data is difficult to obtain, and few studies exist. In this section we survey the studies that use transactional data. Aydın (2018) studied consumer propensity to use credit lines in Europe. The paper finds increased credit availability leads to increased consumption. Agarwal and Qian (2014) test the consumption and debt response to unanticipated income shocks. Agarwal *et al.* (2015b) compare spending changes as people reach retirement.

Closest to our paper are two studies linking housing wealth to consumption. Agarwal and Qian (2017) use Singaporean credit and debit card data to test whether an increase in access to home equity affects consumption. Gan (2010) motivates the use of transactional data by explaining the link between housing wealth and household consumption in Hong Kong. These papers are not in credit risk but find correlation between lifestyle and consumer liquidity changes.

Norden and Weber (2010) link checking account activity and information production to borrower risk. They concentrate on credit line usage, limit violations and aggregate flows but do not include a measure developed through granular transaction level data. Using German data, Puri *et al.* (2017) conclude that a bank can reduce the default risk of a borrower by first strengthening the banking relation with less risky products, such as a transaction accounts. Studying peer-to-peer lending in the US, Miller (2015) finds increasing the amount of information available to lenders greatly increases the efficiency of lending markets, especially among higher risk borrowers.

Our study moves beyond these and links borrower banking transactions directly to mortgage default risk. This link has far-reaching implications as mortgages are the largest asset class for most banks.

3.2.2 Mortgage default

Characteristics at origination

Prior research has analyzed loan characteristics as determinants of mortgage default. Loans with a higher loan to value ratio are more likely to default (Gerardi *et al.* 2017; Bian *et al.* 2018), and there is a positive and nonlinear relation between loan size and default. Default rate changes nonlinearly with loan age, and previously delinquent loans are more likely to become delinquent (Qi & Yang 2009). Broker and correspondent originated loans are riskier compared to bank originated loans (Elul *et al.* 2010). Low documentation level, nontraditional, sub-prime and loan for purchase are more likely to default (Qi & Yang 2009). Property type and owner occupancy have received mixed findings, whereas a higher interest rate at origination and a higher rate for adjustable rate mortgages (ARM) are riskier (Foote *et al.* 2008).

Borrower specific characteristics have also been considered at loan origination. Credit score at origination is a typical inclusion in credit default models, with higher or improving scores indicating less risk (Elul *et al.* 2010). Borrowers with higher debt to income ratio, or credit card utilization are more likely to default (Herkenhoff 2012).

Home equity over time

The large decreases in house prices and the high unemployment rate during the Global Financial Crisis (GFC) have been studied in the recent literature. Demyanyk and Van Hemert (2009) find that delinquency across all loans were affected by the collapse, and that the unsustainable growth of the mortgage market was masked by increased property values. Qi and Yang (2009) find that the current loan to value ratio is the biggest driver of default among high loan-to-value mortgages. Foote *et al.* (2008) find most borrowers with negative equity did not default.

This finding contradicts the notion of "strategic" default, prompting additional studies exploring the relation between double trigger events, and sunk cost fallacy (Agarwal *et al.* 2015a). Guiso *et al.* (2013) found severe negative equity is the strongest factor for strategic default next to race, gender, expectations about future employment, and views about fairness and morality.

Borrower liquidity over time

Liquidity shocks can include job loss, divorce, health shock, or other accident which are generally measured using sparse survey-based data. The link between negative equity and job loss is becoming increasingly clear. Gerardi *et al.* (2017) find that job loss is equivalent to a 35% reduction in home equity. Bricker and Bucks (2016) investigate borrower mobility and negative equity, concluding that a liquidity shock significantly increases the probability of a borrower moving to a new location.

Campbell and Cocco (2015) find liquidity shocks in the form of borrowing constraints are drivers of mortgage default. The high cost of alternate liquidity is offset by the liquidity required to mitigate imminent default. The inability of a household to find further sources of credit leads to default.

Kartashova and Tomlin (2017) indicate a relation between house prices and non-mortgage debt, suggesting that increases in equity are partially negated by increases in other types of debt used for non-housing consumption. These studies motivate our use of borrower spending and credit usage to derive more easily borrower liquidity. We develop a clearer understanding of general consumption and its effects on mortgage default.

Macroeconomic information

Increases in house prices have been associated with decreases in the default rate as refinancing and borrower wealth increases (Crook & Bellotti 2010; Chan *et al.* 2016). Transformations of house price indices (HPI) are included in our payoff model (Carmichael & Coën 2018). Mortgages are more likely to default when geographic unemployment measures increase. Mocetti and Viviano (2017) find income changes after mortgage origination are strong drivers of delinquency.

Higher interest rates are associated with a higher likelihood of default (Aoki *et al.* 2004; Crook & Banasik 2012). Changes in interest rates are also important, as they directly impact the repayments for adjustable rate loans after origination.

3.3 Data

The loan level, origination and performance data for mortgages are provided by a top four Australian retail bank with over 1,000 branches.[16] The bank's business covers all major retail banking products including transaction and savings accounts, credit cards, loans and mortgages.

The data used for this study comprises over three billion individual transaction observations from 6,568,224 account-months and 179,465 distinct accounts. Borrowers in our sample transact on average 46.39 times a month through their debit and credit cards and receive income. The sample period encompasses account-month observations of loan level variables for the period 1 January 2011 to 31 December 2016.

All customers of the bank are attributed a common identification key across all products, allowing for an accurate link from the loan to the borrower's other accounts. The data allows for very accurate measurement of all transactions conducted through and across all the accounts a borrower may hold with the bank.

[16] The data may be comparable in size to with Puri *et al.* (2017) who analyze 296 banks trading under one brand ("Sparkasse").

Table 1: Frequency counts by origination year

This table shows the number of mortgages in the sample by origination year. Default is defined as more than 90 days in arrears. The table includes the number of defaults and the default rate by origination year. Year represents the year that a mortgage was originated, Accounts represents the number of accounts that were originated in that year and are present in our dataset. Default represents the number of accounts originated in the default year within our sample period. The default rate is the number of defaults divided by the number of accounts. The default rate is high due to the stratified sampling. A larger proportion of defaults were kept compared to non-defaults.

Year	Accounts	Default	Default rate
Prior to 1995	362	8	0.022
1996	566	6	0.011
1997	1,110	10	0.009
1998	1,860	25	0.013
1999	2,589	38	0.015
2000	2,919	56	0.019
2001	3,851	57	0.015
2002	4,890	79	0.016
2003	5,307	102	0.019
2004	5,099	122	0.024
2005	7,091	195	0.027
2006	9,890	256	0.026
2007	13,121	402	0.031
2008	14,229	481	0.034
2009	21,266	706	0.033
2010	15,106	456	0.030
2011	14,969	386	0.026
2012	13,024	291	0.022
2013	15,549	286	0.018
2014	14,548	236	0.016
2015	12,119	98	0.008
All	179,465	4,296	0.024

The sample provided by the bank includes a random subset of the entire mortgage book over time, along with any attached savings, checking and credit card accounts. The data is a stratified random sample of defaulting and non-defaulting mortgages for the sample period. The default rate is higher than the bank's true default rate due to the sampling. The sample is constructed to obscure the true default rate and aggregate portfolio characteristics of the bank. To orthogonalize from the shift in average default rate, we use a logit model for which intercepts change while other estimated coefficients remain stable for parallel changes in default rates. Thus, the estimated coefficients in this paper will be similar in magnitude to those estimated on the full loan population (King & Zeng 2001).

Various loan level variables are available including loan to value ratio (LTV), income at origination, credit bureau score at origination, interest rates, and various other variables the firm uses to monitor the current risk in their portfolio. We test model robustness for autocorrelation by estimating the models after keeping only January observations, with the results being materially similar. This test results in only one observation per loan for each year.

We restrict our sample to owner occupier mortgages to increase homogeneity, accounts for which we measure at least one month of income exceeding $2,000 and average monthly card spend over all observations of $500. This ensures we capture a large proportion of a borrower's banking interactions. Non-owner occupier loans may obscure results as more elements and drivers are present with rental income, multiple loan portfolios and a general increase in financial complexity for these borrowers. No other restrictions are placed on the sample. We winsorize predictive variables at the 5% and 95% level to control for potential outliers.

Table 1 illustrates the default rate by loan origination year. Each mortgage is included once. 4,296 of the 179,465 mortgage loans in the sample default at some point in our sample period, representing a default rate of 2.4%. Default is defined as being in arrears for more than 90 days.

Table 2: Sample statistics by observation year

This table shows the account-month distribution of our mortgage sample. The account-months are collapsed into years. Therefore, each account is represented in every month for which it is present. The year column represents the observation year, number of account-months represents the number of observations. Default accounts is the count of distinct accounts that become delinquent in the next 12 months, Default account-months represents the number of observations that default within the next 12 months. Default rate account –months represents the number of account-months delinquent divided by the total number of account-months. The default rate is the number of defaults divided by the number of accounts. The default rate is high due to the stratified sampling. A larger proportion of defaults were kept compared to non-defaults.

Year	Number of account-months	Default accounts	Default account - months	Default rate account-months
2011	1,298,640	909	5,686	0.0044
2012	1,320,964	561	6,197	0.0047
2013	1,329,484	658	7,327	0.0055
2014	1,321,487	765	8,365	0.0063
2015	1,297,649	1,403	13,294	0.0102
All	6,568,224	4,296	40,869	0.0062

Table 2 contains a summary of the account-month observations of our loan sample. There are 6,568,224 account-months, with 40,869 observations becoming delinquent within the following 12-month period. The default accounts column indicates the number of accounts and the default account-months column the number of account-months in default within the following 12-month period. Due to the data being at a monthly level, and default being defined across multiple months, the multiple default observations per account is aligned with the multiple monthly non-default observations. We demonstrate in the robustness section that an alternate methodology, which keeps only one monthly observation per year per account, leads to consistent results.

Table 3 contains the descriptive statistics for our Base model variables, which include traditional default drivers and shows low pairwise correlation. For many of the variables, there is a distinct difference between defaulting and non-defaulting loans. For example, the loan to income ratio (APPLOANINCOME) for the non-default population is lower for the default population (315% vs 344%) and the mean credit bureau score for non-defaulting loans is 636.77 and for defaulting loans is 563.01. A similar relation is indicated by the other variables where defaulting loans have a riskier level.

Table 3: Descriptive statistics of base variables

This table presents the descriptive statistics of monthly mortgage-level variables at both origination and observation times for all observations by default status. The table includes the numerical variables used in the Base model, and the binary transformation of categorical variables used in the Base model. Default is defined as 90-day arrears, and the default flag is equal to one if the account defaults within the next 12 months.

Panel A: Numerical Variables

	Non-default		.	Default		.
Variable	Mean	Median	Standard deviation	Mean	Median	Standard deviation
APPLOANINCOME	314.99	315.00	165.50	344.07	350.00	160.09
BUREAU	636.77	672.00	214.32	563.01	585.00	184.28
IMR	0.72	0.72	0.02	0.72	0.72	0.02
INTMARGIN	2.72	2.71	0.31	2.79	2.78	0.31
FIXEDINT	0.13	0.00	0.33	0.13	0.00	0.34
CCUTIL	30.24	16.00	33.69	48.08	54.00	45.17
CLTV	61.45	67.79	22.13	70.22	76.45	19.07
LNOUTSBAL	11.91	12.15	1.01	12.25	12.44	0.77
LNINCOME	7.19	7.15	0.27	7.13	7.15	0.26
LNPROP	12.68	12.73	0.64	12.63	12.68	0.61
LOWDOC	0.02	0.00	0.14	0.04	0.00	0.19

Table 3 also contains the descriptive statistics for binary transformations of the categorical variables used in the models. In our case, only two variables are used, namely interest rate type (FIXEDINT), which is a variable taking the value of one if fixed and zero otherwise, and a low-documentation flag (LOWDOC), which is equal to one for low documentation loans. We observe the largest effect of the categorical variables at the univariate summary level for low-documentation loans with a higher default rate.

3.3.1 Data pre-processing of transaction data

Debit and credit card transactions

We group together the debit and credit card transactions as they are comparable. The transaction level data includes the date, the amount, the merchant name and the merchant category code (MCC).

Credit and debit card purchases are recorded with the name of the merchant as well as the merchant category code (MCC). To start the classification of MCCs, we use the Visa Merchant Data Standards Manual (VISA 2017). The MCC is unique for merchant type and not merchant name. For example, different supermarkets have the same MCC (5411). The allocation of all MCCs can be found in Appendix 1. An audit of the main merchant types was conducted to ensure code definitions are correct.

Card transactions are extracted from the relevant source system for the primary mortgage holder and matched to our transaction categories using the MCC. These are aggregated to a borrower -month level and merged with the mortgage data using the borrower ID.

Debit and credit card transactions have some exclusive transaction types. Credit cards include interest charges, payments, and cash advance fees. Debit cards conversely include ATM withdrawal as a major transaction type.

Bank transactions

Bank transactions contain all individual transactions which affect the bank balance, including interest received, direct transfers, income and salary payments, periodical debits, and debit card transactions. Information we collect from bank transactions include salary and wage payments, pensions, welfare payments and other income as observed by the bank. These transactions are not currently used as part of the credit risk process and give a clearer indication of income compared to loan application data.

3.3.2 Final sample descriptive statistics

Table 4 summarizes the aggregated data. Panel A of Table 4 contains sample statistics highlighting the size of the data collected from the bank account transactions. Only observations related to income are displayed. In our sample, there are a total of 5.8 million income observations, with total income received of 8.7 billion dollars. We note this includes final income received after tax and other deductions are removed by the employer, representing the final disposable income. Total salary of 8.1 billion dollars indicates most of the income is due to employment, as would be expected for current mortgage holders. The remainder of 0.6 billion dollars marked as Other Income includes welfare, pensions and other smaller amounts of income as systematically captured by the bank.

Panel B aggregates the credit and debit card transactions in a similar fashion to Panel A. Overall, the total amount of expenses is lower than the total amount of income, with the number of transactions being higher. We separate counts of clothing, maintenance, and cash out as these variables are used later in the paper. Although these transaction categories include only a small proportion of the total card transactions present (15%), the predictive power indicates the value of their inclusion in behavioral based credit risk models.

Finally, Panel C contains an example of the average borrower, with means calculated at the account level, then means and medians calculated at the portfolio level. From the table we can see non-defaulting borrowers average a total spend of $3,391 and monthly interest on the mortgage of $1,105, which is lower than average income ($4,890). This differs for the defaulting population where interest plus total spend ($3,674) exceeds income ($2,892) for both the mean and median. Table 3 indicates defaulting and non-defaulting populations have roughly the same income (LNINCOME) at loan origination, suggesting a change in income can be a powerful driver of default, and it is not currently being captured by operational credit default models.

Table 4: New transactional behavioral variables

This table illustrates sample statistics of the transactional data. Panel A and B present the number and dollar amount of transactions that are processed for each month. Panel A specifically concentrates on the salary and income received by the primary mortgage holder and Panel B concentrates on selected transaction categories. Panel C contains mean and median transaction levels for the primary mortgage holder for non-defaulting loans and loans that will default within the next 12 months. Panel C indicates the average borrower and their income and expense levels. These levels show that a large portion of a borrower's total transactions are present in our sample.

Panel A: Descriptive stats for other transaction data

Year	Total Value of income ($)	Number of income transactions	Total Salary ($)	Total salary transactions	Total other income ($)	Other income transactions
2011	1,702,300,433	1,188,686	1,582,121,237	888,370	120,179,195	300,316
2012	1,790,128,262	1,211,631	1,665,972,836	907,563	124,155,426	304,068
2013	1,759,786,421	1,164,249	1,637,631,300	875,782	122,155,121	288,467
2014	1,709,864,572	1,106,156	1,591,434,636	837,673	118,429,936	268,483
2015	1,730,894,865	1,084,910	1,611,792,109	827,778	119,102,756	257,132
All	8,692,974,552	5,755,632	8,088,952,119	4,337,166	604,022,433	1,418,466

Panel B: Descriptive statistics of Credit and Debit card transaction data

Year	Total transaction value	Count total transactions	Count clothing	Count home maintenance	Count cash out	Count other
2011	526,929,727	48,971,830	2,874,864	2,125,791	3,170,993	43,523,306
2012	599,259,001	55,333,333	3,070,974	2,436,699	3,308,106	49,371,567
2013	654,489,938	61,351,758	3,278,619	2,640,442	3,254,324	55,171,722
2014	725,675,679	64,629,815	3,318,945	2,824,768	3,181,680	58,298,438
2015	751,757,281	68,655,024	3,352,214	2,913,359	3,017,419	62,318,698
All	3,258,111,627	298,941,760	15,895,616	12,941,059	15,932,522	268,683,731

Table 4: Continued

Panel C: New transactional/behavioral variables

Variable	Non-default Mean	Non-default Median	Non-default Standard deviation	Default Mean	Default Median	Default Standard deviation
Income related transactions						
Income	4,890.19	4,603.89	3,024.08	2,892.12	2,463.87	2,547.11
Salary	4,580.39	4,376.54	3,141.23	2,550.39	2,078.69	2,619.57
Mortgage calculated expenses						
Mortgage interest	1,104.93	977.83	913.60	1,284.49	1,162.78	915.91
Card expenses						
Alcohol	47.34	21.38	61.23	38.27	9.37	60.97
Cash out	287.63	71.83	396.14	481.76	330.83	489.90
Clothing	191.17	130.85	186.39	75.20	23.15	126.96
Financial services	215.16	135.59	231.10	115.03	18.49	197.27
Groceries	518.95	410.14	431.00	364.71	242.01	379.57
Health services	93.90	63.27	96.94	41.97	10.34	72.86
Home maintenance	273.40	194.41	272.76	129.22	37.52	230.26
Local transport	189.17	154.51	156.24	151.83	104.47	158.58
Other	995.20	804.94	750.41	631.11	426.89	657.99
Restaurants	171.90	120.58	163.74	118.76	61.64	149.37
Travel	204.40	135.41	224.47	99.51	15.48	184.66
Utilities	130.84	74.07	147.07	100.07	40.83	136.18
Vehicles	72.03	51.06	77.20	41.60	4.17	78.03
Total spend	3,391.10	3,030.08	2,049.45	2,389.04	2,025.11	1,817.26

We see the reduction in income between our defaulting and non-defaulting populations is matched by a reduction in total spend as borrowers become more financially constrained. This change in spending forms the basis of this research, with the intention of predicting future default based on spending patterns. The higher monthly interest for defaulting loans aligns to larger loan balances and higher marginal interest rates.

In conclusion, the spending patterns of borrowers at an aggregate level is useful in separating at risk borrowers from those not at risk. Knowing the current serviceability of a borrower based on their financial transactions is a powerful indicator of the current risk of a loan.

3.4. Methodological framework

3.4.1 Two stage regression model

In this section, we establish a baseline with which to test and quantify the effectiveness of the new transactional data. We estimate a two-stage selection model to correct for potential bias due to the payoff selection by less risky borrowers. Typically, borrowers with greater credit quality or higher levels of financial literacy are more likely to pay off their loan (Bajo & Barbi 2018). These borrowers self-select and are removed from the population, increasing the risk in the remaining population.

We compute the inverse Mills ratio based on a regression of the probability of payoff on observable mortgage information in a first stage and include this ratio in the second stage regression for the probability of default. We estimate the following logit model for the first stage regression.

Stage 1

$$Prob(P_{it} = 1|X_{1it-1}) = \frac{1}{1 + e^{-(\alpha_P + \beta_P X_{Pit-1})}} \tag{1}$$

Where P_{it} is a binary variable representing payoff within the next 12 months and X_{Pit-1}represents lagged mortgage level variables correlated to future payoff. i and t represent variable and time respectively. Since the probability of payoff is the expectation of a latent variable process and contrary to OLS regression models, no error term is necessary. The parameters are estimated by maximizing the natural logarithm of the likelihood function:

$$L = \prod_{i=1}^{I}\prod_{t=1}^{T} Prob(P_{it} = 1|X_{Pit-1})^{P_{it}}\left(1 - Prob(P_{it} = 1|X_{Pit-1})\right)^{1-P_{it}} \tag{2}$$

The Inverse Mills Ratio (IMR_{it-1}) is calculated using the estimates of the first stage through the following formula:

$$IMR_{it-1} = \frac{PDF\left(ICDF\left(Prob(P_{it} = 1|X_{Pit-1})\right)\right)}{Prob(P_{it} = 1|X_{Pit-1})} \tag{3}$$

$$IMR_{it-1} = \frac{PDF\left(ICDF(PP)\right)}{PP}$$

PDF is the probability density function and *ICDF* is the inverse cumulative density function of the standard normal distribution.[17] The IMR_{it-1} is then included in the second stage regression estimating the probability of default, which is illustrated in Equation (4).

Stage 2

$$Prob(D_{it} = 1|X_{2it-1}) = \frac{1}{1 + e^{-\left(\alpha_D + \beta_D X_{Dit-1} + \gamma_D X_{Dit-1}^{Transactional} + \delta_D IMR_{it-1}\right)}} \tag{4}$$

Where D_{it} is a binary variable which takes the value of one if the loan is 90 days in arrears within the next year or zero otherwise. X_{Dit-1} is a vector of lagged observable mortgage

[17] Following the literature, we use the standard normal distribution for IMR. Note that simulation studies show consistent results for PDs that are based on logit and probit models. Results are available on request.

information and $X_{Dit-1}^{Transactional}$ includes lagged transactional information. The baseline model includes only non-transactional information and IMR_{it-1}, with the new transactional data being added as test variables to subsequent models.

The maximum likelihood estimator for Stage 2 is represented in Equation 5:

$$L = \prod_{i=1}^{I}\prod_{t=1}^{T} \mathrm{Prob}\left(D_{it} = 1 \middle| X_{Dit-1}, X_{Dit-1}^{Transactional}, IMR_{it-1}\right)^{D_{ti}} \left(1 - \mathrm{Prob}\left(D_{it} = 1 \middle| X_{Dit-1}, X_{Dit-1}^{Transactional}, IMR_{it-1}\right)\right)^{1-D_{it}} \tag{5}$$

Due to the nature of transactional data, many missing values are present. Missing values are observed when there are no transactions present in a category for a measurement period. As an example, if a borrower has no measured cash withdrawals in a period, the cash withdrawal variable will be missing. We cannot conclude from a lack of observation that the borrower did not make a purchase or receive income as alternative arrangements may exist, such as other banking relation to fill this need. From Figure 2, we can see that these observations do not align with the observed portion of data. These observations are distinct from the rest of the series, suggesting a missing value itself can contain specific information distinct from non-missing. To consider this, we create the following spline function for our transactional variables:

$$\gamma X_{Dit-1}^{Transactional} = \begin{cases} \gamma_1 X_{Dit-1}^{Transactional} \ \ if \ X_{Dit-1}^{Transactional} is \ observed \\ \gamma_2 \ \ if \ X_{Dit-1}^{Transactional} \ is \ not \ observed \end{cases} \tag{6}$$

Using this transformation in addition to the variable effect allows for the inclusion of associated observations as missing values are treated while keeping possibly distinct relations separated. Therefore, if a value is missing, the additional dummy variable is set to one, and when the value is not missing, the associated dummy variable is set to zero. This allows for the use of one coefficient if the value is present and another coefficient when the value is missing.

3.4.2 Model fit measures

To highlight the strength and model improvement of the new transactional factors, we modify existing model fit measures to measure marginal effects using Pseudo R^2 and $AUROC$.

Pseudo R^2 calculation

Equation (7) outlines the calculation of Pseudo R^2. Following Cox and Snell (1989), we calculate Pseudo R^2 by comparing the likelihood of an intercept only ($L(0)$) model to the likelihood of a specified model ($L(\beta)$). This is rescaled to the interval $[0,1]$ by comparing the model with the maximum R^2 possible, which we label R^2_{max} (Equation 8). We then rescale R^2 with R^2_{max} to calculate $\tilde{R}^2$ (Equation 9).

$$R^2 = 1 - \left\{\frac{L(0)}{L(\beta)}\right\}^{\frac{2}{n}} \tag{7}$$

$$R^2_{max} = 1 - \{L(0)\}^{\frac{2}{n}} \tag{8}$$

$$\tilde{R}^2 = \frac{R^2}{R^2_{max}} \tag{9}$$

***AUROC* calculation**

The Area Under the Receiver Operating Characteristic $(AUROC)$ curve is a common measure used to compare the discriminatory power of competing binary classifier models. The intuition is to calculate the model discrimination at various thresholds and plot the true positive rate (sensitivity) against the true negative rate (specificity). The sensitivity and specificity pairs are then plotted on a curve and the area under the curve is calculated.

Measure transformation

We propose two novel transformations of our fit. The first compares the increase in the model fit using the improvable area. Intuitively, it becomes more difficult to quantify the improvement due to a diminishing marginal benefit of each subsequent improvement. The second measure calculates the increase in the measure relative to the Base model. These two new measures produce results in the same direction with the magnitude allowing for slightly different conclusions.

The first measure is based on the improvable area. The improvable area is the region between the base estimate and the maximum value. For $\tilde{R}^2$ and $AUROC$, the maximum value is one. The idea is to measure the decrease in the improvable area caused by the improved model. Improvable Reduction ($IMPRED$) is calculated as follows:

$$IMPRED = \frac{M_2 - M_1}{1 - M_1} \tag{10}$$

In Equation (10), the numerator measures the absolute improvement of a measure, and the denominator measures the maximum possible improvement. As an example, if $IMPRED$ is calculated to be 0.5, M_2 improves upon M_1 in such a way as to remove 50% of the improvable area.

The second measure determines the magnitude of the model improvement compared to the Base model, allowing for accurate conclusions based on measure increases or decreases. The Improvement Magnitude ($IMPMAG$) is calculated as follows:

$$IMPMAG = \frac{M_2 - M_1}{M_1} \tag{11}$$

In Equation (11), the numerator again measures the absolute improvement of the second model over the first, and in this case the denominator scales this improvement by the original fit. This measure is not bound in [0,1] and is free to give the relative improvement in percentage terms. As an example, an $IMPMAG$ of 1.25 represents a 125% improvement over the Base model measured as M_1.

$IMPRED$ and $IMPMAG$ measure different aspects. One is concerned with the improvable area only, while the other is concerned with the relative magnitude only. This important distinction allows for conclusions at both ends of the model fit spectrum. For more powerful models, the reduced improvable area can highlight the marginal benefit of improved models when the magnitude is small. On the other hand, for weaker models, a small improvement in model fit can represent a relatively large improvement in the comparative power between models.

In this paper, we apply these new measures to the two model fit measures, namely $\tilde{R}^2$ and $AUROC$. Note that all performance measures should only be used for model comparison based on identical data sets.

3.5. Results

3.5.1 Base model

In this section we estimate and present the Base model, which includes estimation of the first and second stage regressions excluding any of the new transactional variables. This is to establish a baseline with which to assess the additional benefit of the new transactional variables. The Base model results are reported in Table 5.

In the first stage payoff regression we include a binary transformation for fixed interest (FIXEDINT) and interest only (INTONLY). Numeric variables include the dynamic LTV (DLTV), the number of advanced payments by the borrower (ADVPAYMENTS), the tenure (TENURE), the log of the property value (LNPROP), the interest margin (INTMARGIN), the credit card utilization rate (CCUTIL), the credit bureau score (BUREAU), the change in HPI since inception (DHPI), the target interest rate (TARGETINT) and the change in the target interest rate since inception (DTARGETINT). The median income of the postcode (LNINCOME) and the mean payoff rate of the previous period (MEANPAYOFF). These variables hold a causal relation to payoff, and in many cases a directly opposite relation to default. Hence, in these cases the variables are removed from the second stage.

The variables represented by X_{Dit-1} include the borrower loan to income ratio at loan inception (APPLOANINCOME), a low documentation flag (LOW_DOC) equal to one if it is a low-doc loan and zero otherwise. The FIXEDINT flag is included along with the natural logarithm

Table 5: Base regressions

This table depicts the regression results for the first stage payoff and second stage default Base model to which the new models will be compared. Both models are constructed with widely accepted drivers of mortgage payoff and default. Estimate contains the coefficients derived from estimating a logit model on a binary variable equal to one if the loan is paid off or defaults within the next 12 months and zero otherwise for the two models respectively. Chi-square statistics are in parenthesis. All regressions include standard errors clustered at the state level.

Variable	Stage 1 - Payoff		Stage 2 - Default	
INTERCEPT	-7.7759	***	-2.0539	
	(308.72)		(2.02)	
APPLOANINCOME			0.00134	***
			(80.13)	
INTONLY	0.2503	***		
	(194.28)			
LOWDOC			0.5941	***
			(112.64)	
FIXEDINT	-0.8491	***	-0.2096	***
	(134.05)		(9.22)	
DLTV	-0.1834	***		
	(95.53)			
ADVPAYMENTS	0.0059	***		
	(541.08)			
TENURE	-0.0065	***		
	(159.17)			
LNOUTSBAL			0.7067	***
			(57.13)	
LNPROP	-0.1188	***	-0.5486	***
	(109.8)		(65.09)	
INTMARGIN	0.4178	***	0.9416	***
	(450.23)		(92.07)	
CLTV			0.0129	***
			(39.96)	
CCUTIL	0.0027	***	0.0099	***
	(110.75)		(257.45)	
BUREAU	-0.0006	***	-0.0014	***
	(35.65)		(95.7)	
DHPI	0.8678	***		
	(162.24)			
TARGETINT	0.1363	***		
	(24.27)			
DTARGETINT	-0.2813	***		
	(84.21)			
LNINCOME	0.2562	***	-0.8002	***
	(67.78)		(47.23)	
MEANPAYOFF	12.2836	***		
	(71.08)			
IMR			-0.4940	***
			(19.03)	
AUROC	0.6092		0.7426	
Pseudo R-Square	0.0173		0.0697	
Observations	6,568,224		6,568,224	

of the current outstanding balance (LNOUTSBAL). LNPROP and INTMARGIN are included along with the current loan to value ratio (CLTV). Finally, CCUTIL, BUREAU and LNINCOME are included.

Table 5 reports the regression estimates of our Base model with errors clustered at the state level. The results of the Stage 1 model are as expected. Borrowers with interest only loan structures (INTONLY) are more likely to pay off or refinance, in line with a more investing and profit seeking attitude. Fixed interest loans (FIXEDINT) have a lower probability of payoff due to tenure based contractual arrangements including penalties for early exit. A greater reduction in dynamic LTV (DLTV) since inception indicates a lower probability to pay off as these borrowers remain longer with the bank; this is in line with greater TENURE resulting in a lower probability to pay off. The greater the target interest rate (TARGETINT) and the more negative the change (DTARGETINT), the greater the probability of payoff. In our sample, there are reductions in interest rate, giving opportunities for cost savings through refinance. A positive change in house price index (DHPI) increases the payoff probability as increasing house prices are conducive to refinancing. As expected, the lagged payoff rate is positively correlated with the current payoff rate.

For the Base model in Stage 2, the signs for the regression variables match our prior expectations. A higher application loan to income ratio (APPLOANINCOME) increases probability of default. Low-doc (LOWDOC) loans reflect a higher default risk with the coefficient having a positive sign. Fixed interest loans have lower risk compared to variable rate loans as interest rate risk is not a factor. The greater the outstanding balance (LNOUTSBAL) and current loan to value ratio (CLTV), the greater the risk of default. The higher the property purchase price (LNPROP), the lower the risk of default, suggesting significant wealth effects through the interaction of property price and LTV. In Australia, the steadily increasing price of inner-city dwellings allows for reliable and simple refinancing when there is a liquidity shortfall. Consistent with other studies, greater interest margins (INTMARGIN) lead to higher default rates. Interest margin is calculated as the borrower interest rate minus the target interest rate and is risk adjusted for each borrower, leading to this clear effect. High credit card utilization (CCUTIL) leads to higher risk, and credit card default tends to preempt mortgage default according to internal information from the bank. Using other sources of liquidity prior to missing mortgage payments is common. A higher bureau score (BUREAU) and higher income in the geographical area correlate with lower default.

The AUROC of 0.7426 and Pseudo R-Square of 0.0697 represent the fit of our Base model. This will be our main comparator when testing new model variables and the main input into the new fit measures that we develop. An increase in this value represents a better fit, with the difference in AUROC being tested for significance using a statistical test for ROC equality (DeLong *et al.* 1988).

3.5.2. Identification of transactional variables

The variables used consist of raw measures as well as transformed and scaled measures to capture differences across wealth and spend related borrower behavior. We add these variables to the baseline Stage 2 regression and calculate the AUROC for the new model and perform out-of-time and out-of-sample tests to ensure the robustness of the model.

Considering the depth and breadth of the data available in this study, there are countless factors which can be generated and tested. We have outlined the approach to identify key test variables in Appendix 3.

The final variables we select to be included are short term interest coverage ratio (INCOME, i.e., average income in the last six months to interest), salary change (SALARYD), home maintenance expenses to property value (HOMEMAINT), six-month average cash out to total spend (CASH), and non-durable discretionary change (NONDURDUSCD) in the last year. In the following section we investigate the presence of these consumption categories in the literature to motivate our selected variables.

Short-term interest coverage ratio

Borrower income is collected at loan origination, with subsequent updates not incorporated into portfolio monitoring activities leading to reduced accuracy over time. One solution is to capture actual income or salary credits deposited into a borrower's bank account. This can produce an accurate and timely measure of a borrower's current ability to service the loan.

The bank data includes a flag to distinguish between income deposits from employment, government benefits, pensions, and other non-income related transactions. This allows for the creation of a measure which captures the income deposited into the account of a borrower in real time. To incorporate this into the current study we create the following measure:

$$INCOME = \ln\left(\frac{Income_{6m}}{MonthlyInterest_{6m}}\right) \tag{12}$$

The above measure captures the average monthly income received by borrowers in the last six months and scaled by the average monthly interest charge on the loan over the same period. Although a reduction in income can cause financial stress regardless of interest charged, intuitively a greater interest liability will cause greater stress. The natural logarithm transformation is suitable to account for the non-linearity of the relation between default and income.

The impact of income at loan origination on default is well studied, with higher income leading to lower probability of default (Campbell & Dietrich 1983; Riddiough 1992). Income is one of the main variables collected at loan origination used to calculate loan serviceability. Our contribution to the literature in this area is to include a more current and timely measure of income.

Change in income

A powerful feature of our data is the ability to measure the change in income received across periods. A change in income represents a change in financial circumstances, which requires a suitable adjustment of spending behavior or leads to default otherwise.

We find change in salary, which is the main component of income, is a stronger predictor compared to a change in income. This is due to the borrower’s salary payment being riskier than other forms of income such as pension payments or welfare. Therefore, we use salary to measure the change in income. This measure allows for income changes to be separated from income levels, representing a different angle and story. Our change in salary measure is defined as follows:

$$SALARYD = \frac{Salary_{3m}}{Salary_{12m}} \quad (13)$$

Salary change is calculated as the average monthly salary received in the last three months divided by the average monthly salary received in the last 12 months. A decrease in income in the form of job loss has been found to be a significant driver of default (Gerardi *et al.* 2017). Our result indicates both a reduction and an increase in income can increase the risk of default. We conclude that a simple change in financial circumstances can be a powerful driver of default, with an increase in income potentially leading to a disproportionate increase in spending.

Cash withdrawal

An interesting result of this study is the impact of cash withdrawals on observed default. This factor is both highly predictive and positively correlated to default. As the average amount of cash withdrawal increases, so does the default risk. The selected factor is constructed as follows:

$$CASH = ln\left(\frac{CashWithdrawal_{6m}}{TotalSpend_{6m}}\right) \quad (14)$$

Cash expenses are more tangible when compared to other purchase pathways, with its use in decline in recent years. Prior research suggests cash is mainly used for minor transitory purchases such as food and beverages, restaurants, and newspapers or lotteries (Bounie & François 2006; Arango *et al.* 2015; David *et al.* 2016). A significant number of cash users switch to debit and credit cards for amounts marginally larger than \$5 and \$10 (Chen *et al.* 2019) and discounts for cash payments increase the probability of a consumer switching from their preferred payment method (Stavins 2018). Amromin and Chakravorti (2009) find as debit card use increases, the need for smaller cash denominations decreases, and the need for larger denominations increases. This suggests that the use of cash as a storage of value has become more prominent compared with its transactional value.

We show that at relatively low levels of average withdrawal, there is no significant increase in future default, however this increases monotonically after a certain fixed point as is indicated in Figure 2. This suggests a certain level of cash withdrawal and expenditure does not impact the risk of default.

Clothing to monthly interest

We use clothing expenses as an example and proxy for durable discretionary spend. Clothing is usually categorized along with food and local travel expenses as non-discretionary. However, we posit in advanced economies, and especially our sample, borrowers have abundant clothing to meet the requirement for warmth and social acceptance prior to origination. When financial constraints become more binding, it is simpler to reduce clothing spend than other non-discretionary categories due to its durable nature[18].

This relation between discretionary durable expenditure (excluding home specific expenses such as maintenance) and financial distress explains the strength of clothing expenditure as a predictor of default. For any given month, a borrower needs to balance the need to pay their mortgage with their other expenses, and as such we use the monthly interest as a proxy for the minimum payment required in servicing the loan. Our clothing expense factor is constructed as follows:

$$CLOTHING = ln\left(\frac{ClothingExp_{6m}}{MonthlyInterest_{6m}}\right) \tag{15}$$

Where $\mathrm{ClothingExp}_{6m}$ represents the average clothing expenditure for the last six months and $\mathrm{MonthlyInterest}_{6m}$ represents the average monthly interest charges for the last six months.

When financially constrained, individuals are more likely to reduce durable expenditures compared to non-durable as prior durable purchases (particularly clothing) can be used as a buffer against transitory negative shocks to income (Browning & Crossley 2008). Overall, clothing

[18] For example, groceries are likely to be purchased before clothing under financial constraints.

expenditures as a proxy for other non-essential durable expenditures is an effective addition to our model as it is affected by a household's financial circumstances.

Home maintenance

To calculate the total home maintenance, we collect all credit and debit card expenses relating to home maintenance and furniture. We find this variable is highly predictive when scaled by property value as the maintenance a property requires is a function of its size and value. Note, debit and credit card expenses do not include expenses, which were paid by direct transfer or by checks. The measure is calculated as follows:

$$HOMEMAINT = LN\left(\frac{HomeMaint_{6m}}{PropertyValue}\right), \tag{16}$$

Where $HomeMaint_{6m}$ represents the average monthly dollar value spent on home maintenance in the last six months and $PropertyValue$ is the property value collected at loan origination.

The effect of home maintenance on mortgage default is not well studied. Choi *et al.* (2014) study owner speculation of future house prices through home improvement and find that the recoup value of home improvements simultaneously increases with home price appreciation and falls with the growth of construction costs. Furthermore, Montgomery (1992) develops an analytical framework and demonstrates how home improvements might fit into a model of household investments, but does not link this to mortgage default. Finally Helms (2003) studies the impact of building and neighborhood characteristics on the decision to renovate. Although home renovation has been studied in some respects, its link to default is yet to be determined; our study introduces and tests the relation.

When a home is purchased, required home maintenance may be rationally included or excluded from the purchase price; therefore, home maintenance is a real option. Since the expense is not experienced at purchase, a borrower can later decide to undertake the expense based on financial circumstances. From our preliminary results, we see this is indeed the case, with maintenance indicating a favorable financial position for the borrower.

Change in non-durable discretionary expenses

We include the change in non-durable discretionary expenses as a final new factor to capture information not already captured through other factors. Transaction categories such as restaurants, gambling, alcohol, travel and financial services are included.

The nature of non-durable discretionary expenses suggests this category can be reduced in the presence of financial constraints. This explains why there is a significant increase in credit risk when spending in this category reduces (see Figure 2). We calculate the change in non-durable discretionary expenses as follows:

$$NONDURDISC = \frac{Nondurable\ discretionary_{3m}}{Nondurable\ discretionary_{12m}} \tag{17}$$

Where $Nondurable\ discretionary_{3m}$ represents the average non-durable discretionary expenditure for the last three months and $Nondurable\ discretionary_{12m}$ represents the non-durable discretionary expenditure for the last 12 months.

The relation of this factor to default has not been studied to our knowledge. During financial stress, durable expenditure can be reduced to allow for more stable non-durable expenditure (Browning & Crossley 2009). Browning and Crossley (2009) use clothing and food expenses to proxy for durable and non-durable expenses respectively. In our paper we further split the analysis into discretionary and non-discretionary spending.

3.5.3 Factor transform and relation to default

We now investigate the general univariate relations present between our transactional factors and default. To do this, we sort our observations into 400 bins, ranking the observations by the factor. Binning is required as default is a binary variable and default rates for bins are more robust than binary default events.

Figure 2 depicts the relation between default and factors. For all plots the y-axis contains the average default rate for each of the bins, while the x-axis contains the factor value. A declining cluster indicates that as the factor increases, the default rate decreases and vice versa. The two main relations suggested are linear and quadratic for the moving average and change variables respectively.

Panel 1 (top left chart) represents the short-term interest coverage ratio. Although interest coverage ratio is a common measure in credit risk models, an almost "live" income feed from borrower bank accounts has not been studied. As expected, there is a negative relation between income and risk of default. The relation changes slightly at the fringes, the ratio above three indicating more income does not affect the default rate, and the increased variance below zero indicating that for some loans there may be income not captured by the bank. Overall, although there may be uncaptured income, the default rate for these borrowers is still high.

This result is intuitively similar to other default risk studies regarding income level, however, this measure is fundamentally different due to the inclusion of income recently received by the borrower. We expect this to be a stronger predictor of default compared to the salary measured at origination, which can change throughout the loan period.

Panel 2 (top right chart) reports the effect of a change in salary on the probability of default. The clustering around one indicates risk is lowest for borrowers who have had no significant change in income over the last 12 months. As borrowers experience a negative change in income, there is a clear and linear increase in the default rate. Interestingly, it appears an increase in income also results in an increase in default risk, leading to the conclusion that a simple change in circumstances can have a large effect on default risk. This is in contrast with an increase in salary allowing for greater ease of mortgage payments. It may be that borrowers overestimate future wealth and adjust spending accordingly, resulting in a reduced capacity to service the loan. The observation at zero represents borrowers for whom we do not have a salary payment within the last 12 months, with the lower observed default rate suggesting alternate income sources or income being paid into a different bank.

Subsequent change in borrower income is a powerful indicator of default risk generally not captured by financial institutions. Furthermore, other studies that use proxies for job loss include only total job loss, and not income shrinkage, which is an important driver of serviceability. This measure is more accurate than regional measures of the employment rate as it captures change in income on a borrower level.

Panel 3 contains ATM withdrawals as a fraction of total spend. This factor represents a novel measure of risk, with higher withdrawals representing higher risk. Although we cannot ascertain the final use of these funds, intuition suggests goods or services bought with cash may not be essential. Generally, cash is used on smaller transitory expenses such as snacks, entertainment or potentially gambling. In many countries, including Australia, it is illegal for gaming machines to accept card payments, requiring patrons to first withdraw cash. This factor and its presence in prior literature is discussed in Section 5.

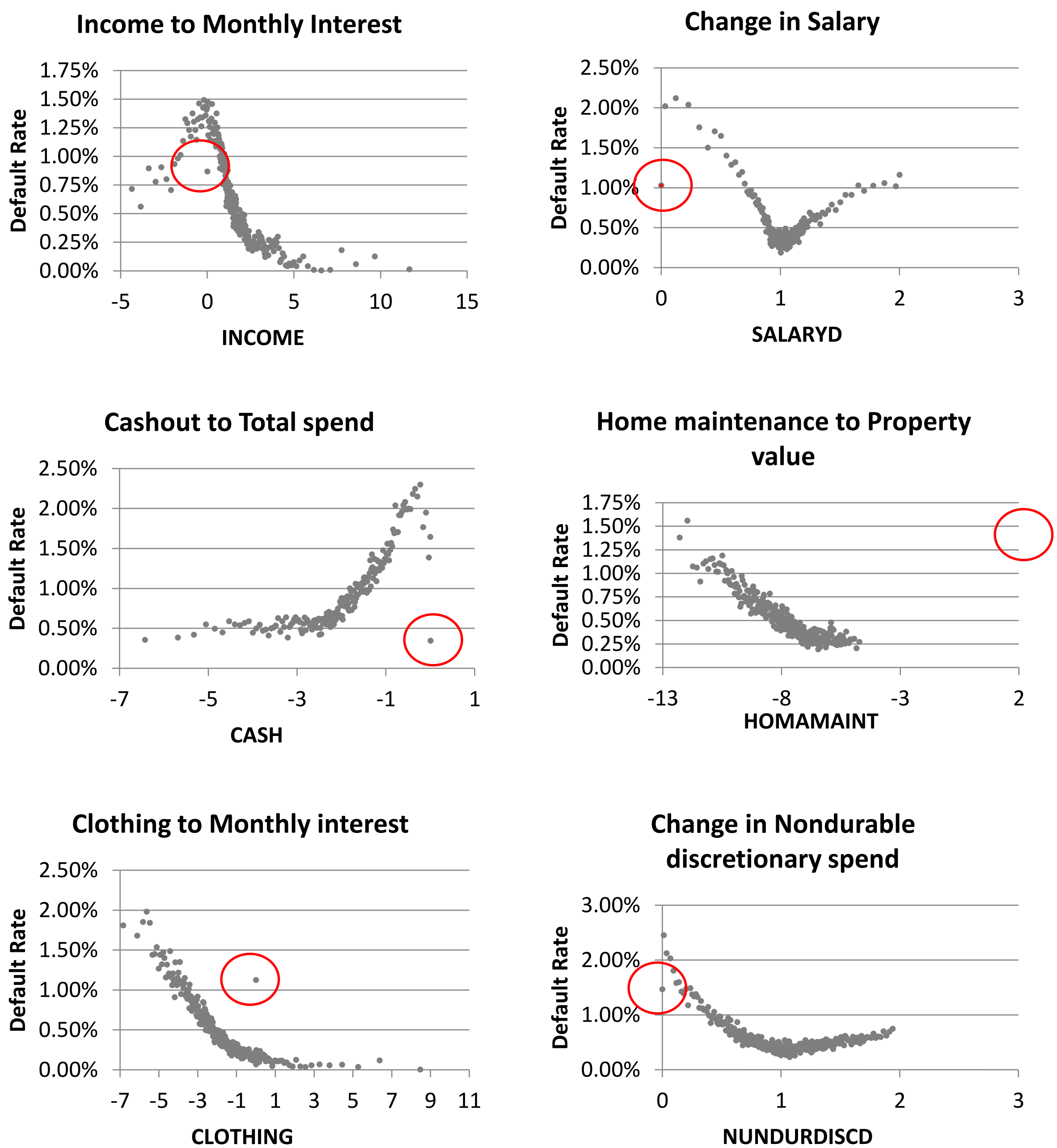
The relation between cash withdrawal and default is interesting. At the lower levels, the default rate does not increase, however the slope rapidly rises as the proportion of withdrawals rises, suggesting there is an acceptable level of withdrawal with no impact on credit risk, however as this increases, so too does risk.

The relation between default and the natural logarithm of average monthly home maintenance spend to property value for the last six months is depicted in panel 4. There is a clear linear and negative relation between this factor and default. Note that only credit and debit card maintenance expenses are captured as only these can be mapped to the respective expenditure codes. We do not measure other payment types including checks and direct transfers to maintenance or renovation firms. We do believe however, that our captured expenses act as a proxy for total expenses.

A quadratic relation is also present for our nondurable discretionary change variable, suggesting a change in behavior can cause an increase in credit risk as seen in Panel 6. This is clear as the risk of default increases above and below the no-change point one. The smooth curving relation shows low variance in the factor.

Figure 2: Univariate relation between transactional variables and default rates

This group of figures illustrates the general relations that are present in the data for different variable types. CASH and HOMEMAINT are indicative of the relation between the moving average transformation and observed default within the next 12 months, the linear relation is modeled using simple linear parameters. The SALARYD and NONDURDISCD depict the general relation held by change based factor construction. These are modeled by using a quadratic relation due to the increase in observed default around one. CLOTHING and INCOME are simple moving averages in line with CASH and HOMEMAINT. The plots are constructed by sorting observations into 400 bins and measuring the observed default rate for each of the bins, each dot represents one bin. Grey circles bring attention to the bin containing missing values where the factor is set to one. For HOMEMAINT, the missing value bin is to the right of the chart, with the grey circle set at the level of the bin (the circle indicates the level, but the point is not displayed). A missing occurs when there is no observation present to create the factor.



Overall, the transformation of the underlying factor creates the need for different parameters to accurately model the factor relation to mortgage default. For simple moving average factors, a natural logarithm transform is used, and for changes in factors, a quadratic spline is used to capture the nonlinearity. Many other nonlinear spline and structures were tested, with the results remaining materially similar. We opt for these simpler structures for a better transparency.

3.5.4 Full model

We look more closely at the additional benefit of our transaction-based variables in standard probability of default models. We now augment the Base model with our new variables. Standard errors are clustered at the state level.

Table 6 contains the estimates of the augmented models. Model 1 adds the short-term interest coverage ratio to the regression. The augmented model is estimated by including a binary variable to capture missing values, measuring the underlying effect of either not having a transaction, or having a transaction through an uncaptured account (the spline term is indicated by *P*). As expected, higher short-term interest coverage leads to lower default risk, as indicated by the negative and significant coefficient. There is an increase in the model AUROC over the original model of 0.0123, suggesting that including a short-term income measure is an effective addition to the Base model, even though a measure for the original application income already exists.

As expressed earlier, income at application is captured at mortgage origination and is not updated to reflect changes after this point; therefore, any information a bank can capture regarding subsequent salary changes would be beneficial for capturing the current serviceability of any client. Although a bank cannot necessarily change the interest rate a specific borrower is charged, it is able to make decisions regarding its mortgage portfolio more accurately if this information is available.

Model 2 includes a quadratic spline to capture the parabolic relation between change in salary and probability of default. The coefficients are in line with the scatter plot depicted in Figure 2. A reduction and an increase in salary received result both in increased default risk. The addition here gives a sharp rise in the AUROC measure to a value of 0.7706.

Model 3 includes CASH, which is total borrower cash withdrawals divided by their total spend on both credit and debit cards. The positive coefficient indicates borrowers who withdraw more cash are more likely to default. The AUROC of 0.7745 suggests this model is more informative. The negative coefficient on CASHP indicates a missing observation is less risky compared to a high proportion of withdrawals to total spend.

The result of adding CLOTHING is depicted in Model 4. The negative coefficient means greater clothing expenses are an indication of a stronger financial position. Risk for borrowers who have no measured expense in this category is increased as measured by the positive coefficient on CLOTHINGP. The additional parameter creates an increase in the AUROC measure of 0.0378 over the Base model to 0.7804.

Model 5 includes the effect of home maintenance expense (HOMEMAINT) on the probability of default. The highly significant negative coefficient of -0.2377 implies a greater maintenance expense in the previous six months leads to a reduced risk of default in the next 12 months. The AUROC measure increases by 0.0290 over the Base model to 0.7716, showing a strong effect in separating defaulting from non-defaulting loans. The HOMEMAINTP variable allows for the borrowers who have no observed maintenance spend to be treated differently to those that do. Consistent with what we can see in Figure 2, the probability of default for borrowers who have zero spend on home maintenance is in line with borrowers who spend very little. Therefore, as borrowers spend more on maintaining their home, their probability of default within the next 12 months reduces.

Model 6 includes change in nondurable discretionary spending as the additional factor. We use a quadratic spline to accurately transform the relation depicted in Figure 2. Consistent with Figure 2, the parabolic nature is adequately captured by the quadratic spline, with the model improvement of 0.0222 over the Base model.

Table 6: Full regressions

Panel A shows results of regressions with the new variables. Transaction variables include a binary variable equal to one if there are observations available; for example, HOMEMAINTP is equal to one if there are no observations available for HOMEMAINT. This reduces the bias for borrowers who have no observations available with this bank. CONTROL indicates the new variables are added to the Base model. Panel B contains the new model fit measures. AUROC refers to the Area Under Receiver Operating Characteristic curve. *** and ** indicate significance at the 1% and 5% level respectively. Chi-square statistics are in parenthesis. Standard errors are clustered at the state level.

Panel A – Model results

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Full model	
INCOME	-0.1782	***											-0.0988	***	0.0012	
	(543.17)												(53.12)		(0.01)	
INCOMEP	0.3845	***											0.0507		-0.1068	***
	(100.80)												(2.53)		(6.95)	
SALARYD			-2.7739	***									-2.3001	***	-2.5831	***
			(2499.53)										(1331.73)		(768.35)	
SALARYD2			1.2391	***									1.0401	***	1.1471	***
			(2535.82)										(1510.27)		(986.21)	
SALARYDP			-0.8098	***									-0.8372	***	-0.7242	***
			(1241.62)										(826.80)		(838.16)	
CASH					0.4700	***							0.1193	***	0.1730	***
					(92.30)								(9.56)		(15.05)	
CASHP					-1.7075	***							-1.2181	***	-1.2980	***
					(212.44)								(216.06)		(144.97)	
CLOTHING							-0.3320	***					-0.3048	***	-0.2097	***
							(1340.44)						(370.31)		(419.04)	
CLOTHINGP							1.8287	***					1.3624	***	1.1174	***
							(1067.71)						(368.99)		(335.11)	
HOMEMAINT									-0.2377	***			-0.1110	***	-0.1323	***
									(475.35)				(60.06)		(125.77)	
HOMEMAINTP									2.8569	***			1.3504	***	1.5887	***
									(689.01)				(79.87)		(129.81)	
NONDURDISCD											-2.5648	***	-1.2592	***	-1.2226	***
											(2768.49)		(764.59)		(397.39)	
NONDURDISCD2											1.0440	***	0.4656	***	0.4542	***
											(3128.03)		(333.86)		(252.32)	
NONDURDISCDP											-0.6110	***	-0.9583	***	-0.6635	***
											(75.44)		(356.18)		(146.64)	
CONTROL	YES		YES		YES		YES		YES		YES		NO		YES	
AUROC	0.7549		0.7706		0.7745		0.7804		0.7716		0.7648		0.7850		0.8300	
AUROC - BASE	0.0123	***	0.028	***	0.0319	***	0.0378	***	0.029	***	0.0222	***	0.0424	***	0.0874	***
Pseudo R-Square	0.0782		0.0918		0.0943		0.0986		0.0911		0.0891		0.1000		0.1483	
Observations	6,568,22		6,568,224		6,568,224		6,568,224		6,568,224		6,568,224		6,568,224		6,568,224	

Table 6: In sample continued

Panel B - Model improvement measures

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Full model**
AUROC – IMPRED	0.05	0.11	0.12	0.15	0.11	0.09	0.16	0.34
AUROC – IMPMAG	0.02	0.04	0.04	0.05	0.04	0.03	0.06	0.12
R-Square – IMPRED	0.01	0.02	0.03	0.03	0.02	0.02	0.03	0.08
R-Square – IMPMAG	0.12	0.32	0.35	0.41	0.31	0.28	0.43	1.13

To test the predictive accuracy of the new factors without the Base model variables, we include Model 7. Model 7 only contains the new variables and results in an AUROC of 0.7850, which is higher than the Base model. This result shows the new factors are not simply an addition to an already robust framework but can constitute a robust framework itself. All the signs across the regressions are the same, representing robust relations between the new factors and default. Overall, these new behavioral and transaction-based variables are at least equivalent to other well-studied factors commonly used in credit risk models.

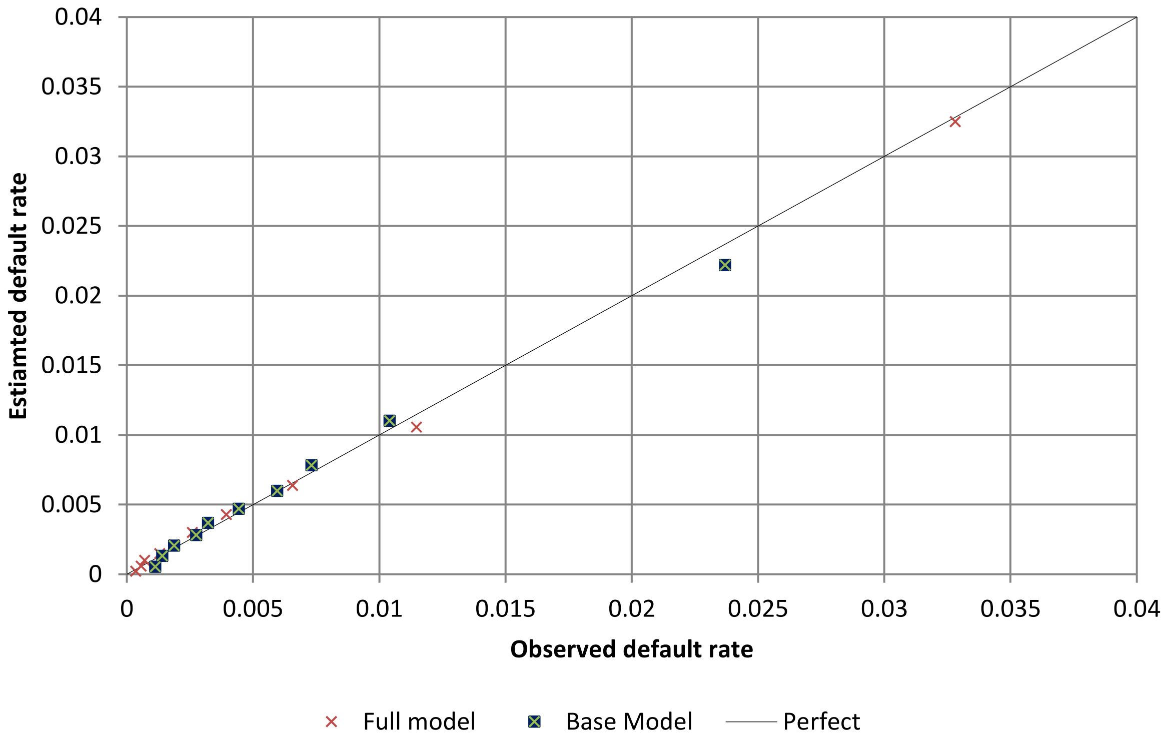
Finally, the Full model includes all transaction variables. This model increases the AUROC from the Base model by 0.0874 to 0.8300. Interestingly, the coefficient on INCOME is not significant, suggesting that income at origination, along with a change in income measure, adequately captures the current income level. The changes in income factors (SALARY and SALARYD) capture the change from the original application income and continues to be highly significant.

Our new model improvement measures are tabulated in panel B of Table 6. The first row containing the IMPRED calculation of AUROC suggests a large decrease in the improvable area. An IMPRED of 0.34 for the Full model indicates a reduction in the improvable area of 34%, and an IMPMAG shows that the Full model is 12% more powerful than the Base model. A similar conclusion is drawn from the increase in the Pseudo R-Square; the IMPRED shows that there is reduction in the improvable area of 8%, and the IMPMAG shows a magnitude increase of 113% over the Base model. This final IMPMAG shows that the Pseudo R-Square is over double that of the Base model.

Figure 3 compares the observed default and estimated probability of default for the sample segmented into deciles sorted by the Base model probability of default. Both models track the optimal line where the observed and estimated probabilities are identical (diagonal line). This indicates that the model is well calibrated to capture the default rates in the sample.

Figure 3: Real-fit diagrams for default probabilities

This figure illustrates the comparative strength of the Base and Full model. The data is split into deciles by estimated default rate. The bottom axis represents the observed default rate for each decile, and the side axis represents the estimated default rate. The Base model refers to the second stage regression including all base variables, and the Full model refers to the model including both base and transactional variables. The grey line represents the relation under a perfect model where the estimated and observed default rates in each decile are identical.



The greater dispersion by the Full model indicates greater effectiveness of separating higher risk loans from the rest of the population. The upper two right most markers show the estimated and observed risk of the highest decile, with the Full model marker being above the Base model marker. The distance of the riskiest Full model decile from the most risky Base model decile demonstrates that the Full model is better at separating truly risky loans from the rest of the population.

In conclusion, these additional transaction-based variables can be used to increase the model effectiveness by adding information not captured by more traditional credit default variables, and the Full model is superior to the Base model.

3.5.5 Timeliness of estimates

One aspect of transactional data not yet explored is its usefulness in determining changes in borrower risk over time. In this section we further contrast our Base model with our Full model, specifically investigating how risk estimates change as a borrower approaches default. Probability of default models built using application data suffer from stale information as these variables do not account for borrower circumstantial changes. The transactional data can capture patterns and changes in borrower risk through spending patterns.

The data in this section is limited to borrowers who eventually default, all observations for these borrowers are kept regardless of their status at any point in time. We calculate the number of months until default and display this on the x-axis, with the average probability of default estimate for each respective model displayed on the y-axis. We produce three sets of models changing only the default horizon, specifically, our default flag is calculated with lags of three months, 12 months (identical to the Full model, previous results), and 24 months. The purpose of this is to showcase the effectiveness of the transactional data in predicting default over different time horizons.

Figure 4: Average default rate by month to default

This figure compares the use of different modelled default lags on the estimated probability of default for loans approaching default. Default is at t = 0, and negative t represents months to default. The solid lines represent the average probability estimate for the Full model and the dashed lines represent the Base model. The Full model estimates are higher than the Base model estimates, particularly as default becomes imminent.

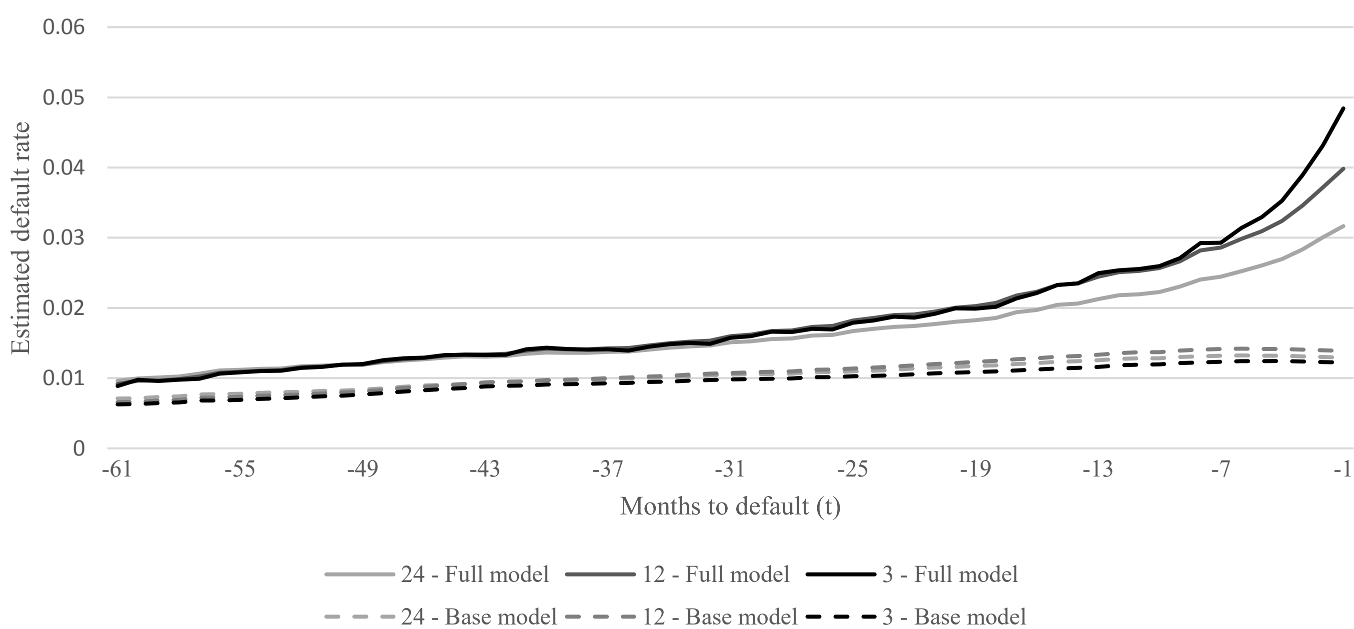


Figure 4 contains a graphical depiction of the results. The main finding is that the mean predicted default probabilities based on transactional data rise sharply as borrowers approach default, and this suggests the Full model can identify increasing risk 60 months before default. This relation is also present for the Base model parameterizations, however there is no sharp increase as default approaches, despite including credit card utilization, macroeconomic variables and current loan derived variables are included in the Base model.

3.5.6 Scoring at origination

Scorecards used at origination make a first assessment of the riskiness of a potential loan. In this section, we use the score generated by the Base and Full model at origination to assess different loan origination outcomes to indicate how loan decisions based on transaction data differs to more traditional models. To do this, we filter our sample for loans where the origination date is the same as the scoring date, resulting in 62,905 mortgages. We then measure the number of default incidents by the end of our sample period[19]. The Base model under this circumstance uses up to date origination factors, with the Full model supplementing behavioral factors. We use the risk estimates from the current models, and do not re-estimate the parameters.

Using the sample of originated loans, we measure the number of loans which would have been originated using different estimated default risk cutoffs. For example, a bank may insist that it will not originate any loans for which the estimated one-year default risk is above 1%. The results are depicted in Table 7.

Table 7 compares the marginal difference between the Base and the Full model. Using 1% as the barrier to origination, 1,031 fewer loans are originated using the Full model, however 66 of these loans default within our sample period, leading to a 6% default rate for the marginal loans. A stylized calculation for the data shows the transactional model results in a cost saving of approximately $18 million[20] for our small sample. Further benefits are the avoidance of adverse

[19] Although many loans may default after the end of the sample period, since both models are using the same period we can assume measured relationships to continue.

[20] The assumptions are: an average life on books of 5 years (as loans are often refinanced), an average loan amount of $500,000, a profit margin of 0.5% and a loss rate of 20% (the Australian LGD floor imposed by the regulator for mortgages), and a default at loan start. 500,000*1,031*(0.2*0.06-0.005*0.94*5) = 18,042,500.

selection in the origination process and lower loan loss provisioning as well as regulatory capital allocation.

Table 7: Model power for origination

This table depicts how the Base and Full model impact the decision to originate a mortgage, using a sample constructed where the origination month is the same as the reporting month. Cut-off represents the % probability of default barrier above which a new mortgage is not originated. Loans represent the total loans in our sample that would be originated using the model, with the Defaulting column indicating which of these will default within our sample period. The default rate is Defaulting divided by Total. The right most section of the table shows the marginal impact of the Full model. Marginal loans not originated is the number of loans that would not have been originated if using the Full model, with Marginal defaults being the number of accounts that would not subsequently default. Finally, Marginal default rate shows the default rate of the marginal accounts.

	Base model			Full model			Difference		
Cut-off	Loans	Defaulting	Default rate	Loans	Defaulting	Default rate	Marginal loans not originated	Marginal Defaults	Marginal default rate
1	57,961	960	0.0166	56,930	894	0.0157	1,031	66	0.0640
2	62,332	1,152	0.0185	61,519	1,094	0.0178	813	58	0.0713
3	62,831	1,189	0.0189	62,454	1,146	0.0184	377	43	0.1141
4	62,887	1,194	0.0190	62,730	1,172	0.0187	157	22	0.1401
5	62,900	1,194	0.0190	62,825	1,179	0.0188	75	15	0.2000
6	62,905	1,194	0.0190	62,860	1,185	0.0189	45	9	0.2000
6+	62,905	1,194	0.0190	62,905	1,194	0.0190			

We cannot measure the number of additional loan applications approved under the Full model with our sample, however we can assume that the default rate would be consistent. Additionally, the default rate measured in Table 7 is higher than the cutoff due to the longer outcome period, with the outcome period being to the end of the sample for any one loan.

Therefore, using the Full model instead of the Base model generates a material improvement in loan book performance when a model includes transactional and behavioral variables.

3.6. Robustness tests

3.6.1 Out-of-sample testing

Both the Base and Full model are estimated out-of-sample tests. In each case, the model is estimated on the in-sample portion of the data and scored on the out-of-sample portion of the data. These models are then compared on their out-of-sample predictive AUROC power.

Table 8: Out of sample testing

This table depicts the out of sample testing of the Base model and the Full model. Panel A depicts the out of time testing, with the model built using the year modeled period, and tested using the out of time years. Panel B depicts out of sample testing where loans are separated into random proportional buckets referenced in the table header. For example, 20%/80% indicates that 20% of loan accounts were used for estimating the model, and 80% of the loan accounts were used to test the model. The numbers in the table refer to the out of sample Area Under the Receiver Operating Characteristic curve. For all instances, the Full model is significantly better than the Base model at the 1% level. All regressions include errors clustered at the state level.

Panel A - Out of time		**Prediction year**			
Model	Estimation period	2012	2013	2014	2015
Base	2011	0.721	0.7203	0.7402	0.7466
Base	2011 - 2012		0.725	0.7437	0.7479
Base	2011 - 2013			0.7467	0.7484
Base	2011 - 2014				0.7486
Full	2011 - 2011	0.8041	0.8096	0.8158	0.8122
Full	2011 - 2012		0.8274	0.8343	0.8262
Full	2011 - 2013			0.8462	0.8343
Full	2011 - 2014				0.8371

Panel B - In sample/out of sample				
Model	20%/80%	40%/60%	60%/40%	80%/20%
Base	0.7516	0.752	0.7511	0.7514
Full	0.839	0.8358	0.8366	0.8328

Panel A of Table 8 contains the out-of-time tests, with the model estimated using data in the "Estimation period" column and tested on data represented by the preceding columns. There is a clear difference between the Base and Full model for all parameters, with these changes all being significant at the one percent level. Model strength increases as time progresses due to an increase in the default rate of the sample. There is also an increase in model strength down the rows as more data is used to estimate the in-sample model.

Panel B of Table 8 includes the out-of-sample tests. Here, the sample is split based on loan IDs into an in-sample and out-of-sample portions according to the headings, with the first number indicating the in-sample and the second the out-of-sample allocation. Again, consistent with prior results, there is an increase in model strength when comparing the Base and Full model, with all the differences being significant at the one percent level.

In conclusion, our results are robust to both out-of-time and out-of-sample model tests.

3.6.2 Robustness over different time horizons

We adjust the default horizon in our models from 12 months to three and 24 months to assess the effectiveness of our models with less and more time to default. The purpose of this section is to show that the transactional variables continue to have the same relation to default across different time horizons, and that borrower habits remain significant.

Table 9 contains the regression results for this section. As expected, the signs across regressions remain the same for all variables, and the AUROC and Pseudo R-Square remaining in a similar range. It is important to note that the fit statistics cannot be directly compared due to recoding the default flag causing different sample structure. For AUROCs to be directly comparable it is required that the modeling sample is identical between models. The default flag is defined as an account-month being in default in the next number of months, thus the longer the default horizon the more default observations. The number of default observations for each model is included at the bottom of the table.

Interestingly, as the default lag is extended, the model coefficients move toward zero. This shows that although still being significant, the transactional data becomes less effective over very long horizons. If the lag is great enough, we can expect this data to become comparable to origination variables. Note that although changes in borrower circumstances may become less clear, habits are still being captured.

The clearest example of longer time horizons affecting coefficients to move toward zero is in the INCOME variable. For both INCOME and INCOMEP, the absolute coefficient is largest for the 3 Month model, and smallest for the 24 Month model. This shows the 6-month average income just prior to default is a greater driver compared to income up to 24 months prior to default. Furthermore, as horizons increase, the income will converge to the income level collected at origination.

Table 9: Alternate default lags

This table depicts results of regressions using alternate default lags. The dependent variable is a flag equal to one if an account-month will be in default in the next three months and the next 24 months compared to our main result case of 12 months (Full model). The target variable is equal to zero otherwise. Model specifications are otherwise identical to the main results. AUROC refers to the Area Under Receiver Operating Characteristic curve. *** and ** indicates significance at the 1% and 5% level respectively. Standard errors are clustered at the state level.

Panel A - Model results

Variable	3 Month		Full model		24 Month	
	-0.0649		0.0012		-0.0139	
INCOME	(41.64)	***	(0.01)		(0.87)	
	-0.2280		-0.1068		-0.0103	
INCOMEP	(42.39)	***	(6.95)	***	(0.19)	
	-2.8583		-2.5831		-2.1511	
SALARYD	(498.06)	***	(768.35)	***	(408.25)	***
	1.2417		1.1471		0.9671	
SALARYD2	(667.52)	***	(986.21)	***	(379.72)	***
	-0.9626		-0.7242		-0.5590	
SALARYDP	(780.09)	***	(838.16)	***	(191.03)	***
	0.2273		0.1730		0.1833	
CASH	(18.25)	***	(15.05)	***	(14.78)	***
	-1.4306		-1.298		-1.2305	
CASHP	(111.41)	***	(144.97)	***	(152.73)	***
	-0.2214		-0.2097		-0.2017	
CLOTHING	(399.58)	***	(419.04)	***	(1252.68)	***
	1.1781		1.1174		1.0607	
CLOTHINGP	(615.97)	***	(335.11)	***	(752.5)	***
	-0.1793		-0.1323		-0.1050	
HOMEMAINT	(222.28)	***	(125.77)	***	(72.19)	***
	2.0951		1.5887		1.2516	
HOMEMAINTP	(228.58)	***	(129.81)	***	(77.68)	***
	-1.6982		-1.2226		-1.0046	
NONDURDISCD	(241.23)	***	(397.39)	***	(792.03)	***
	0.5992		0.4542		0.3960	
NONDURDISCD2	(235.78)	***	(252.32)	***	(523.37)	***
	-0.8259		-0.6635		-0.5096	
NONDURDISCDP	(683.86)	***	(146.64)	***	(124.18)	***
Controls	YES		YES		YES	
AUROC	0.7965		0.8300		0.8176	
Pseudo R-Square	0.1493		0.1483		0.1465	
Observations	6,568,224		6,568,224		6,568,224	
Default observations	9,097		40,869		79,829	

3.6.3 Alternate sampling and factor design

We apply further tests to ensure the robustness of the previously presented results. The two main concerns we are addressing with these tests are whether the new factors are dependent on an exact specification, and whether the account-month structure of the sample is generating spurious results.

Table 10: Alternate factor and sample specification

This table depicts different parameterizations of the models to illustrate robustness. The Transform model contains 12-month moving averages instead of six-month moving averages that are in the main models. The January model limits the sample to only January observations; this limits the potential for bias originating from autocorrelation due to using account-month observations. AUROC refers to the Area Under Receiver Operating Characteristic curve. *** and ** indicates significance at the 1% and 5% level respectively. Chi-square statistics are in parenthesis. All regressions include standard errors clustered at the state level.

Variable	Full Model		Transform		January	
INCOME	0.0012		0.0149		-0.0207	
	(0.01)		(0.92)		(1.99)	
INCOMEP	-0.1068	***	-0.1402	***	-0.2429	***
	(6.95)		(8.74)		(12.52)	
SALARYD	-2.5831	***	-2.5707	***	-2.4356	***
	(768.35)		(796.43)		(275.35)	
SALARYD2	1.1471	***	1.1476	***	1.0321	***
	(986.21)		(844.88)		(215.24)	
SALARYDP	-0.7242	***	-0.6761	***	-0.7788	***
	(838.16)		(314.63)		(259.78)	
CASH	0.1730	***	0.1868	***	0.1852	***
	(15.05)		(16.87)		(21.92)	
CASHP	-1.298	***	-1.3769	***	-1.3662	***
	(144.97)		(186.11)		(144.37)	
CLOTHING	-0.2097	***	-0.1994	***	-0.1715	***
	(419.04)		(479.94)		(274.62)	
CLOTHINGP	1.1174	***	0.9879	***	0.9411	***
	(335.11)		(305.8)		(687.84)	
HOMEMAINT	-0.1323	***	-0.1291	***	-0.1435	***
	(125.77)		(115.56)		(79.02)	
HOMEMAINTP	1.5887	***	1.4518	***	1.7634	***
	(129.81)		(91.36)		(134.44)	
NONDURDISCD	-1.2226	***	-1.3598	***	-1.4913	***
	(397.39)		(516.3)		(82.61)	
NONDURDISCD2	0.4542	***	0.4908	***	0.5782	***
	(252.32)		(331.89)		(69.51)	
NONDURDISCDP	-0.6635	***	-0.6707	***	-0.817	***
	(146.64)		(155.23)		(178.08)	
CONTROL	YES		YES		YES	
AUROC	0.8300		0.8291		0.8277	
Pseudo R-Square	0.1483		0.1451		0.1426	
Observations	6,568,224		6,568,362		545,665	

To assess the factor specification concern, we adjust our model to include 12 month moving averages instead of the 6 month moving averages used in the main results. Using 12 month moving averages can give a longer-term view of the behavior of borrowers and may include a period where borrowers are not yet financially constrained. This will give a more normal view of the borrower's behavior over a longer period. The model is estimated in Table 10 (column 12 Month MA).

From the 12 Month MA model in Table 10, the model is not subject to factor specification, with the estimated coefficients and model fit being similar to that shown in the main results. Therefore, our results are robust to alternate factor specifications.

Another concern may be the account-month sample structure, which may include an autocorrelation bias due to auto correlated panel observations. To reduce this potential bias, we keep only January months, meaning that each account is only depicted once per year, with only one default observation for each defaulting account. The results can be seen in the January model of Table 10. Results are in line with our main results, with similar coefficient magnitudes and significance. Thus, our model is also robust to this alternate sample specification.

From these additional tests, our results are robust to alternate factor and sample specifications.

3.7. Conclusion

We use borrower bank and card transaction data and study the effectiveness of transactional data in predicting the probability of default for owner occupier mortgages. We find borrower financial transactions can be used to ascertain an accurate indication of financial position and general behavior. These measures can be used to more accurately separate defaulting and non-defaulting borrowers than traditional default drivers.

We further show default risk changes according to a borrower's change in circumstances. If a borrower experiences a change in income, they are more likely to default. Borrowers who are currently renovating, or actively maintaining their home are less likely to default, indicating their outlook on both property value, as well as their ability to service the loan. Borrowers who withdraw large amounts of cash or purchase clothing are more and less likely to default, respectively.

Our data and methodology greatly improve the model fit of a traditional variable model and provide a timelier prediction of credit risk.

This study is of interest to regulators, financial institutions, and mortgage holders as it outlines some behaviors that may indicate or lead to financial distress for borrowers. Effectively managing this financial stress can lead to lower default rates, less stress, and lower collection costs for financial institutions. Benefits can also accrue to homeowners if bank intervention can be earlier with more positive long-term outcomes.

The ethics surrounding data usage has received high attention. Private data at origination including details regarding income, expenses, and other credit facilities is already provided in loan applications and may not change this process. Transaction data which is to be considered for risk measurement, loan covenants and later finance applications may require borrower consent.

The debate on the proper use of transaction data is ongoing and fraught with grey areas, and it is important for banks to thoroughly consider both the commercial aspects and its impact on borrowers who trust banks to act in their best interest. The model expressed in this paper may benefit some prospective borrowers and disadvantage others. It needs to be considered whether the data used in this study requires treatment significantly different to other private data that already forms part of existing credit risk models.

Transaction data can highlight the impact of loan approval on consumer behavior. Theory states that individuals tend to be optimistic regarding their financial future, often expecting income to increase and expenses to decrease. The impact of a new loan on household finances is an interesting area for future research.

Internet appendix for Chapter 3

Appendix 1: Categorization of merchant codes

Table A1: Categorization of merchant codes

This table contains the Merchant Category Codes (MCC) and their associated grouping used for this study. These are originally sourced from the VISA Merchant Data Standards Manual (VISA 2017) and are formatted appropriately to be useful for our analysis. MCCs present in our data that were not part of the original list are manually added after inspecting the predominant merchant name within the category.

Our category	Merchant Category Code
Alcohol	2188, 5921
Cash out	Blank
Clothing	4843, 4844, 4845, 4846, 5137, 5139, 5311, 5611, 5621, 5631, 5641, 5651, 5655, 5661, 5681, 5691, 5697, 5698, 5699, 7296, 9340
Credit Card Repayment	6010
Financial services	5960, 6012, 6051, 6211, 6300, 6310, 6371, 6381, 6399, 6513, 7012, 7276, 7321, 8931, 9311
Gambling	7800, 7801, 7802, 7955, 7995
Groceries	5411, 5422, 5441, 5451, 5462, 5499
Health services	4891, 7297, 8011, 8021, 8031, 8041, 8042, 8043, 8044, 8049, 8050, 8062, 8071, 8099, 8126, 8151, 8152, 8158
Home maintenance	0780, 1520, 1711, 1731, 1740, 1750, 1761, 1771, 1799, 2842, 4232, 4233, 4242, 4727, 4848, 4849, 4856, 4857, 5065, 5072, 5074, 5085, 5198, 5200, 5211, 5231, 5251, 5261, 5712, 5713, 5714, 5718, 5719, 5722, 5732, 6240, 7217, 7342, 7349, 7623, 7629, 7641, 7692, 8911
Local transport	4111, 4112, 4121, 4131, 4784, 5107, 5108, 5541, 5542, 5544, 7523, 7532, 9000
Other	0742, 0763, 1234, 2130, 2741, 2791, 4011, 4119, 4214, 4215, 4225, 4468, 4717, 4751, 4789, 4815, 4847, 4854, 4855, 4879, 4892, 4893, 4894, 4895, 4896, 5013, 5021, 5039, 5044, 5045, 5046, 5047, 5051, 5094, 5099, 5111, 5122, 5131, 5169, 5172, 5192, 5193, 5199, 5300, 5309, 5310, 5331, 5399, 5733, 5734, 5735, 5811, 5815, 5816, 5817, 5818, 5832, 5912, 5931, 5932, 5933, 5935, 5937, 5940, 5941, 5942, 5943, 5944, 5945, 5946, 5947, 5948, 5949, 5950, 5961, 5962, 5963, 5964, 5965, 5966, 5967, 5968, 5969, 5970, 5971, 5972, 5973, 5975, 5976, 5977, 5978, 5983, 5992, 5993, 5994, 5995, 5996, 5997, 5998, 5999, 6011, 6390, 6540, 7112, 7210, 7211, 7216, 7221, 7230, 7251, 7261, 7273, 7277, 7278, 7298, 7299, 7311, 7332, 7333, 7338, 7339, 7361, 7372, 7375, 7379, 7392, 7393, 7394, 7395, 7399, 7511, 7622, 7631, 7699, 7829, 7832, 7841, 7911, 7922, 7929, 7932, 7933, 7941, 7991, 7992, 7993, 7994, 7996, 7998, 7999, 8111, 8160, 8211, 8220, 8231, 8234, 8241, 8244, 8249, 8299, 8305, 8351, 8398, 8641, 8651, 8661, 8675, 8699, 8734, 8999, 9133, 9136, 9141, 9211, 9222, 9223, 9351, 9399, 9402, 9405, 9700, 9701, 9702, 9950
Restaurants	0020, 5812, 5813, 5814, 6050, 7997, 9212, 9231, 9232, 9241

Travel- local and overseas	3000, 3001, 3002, 3003, 3004, 3005, 3006, 3007, 3008, 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018, 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028, 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038, 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047, 3048, 3049, 3050, 3051, 3052, 3053, 3054, 3055, 3056, 3057, 3058, 3059, 3060, 3061, 3062, 3063, 3064, 3065, 3066, 3067, 3068, 3069, 3070, 3071, 3072, 3073, 3074, 3075, 3076, 3077, 3078, 3079, 3080, 3081, 3082, 3083, 3084, 3085, 3086, 3087, 3088, 3089, 3090, 3091, 3092, 3093, 3094, 3095, 3096, 3097, 3098, 3099, 3100, 3101, 3102, 3103, 3104, 3105, 3106, 3107, 3108, 3109, 3110, 3111, 3112, 3113, 3114, 3115, 3116, 3117, 3118, 3119, 3120, 3121, 3122, 3123, 3124, 3125, 3126, 3127, 3128, 3129, 3130, 3131, 3132, 3133, 3134, 3135, 3136, 3137, 3138, 3139, 3140, 3141, 3142, 3143, 3144, 3145, 3146, 3147, 3148, 3149, 3150, 3151, 3152, 3153, 3154, 3155, 3156, 3157, 3158, 3159, 3160, 3161, 3162, 3163, 3164, 3165, 3166, 3167, 3168, 3169, 3170, 3171, 3172, 3173, 3174, 3175, 3176, 3177, 3178, 3179, 3180, 3181, 3182, 3183, 3184, 3185, 3186, 3187, 3188, 3189, 3190, 3191, 3192, 3193, 3194, 3195, 3196, 3197, 3198, 3199, 3200, 3201, 3202, 3203, 3204, 3205, 3206, 3207, 3208, 3209, 3210, 3211, 3212, 3213, 3214, 3215, 3216, 3217, 3218, 3219, 3220, 3221, 3222, 3223, 3224, 3225, 3226, 3227, 3228, 3229, 3230, 3231, 3232, 3233, 3234, 3235, 3236, 3237, 3238, 3239, 3240, 3241, 3242, 3243, 3244, 3245, 3246, 3247, 3248, 3249, 3250, 3251, 3252, 3253, 3254, 3255, 3256, 3257, 3258, 3259, 3260, 3261, 3262, 3263, 3264, 3265, 3266, 3267, 3268, 3269, 3270, 3271, 3272, 3273, 3274, 3275, 3276, 3277, 3278, 3279, 3280, 3281, 3282, 3283, 3284, 3285, 3286, 3287, 3288, 3289, 3290, 3291, 3292, 3293, 3294, 3295, 3296, 3297, 3298, 3299, 3351, 3352, 3353, 3354, 3355, 3356, 3357, 3358, 3359, 3360, 3361, 3362, 3363, 3364, 3365, 3366, 3367, 3368, 3369, 3370, 3371, 3372, 3373, 3374, 3375, 3376, 3377, 3378, 3379, 3380, 3381, 3382, 3383, 3384, 3385, 3386, 3387, 3388, 3389, 3390, 3391, 3392, 3393, 3394, 3395, 3396, 3397, 3398, 3399, 3400, 3401, 3402, 3403, 3404, 3405, 3406, 3407, 3408, 3409, 3410, 3411, 3412, 3413, 3414, 3415, 3416, 3417, 3418, 3419, 3420, 3421, 3422, 3423, 3424, 3425, 3426, 3427, 3428, 3429, 3430, 3431, 3432, 3433, 3434, 3435, 3436, 3437, 3438, 3439, 3440, 3441, 3501, 3502, 3503, 3504, 3505, 3506, 3507, 3508, 3509, 3510, 3511, 3512, 3513, 3514, 3515, 3516, 3517, 3518, 3519, 3520, 3521, 3522, 3523, 3524, 3525, 3526, 3527, 3528, 3529, 3530, 3531, 3532, 3533, 3534, 3535, 3536, 3537, 3538, 3539, 3540, 3541, 3542, 3543, 3544, 3545, 3546, 3547, 3548, 3549, 3550, 3551, 3552, 3553, 3554, 3555, 3556, 3557, 3558, 3559, 3560, 3561, 3562, 3563, 3564, 3565, 3566, 3567, 3568, 3569, 3570, 3571, 3572, 3573, 3574, 3575, 3576, 3577, 3578, 3579, 3580, 3581, 3582, 3583, 3584, 3585, 3586, 3587, 3588, 3589, 3590, 3591, 3592, 3593, 3594, 3595, 3596, 3597, 3598, 3599, 3600, 3601, 3602, 3603, 3604, 3605, 3606, 3607, 3608, 3609, 3610, 3611, 3612, 3613, 3614, 3615, 3616, 3617, 3618, 3619, 3620, 3621, 3622, 3623, 3624, 3625, 3626, 3627, 3628, 3629, 3630, 3631, 3632, 3633, 3634, 3635, 3636, 3637, 3638, 3639, 3640, 3641, 3642, 3643, 3644, 3645, 3646, 3647, 3648, 3649, 3650, 3651, 3652, 3653, 3654, 3655, 3656, 3657, 3658, 3659, 3660, 3661, 3662, 3663, 3664, 3665, 3666, 3667, 3668, 3669, 3670, 3671, 3672, 3673, 3674, 3675, 3676, 3677, 3678, 3679, 3680, 3681, 3682, 3683, 3684, 3685, 3686, 3687, 3688, 3689, 3690, 3691, 3692, 3693, 3694, 3695, 3696, 3697, 3698, 3699, 3700, 3701, 3702, 3703, 3704, 3705, 3706, 3707, 3708, 3709, 3710, 3711, 3712, 3713, 3714, 3715, 3716, 3717, 3718, 3719, 3720, 3721, 3722, 3723, 3724, 3725,

	3726, 3727, 3728, 3729, 3730, 3731, 3732, 3733, 3734, 3735, 3736, 3737, 3738, 3739, 3740, 3741, 3742, 3743, 3744, 3745, 3746, 3747, 3748, 3749, 3750, 3751, 3752, 3753, 3754, 3755, 3756, 3757, 3758, 3759, 3760, 3761, 3762, 3763, 3764, 3765, 3766, 3767, 3768, 3769, 3770, 3771, 3772, 3773, 3774, 3775, 3776, 3777, 3778, 3779, 3780, 3781, 3782, 3783, 3784, 3785, 3786, 3787, 3788, 3789, 3790, 3816, 3835, 4411, 4457, 4511, 4582, 4722, 4723, 5514, 7011, 7032, 7033, 7512, 7513, 7519
Utilities	4812, 4814, 4816, 4818, 4821, 4829, 4899, 4900
Vehicles	4243, 4247, 4731, 4732, 4861, 4865, 4866, 4867, 4868, 5271, 5511, 5521, 5531, 5532, 5533, 5551, 5561, 5571, 5592, 5598, 5599, 7310, 7531, 7534, 7535, 7538, 7542, 7549

Appendix 2: Sample transactions

Table A2: Sample transactions

This table contains an adapted sample of the transaction data. The data contains a timestamp, amount, merchant name, merchant category code (MCC), our transaction categorisation and the location of the transaction. The last column is generated by the link table from Appendix 1

Month key	Amount	Merchant name	MCC	Category	Suburb
1/12/2013	-67.00	MEGAN PARK RETAIL PTY	5621	Clothing	ARMADALE
3/12/2013	-38.45	COLES MALVERN	5499	Groceries	MALVERN
3/12/2013	-68.10	TATTS ONLINE PTY LTD	7995	Gambling	ALBION
4/12/2013	-33.00	TATTS ONLINE PTY LTD	7995	Gambling	ALBION
7/12/2013	-160.25	HONEYBEE TOYS	5945	Other	MALVERN
8/12/2013	-142.80	TATTS ONLINE PTY LTD	7995	Gambling	ALBION
9/12/2013	-632.00	CABRINI HOSPITAL 000	8062	Health Services	MALVERN
12/12/2013	-36.65	FRUIT NEST MALVERN	5499	Groceries	MALVERN
14/12/2013	-34.90	COUNTRY ROAD-SOUTH YAR	5691	Clothing	SOUTH YARRA
14/12/2013	-70.90	DAVID JONES MALVERN CE	5311	Clothing	MALVERN
15/12/2013	-29.56	WOOLWORTHS 3162 MALVER	5411	Groceries	MALVERN
15/12/2013	-330.00	ZIMMERMANN WEAR P/L	5621	Clothing	SOUTH YARRA
19/12/2013	-81.90	DAVID JONES MALVERN CE	5311	Clothing	MALVERN
19/12/2013	-121.50	BIG W 0367 SOUTH YARRA	5311	Clothing	SOUTH YARRA
20/12/2013	-98.00	JB HI FI	5733	Other	CHADSTONE
20/12/2013	-24.00	NEWS LIMITED - PAID CO	7311	Other	SURRY HILLS
21/12/2013	-89.35	STOCKED MALVERN	5499	Groceries	MALVERN
23/12/2013	-43.95	COUNTRY ROAD	5691	Clothing	RICHMOND
23/12/2013	-243.00	TORSA PRAHRAN VIC	5621	Clothing	PRAHRAN
25/12/2013	-93.40	COLES EXPRESS MALVERN	5541	Local Transport	MALVERN
26/12/2013	-17.95	SEED 118 MALVERN CENTR	5621	Clothing	MALVERN
26/12/2013	-43.70	WATTLETREE ROAD LPO	9402	Other	MALVERN EAS
27/12/2013	-117.30	MARIMEKKO AUSTRALIA PT	5691	Clothing	SOUTH YARRA
27/12/2013	-30.96	DAN MURPHYS 3610 PRAHR	5921	Alcohol	PRAHRAN
27/12/2013	-50.85	GARDEN LOVERS ON WATTL	5261	Home Maintenance	MALVERN EAS
28/12/2013	-144.80	COUNTRY ROAD	5691	Clothing	RICHMOND
29/12/2013	-66.00	SISSI AND CO FINE FOOD	5812	Restaurants	MALVERN

Appendix 3: Pre-processing of transactional data

In this section we outline our methodology for identifying factors and display some examples of factors with high association to default.

Our main intention in building the large aggregate categories is to isolate the main areas of individual spending, with a focus on factors that may be important for default. For example, we assume that grocery, and local travel intuitively constitute large portions of individual expenditure, while at the same time dedicated home expenses such as home improvement can be directly correlated to default. Although more categories can be created to tease out effects within the "Other" category, along with rearranging the dedicated categories, our focus is to document an initial foray into this new data.

The data mining aspect is concentrated in the testing of the variables and not the categorization of the transactions. Specifically, once the categories are established based on manual inspection of the MCC code description, these are not changed unless tests indicate there is data for an unlisted MCC code, or if the transactions marked with an MCC code do not match the description. There is a risk of incorrect MCC categorization when a merchant facility is installed, as this depends on correct input of the installer. We take care to ensure the categories are materially correct[21]. Since all active MCC codes are allocated to a high-level aggregated category, each has a significant number of transactions able to be used in testing.

Even though our sample contains a high quantity and frequency of transactions, to obtain a more consistent view of borrower behavior we make use of moving averages. This serves two purposes: it spreads any large one-time expense over subsequent months and creates a more stable borrower view. This is important for transactions that do not necessarily occur on a consistent monthly basis like home maintenance. The length of the moving average does not impact the result.

Another point of interest is how changes in behavior affect the risk of a borrower. To capture this effect, we measure changes over time. Interpretation of this relation can be more difficult as observations in some spend categories may be inconsistent. For example, although home maintenance may be needed on a regular basis, it may only occur every few months, therefore looking at a change over a period (perhaps a year) may not indicate a legitimate change in behavior, although a change will be measured.

[21] Some of the MCC codes are rarely used. These MCC codes are included nonetheless, however their impact is small.

While we calculate changes for certain categories over time, it is more difficult to interpret the results that stem from these tests.

Our test variables are scaled to account for different effects. This is to allow for heterogeneous endowments and their potential effect on transactions. For example, an individual with low endowment may be more adversely affected by overspending in certain categories compared to individuals with higher endowment. The scalars tested include property value, mortgage interest charged, total spend on credit and debit cards and application income. An unscaled variant is also tested. The scalars result in minor differences, and the final scalar used in subsequent regressions is selected based on which scalar received the highest AUROC.

Table A3 outlines some of the variables we created and tested, along with an Area Under the Receiver Operator Characteristic (AUROC) score. The AUROC is a measure between 0.5 and one that indicates the discriminatory power of a model.[22] An AUROC of 0.5 shows the model provides no additional separation between defaulters and non-defaulters compared to a random allocation, and an AUROC of one indicates perfect separation of defaulting and non-defaulting loans.

The columns represent different transformations of the test variable to capture linear and non-linear relations to default. The first column (6 Month MA) contains the result of using the untransformed six-month moving average (MA6) as the test variable ($X_{Dit-1}^{Transactional} = MA6$). This simple transformation is adequate to capture the linear relations across all test variables. A sample of the selected test variables and their relation to default can be seen in Figure 2.

To capture factor changes over time we include a simple transformation which compares the three-month moving average (MA3) to the 12-month moving average (MA12). If there is a recent increase (decrease) in a transaction category, the ratio $\frac{MA3}{MA12}$ will be greater (less) than one, if the borrower's behavior is consistent over time then the ratio will remain near one. Investigation of the relation of this factor to default resulted in a clear nonlinear relation as illustrated in Figure 2. To capture this nonlinearity, we include a quadratic term. Both the linear and quadratic change factor is included in

[22] It can be shown that other popular measures for discrimination, such as the Gini Coefficient or the Accuracy Ratio, are strict monotone transformations.

Table A3 in the columns MA3/MA12 ^1 ($X_{Dit-1}^{Transactional} = \frac{MA3}{MA12}$) and MA3/MA12 ^2 ($X_{Dit-1}^{Transactional} = \frac{MA_3}{MA_{12}}, \frac{MA_3}{MA_{12}}^2$) respectively.

We test a range of scaling variables including property purchase price (PROP), original application income (LNAPPLINC), monthly interest calculated using the interest rate and the current loan balance (MNTHINT), the total purchase amount (TOTSPEND), and an unscaled variable (ONE). The purpose of these scaling factors is to allow for different asset endowments, general income levels, and competing expenses. We use loan level variables (except TOTSPEND, which is a spend variable) since these are available for all loans. The use of purchase price and application income is intuitive, and we include monthly interest as a proxy for competing spending needs originating from the loan. For succinctness, Property Value is the main scalar included in Table A3, with only selected variables presented for other scalars.

Table A3 indicates many of the new variables add information to the Base model. The strongest factor overall is durable spend to monthly interest, which increases the AUROC of the Base model by 0.0467 from 0.7426 to 0.7893. For our purpose, we elect to use two of the constituents of durable spend instead of the aggregate as we believe these represent different spending motives. Specifically, we elect to use home maintenance and clothing spend separately. Home maintenance spend is generally conducted under the assumption the borrower will be able to experience the benefit of the expense which would be lost if there is a default on the loan (Choi *et al.* 2014), therefore home maintenance spend can be a proxy for a borrower's belief in their own loan serviceability. Clothing is the second factor selected. It can be argued that clothing is a non-discretionary expense; however, in general, borrowers are not in dire need of new clothing and are likely to service their mortgage instead if financially constrained. Therefore, we select home maintenance to property value and clothing to monthly interest as two variables to investigate further in the following sections. Both factors have a negative relation to default, i.e., an increase in spend is accompanied by a decrease in default risk.

Table A3: AUROC results for multivariate regression

This table depicts the AUROC of a logit regression of the Stage 2 Base model and a transformed transaction variable on a binary variable representing default in the next 12 months. The columns represent different transformations of the transaction type variable, with the rows representing the transaction type and the scaling variable. Values in bold represent variables selected for modeling. Letters are used to indicate construction of aggregated variables. 6 Month MA is the average of a category for the last six months as at reporting date. MA3/MA12 ^1 is calculated as the three-month moving average over the 12-month moving average. The MA3/MA12 ^2 model includes MA3/MA12 ^1 as well as the same factor squared. All regressions include errors clustered at the state level.

Scaling variable	Spend type	6 Month MA	MA3/MA12 ^1	MA3/MA12 ^2
Property value	Home maintenance (a)	**0.7716**	0.7558	0.7648
	Clothing (b)	0.7790	0.7583	0.7744
	Durable discretionary (a+b = O)	0.7828	0.7551	0.7739
	Vehicles (c)	0.7567	0.7519	0.7541
	Durable non-discretionary (c = C)	0.7567	0.7519	0.7541
	Durable (a+b+c = P)	0.7837	0.7551	0.7754
	Groceries (d)	0.7487	0.7476	0.7582
	Utilities (e)	0.7449	0.7439	0.7552
	Local transport (f)	0.7461	0.7463	0.7548
	Non-durable non-discretionary (d+e+f = Q)	0.7495	0.7484	0.7650
	Restaurant (g)	0.7500	0.7472	0.7547
	Health services (h)	0.7628	0.7510	0.7579
	Gambling (i)	0.7426	0.7426	0.7426
	Alcohol (j)	0.7436	0.7433	0.7458
	Travel (k)	0.7640	0.7541	0.7595
	Financial services (l)	0.7650	0.7585	0.7661
	Non-durable discretionary (g+h+i+j+k+l = R)	0.7682	0.7503	**0.7648**
	Non-durable (Q+R = S)	0.7589	0.7503	0.7701
	Cash out (m)	0.7664	0.7684	0.7692
	Other (n)	0.7616	0.7500	0.7677
	Discretionary (O+Q = S)	0.7536	0.7497	0.7670
	Non-discretionary (Q+C=T)	0.7550	0.7496	0.7680
	Total spend (S+T = U)	0.7551	0.7507	0.7726
	Proportion discretionary (S/U)	0.7454	0.7456	0.7567
	Salary	0.7648	0.7520	0.7706
	Income	0.7563	0.7482	0.7644
Selected variables with different scaling factors				
	Salary	0.7648	0.7520	**0.7706**
Mortgage interest	Clothing	**0.7804**	0.7583	0.7744
Mortgage interest	Durable	0.7893	0.7551	0.7754
Mortgage interest	Income	**0.7549**	0.7482	0.7644
Total spend	Cash out	**0.7745**	0.7684	0.7692

An interesting finding in our data is the predictive power of cash withdrawals from ATMs. In contrast to many of the other variables that are tested, cash out has a positive relation to default, with an increase in the level of cash out accompanied with an increase in probability of default. We select cash out scaled by total spend as the variable for subsequent inclusion as it results in the highest AUROC at 0.7745. Due to the strength and relation of this factor to default we include this variable for further study.

Salary change, and short-term interest coverage ratio (income scaled with mortgage interest) are two income-related factors we include. Although income is included at application, this is not generally updated until there is a revision in the loan. We include change in salary as it represents a change in circumstances which directly impacts loan repayments. We also test short-term interest coverage measured as average income in the last six months divided by average interest accrued in the last six months as a direct measure of a borrower's point in time serviceability. The income data allows for these measures to be updated on a regular basis, and as income is a large driver in loan serviceability, these are included.

To include the information that may be missed in the currently selected spending variables, we select a transformation of non-durable discretionary expenses which includes spend categories such as restaurants, alcohol and local travel which aren't otherwise considered. We choose the quadratic transformation of the change in spending in the last three months compared to the 12-month average. The AUROC for this transformation is 0.7648, which is 0.0222 higher than the Base model. This gives a general indication of changes in non-discretionary spending without including the factors we have above.

Although there are more factors that can be included in this study, many are simply transformations and aggregations of the lowest level spend categories, and as such they are not studied further. The aim of this study is to investigate the potential benefits of this new behavioral data source for credit risk modeling without providing a comprehensive study of all factors and circumstances.

To ensure our chosen factors are not highly correlated, and thus impacting the robustness of our results, we compare the correlation of all variables that may be included in our final model. Our goal is to ensure we are not including highly correlated variables in the model. Although we have several highly predictive variables, they may contain the same informational component.

From Table A4 it is clear the only highly correlated variables are the square terms of salary change and short-term interest coverage ratio due to the same underlying variable. Clothing and income are also moderately correlated to outstanding balance due to the denominator (monthly interest) being directly related to outstanding balance. Finally, outstanding balance is understandably related to property value, with both variables forming part of the Base model. Some variables were removed from the Base model due to a correlation to other Base model variables. A cutoff value of 0.6 was used for the correlation.

The final variables we select to be included in the modeling stage are home maintenance to property value, short term interest coverage ratio (average income in the last six months to interest), six-month average cash out to total spend, salary change, and non-durable discretionary change over the last six months.

Table A4: Pearson correlations

This table contains the correlation between all variables that are used in the modeling stage of this paper. Overall, most variables are relatively uncorrelated, ensuring minimal risk of multicollinearity. Variables in this table represent the final transformed modeling factors as presented in the text, hence the use of capitalization.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
LOWDOC	(1)	1.00	-0.01	0.02	-0.01	-0.01	0.03	0.03	-0.01	0.01	-0.04	-0.03	0.00	-0.02	0.01	-0.01	0.00	-0.08	-0.08	-0.05
FIXEDINT	(2)		1.00	0.11	-0.01	-0.08	-0.03	0.17	0.12	0.02	-0.23	0.25	-0.05	-0.09	-0.01	0.01	0.00	-0.05	0.03	0.02
LNOUTSBAL	(3)			1.00	0.14	0.32	0.51	-0.21	0.47	0.08	-0.03	0.21	-0.07	-0.51	-0.08	0.01	0.00	-0.54	0.07	0.05
LNINCOME	(4)				1.00	0.06	0.27	-0.15	-0.04	-0.01	0.02	-0.16	0.04	0.01	-0.06	0.02	-0.01	-0.03	0.04	0.02
APPLOANINCOME	(5)					1.00	0.32	-0.13	0.22	-0.02	0.23	0.12	-0.04	-0.18	-0.05	0.00	0.01	-0.21	0.00	0.01
LNPROP	(6)						1.00	-0.21	0.04	-0.03	0.07	0.23	0.01	-0.19	-0.18	0.02	-0.01	-0.24	0.03	0.02
INTMARGIN	(7)							1.00	-0.23	0.00	-0.04	-0.37	-0.01	0.06	0.00	0.00	0.03	0.08	-0.06	-0.03
CLTV	(8)								1.00	0.07	-0.07	0.14	-0.11	-0.28	0.00	0.01	0.00	-0.22	0.09	0.07
CCUTIL	(9)									1.00	-0.04	-0.16	0.01	-0.08	-0.05	0.05	0.02	-0.05	0.01	0.01
BUREAU	(10)										1.00	0.25	0.04	0.05	-0.02	0.01	0.00	0.05	0.01	0.00
IMR	(11)											1.00	0.06	-0.12	0.03	-0.03	-0.03	-0.14	0.03	0.01
CASH	(12)												1.00	0.09	0.13	-0.09	-0.05	0.02	-0.03	-0.03
CLOTHING	(13)													1.00	0.15	-0.05	-0.05	0.40	-0.04	-0.03
HOMEMAINT	(14)														1.00	-0.17	-0.09	0.00	-0.06	-0.05
NONDURDISCD	(15)															1.00	0.94	0.03	0.06	0.05
NONDURDISCD2	(16)																1.00	0.01	0.04	0.04
INCOME	(17)																	1.00	0.36	0.22
SALARYD	(18)																		1.00	0.91
SALARYD2	(19)																			1.00

Chapter 4: Announcement effect of Trump Tariffs

4.1. Motivation

The final impact of the trade tariffs imposed by US president Donald Trump is unknown, with Trump supporters claiming a benefit and Trump opposition claiming a detriment to the US economy and its relationship with neighbors. Measuring the impact of the tariff announcements carries unresolved uncertainty concerning its final shape and execution, with traditional event studies not able to resolve the ambiguity in the announcement.

In this paper we quantify the announcement effect of the solar and washing machine tariff (SOLAR), and the steel and aluminum tariff (STEEL), accounting for unresolved uncertainty while measuring the value and news effects. We accomplish this through a novel options-based methodology developed by Barraclough *et al.* (2013) and modified by Han *et al.* (2019).

We find that both the SOLAR tariff and STEEL tariff demonstrate a negative and significant value effect, indicating a large reduction in the value in US firms. The news effect is not significant, consistent with tariffs being part of Trump's election promises. In this case the news of impending tariffs was already priced, but the final value effect of the execution was not known.

We contribute to the literature of the use of tariffs by quantifying the value and wealth effects of the announcement, controlling for the possibility that the announced event does not come into effect. We show the effect of protectionist tariffs on growth can be negative (Osang & Pereira 1996), and show that tariffs can generate large welfare losses in the medium and long term (Akcigit *et al.* 2018).

We also contribute to the literature on announcements with unresolved uncertainty. The model used in this paper allows for a separation of the different effects present in any announcement. Using a traditional event study methodology when the final implementation is uncertain can cause an underestimation of the announcement effect. Our methodology captures this uncertainty and assigns a probability to a successful outcome. Our methodology also separates the new information entering the market from its effect on the economic value of firms. We find the announcements released an insignificant amount of news and caused a significant reduction in the value of firms. The effect of indirect consequences is also captured, the effect measured when the tariff is implemented includes any secondary effects and their probability. For this paper, the secondary effects of additional trade conflicts with Canada and Europe are captured within the estimation results.

This paper has implications for future tariff decisions. Trump's aim with the tariffs is to increase the value of the US market by reducing the trade deficit to other countries. Although this may eventually be the case, the increase in the cost of goods created by the tariffs results in a large value drain.

This paper continues as follows. Section 2 outlines a review of related literature and outlines the timeline of the two announcements presented in this study. Section 3 presents the data and some summary statistics. Section 4 outlines the methodological framework with Section 5 presenting the results and Section 6 concluding.

4.2. Literature review

Tariffs

Tariffs and duties play a significant role in global economies, with the changes within major players impacting neighbors. China has been a major contributor to international trade for many years, with modern technology allowing for global trade with relative ease. China's rapid growth and recent slowdown has had a significant impact on other global economies, and in this case, is purported to have a significant impact on the US economy. The slowing of the Chinese economy has resulted in an oversupply of materials for local markets, with this oversupply spilling into other countries. Steel is an example of this, with subsidized steel being imported to the US, reducing the competitiveness of the US steel industry.

Import tariffs are instituted to increase the price of imported goods or services to restrict trade by making imported goods more expensive than locally produced goods. These are collected by local governments as an additional source of revenue. Tariffs can be used to shift demand to locally produced goods, with the welfare being distributed within the local population (Felbermayr *et al.* 2015).

Tariffs can be used to protect local producers by providing a price advantage (Bhagwati & Ramaswami 1963). According to World Trade Organisation (WTO 2019), import tariffs provide a price advantage to locally manufactured products or services by making the imported product or service more expensive. Tariffs can also reduce growth. Osang and Pereira (1996) find that most tariffs reduce growth, even in the short term and all tariffs are welfare reducing in the long run.

Tariffs between identical countries slow technological change and economic growth at the margin (Rivera-Batiz & Romer 1991). Grossman and Helpman (1990) find that a small import tariff on final goods only increase economic growth if the country enacting the tariff has a disadvantage in R&D. These papers assume there can be no shift in the comparative advantage of firms, either through a switch in advantage or a country taking a greater lead. Whether the US has a comparative advantage in R&D with regards to solar or steel production is unclear. The model by Dinopoulos and Segerstrom (1999) finds that large 'protective' contingent tariffs allow domestic firms to capture the local market despite being technologically behind and is positively related to the global rate of technological change in the short run.

Akcigit *et al.* (2018) find that reduced trade tariffs increase domestic innovation through induced international competition, and that a tax credit is an effective policy response to foreign competition. Furthermore, tariffs generate large welfare losses in the medium and long term, or when there is retaliation from a foreign economy. Protectionist measures reduce the openness of an economy, distorting the innovation incentives and productivity growth.

As the tariffs in our paper currently stand, they are simply a demand shifting tariff which increases the price of goods, with prior theory and evidence suggesting negative welfare effects for the US economy.

Impact of China on the US

China has seen staggering growth in recent history, with lower labor costs giving a competitive advantage compared to the higher labor costs of the US. Lin and Wang (2018) show that East Asian countries have a comparative advantage over the US in labor cost, which lead to these countries being a major source of the US trade deficit. This comparative advantage resulted in manufacturing being moved to China, exacerbating the trade deficit. The authors expect the gentrification of China to cause manufacturing to shift away from China to other low labor cost countries, resulting in a decrease in the China-US trade imbalance, but not greatly affecting the US aggregate trade imbalance.

David *et al.* (2013) find that rising imports cause higher unemployment and reduced wages in import competing industries, with import competition explaining one-quarter of the contemporaneous aggregate decline in US manufacturing employment. Further consequences of these changes include increase transfer benefits payments for unemployment, disability, retirement, and healthcare in more trade-exposed labor markets.

Announcement effect with uncertainty

Estimations of wealth effects are difficult in situations where there is uncertainty regarding whether a policy will eventuate. For example, announcing a tariff does not mean that the tariff will eventually be implemented, and furthermore, what the structure of the tariff will be. This is especially true given President Trump's negotiation tactics. Due to this identification problem, a traditional event study method like Fama *et al.* (1969) and Ball and Brown (1968) cannot be applied as the probability of outcome is unknown. In this situation there is a probability parameter, as well as outcome price parameters (Bhagat *et al.* 2005; Barraclough *et al.* 2013; Linnenluecke *et al.* 2016).

To overcome these limitations, Han *et al.* (2019) propose a novel approach based on the model developed by Barraclough *et al.* (2013) and Borochin and Golec (2016). We implement this model using stock and call option prices to estimate the probability and outcome parameters.

4.2.1. Timeline of the Trump tariffs

One of Trump's election promises was to "bring manufacturing back", and one of the ways Trump has chosen to pursue this is through import tariffs on certain goods, protecting and encouraging local manufacturers to expand their operations.

The path of the tariffs started in the middle of 2017 with International Trade Commission (ITC) initiating an injury investigation on solar cells and residential washers. In mid to late 2017, the ITC concluded that that there was injury sustained due to import of solar components and residential washers. In late 2017 the ITC submitted a series of recommendations to the president, followed by the president announcing a tariff on solar components and large residential washers on 23rd January 2018, effective February 7, 2018. Although the executive order was signed on 23rd

January 2018, Trump announced on 18th January 2018 that a decision regarding the SOLAR tariffs was coming, thus we use the 18th January 2018 as our announcement event date.

Table 1: Key events

This table outlines the timeline and key events present in the SOLAR and STEEL tariffs.

Solar and washing machines	
May - June 2017	U.S. industry petition initiates International Trade Commission injury investigation on solar cells/modules and large residential washers.
September - October 2017	International Trade Commission makes affirmative solar cells/modules and large residential washers injury determination.
November - December 2017	International Trade Commission submits report and recommended action on solar cells/modules and large residential washers to President.
Thursday, 18 January 2018	Trump states that an announcement is coming to Reuters.
Tuesday, 23 January 2018	President proclaims actions on solar cells/modules and large residential washers, effective February 7, 2018.
Steel and aluminum	
Wednesday, 4 January 2017	Commerce initiates investigations on effects on national security of U.S. steel (4/19) and aluminum (4/26) imports. President signs memoranda prioritizing steel and aluminum investigations.
Monday, 1 January 2018	Commerce submits steel (1/11) and aluminum (1/17) findings and recommendations to President.
Thursday, 1 March 2018	President announces his intention to impose a 25% tariff on steel and 10% tariff on aluminum imports.
Thursday, 8 March 2018	President proclaims steel and aluminum duties, effective March 23, 2018, temporarily exempting Canada and Mexico.

Initiation of investigations leading to the STEEL tariffs were happening concurrently, with Commerce initiating investigations in April 2017 on the effects of steel and aluminum imports on national security. In January 2018, Commerce submitted the steel and aluminum findings and recommendations to the president. On March 1, 2018, Trump announced his intention to impose a 25% tariff on steel and 10% tariff on aluminum imports, issuing an executive order on 8th March 2018, with the tariffs becoming effective February 7, 2018. Eventually through negotiations some countries were exempted from this tariff, including Argentina, Australia, Brazil, South Korea, the European Union (EU) in addition to Canada and Mexico.

These announcements were the start of the international trade war, concentrated between the US and China. At the current date the outcome of the war is unknown. The increase in the final cost of goods to the consumer would make consumer the main casualty.

4.3. Data and summary statistics

Stock and option data are extracted from Thomson Reuters Tick History for all firms in the S&P 100 around the two announcement dates 18 January 2018 and 1 March 2018. Firms are separated based on their industry classification group sourced from WRDS. Some firms may be manually classified depending on investigation of announcement effects for a firm. We exclude firms that are close to delisting or a share split. There is no need for firms to have survived for the whole period as our estimates are made ex ante at the time of the announcement. The relevant fields extracted for the stock series are the daily closing price, dividend payments, ex-dividend date, trading volume and shares outstanding. The options fields extracted include option price, trading volume, open interest, exercise price and maturity. Delta is calculated using actual dividends that will be received. The total market capitalization captured is approximately 12 trillion dollars. Actual future dividends give the best estimate of expected future dividends. Although there is a possibility that dividends may change as a result of the new tariffs, the impact would be low as it will impact an insignificant number of firms.

The largest firm in our sample is AAPL with market capitalization of approximately 900 billion US dollars, and the smaller firm is FOX with a market capitalization of approximately 30 million US dollars indicating that we are covering a wide range of firm values. The average firm size in our sample is 159 million US dollars, SOLAR affected, and unaffected firms are on average valued at 144 and 187 million US dollars respectively. STEEL affected and unaffected firms are on average valued at 169 and 122 million US dollars respectively.

An estimate of the risk-free rate is required for the computation of implied volatilities. Following Barraclough *et al.* (2013) we use the Eurodollar nominal spot rate for one-day, seven-day, one-month, three-month, six-month and one-year intervals downloaded from DataStream. The interest rate for any maturity t is equal to the linear interpolation of the neighboring LIBOR maturities. We use actual dividends paid to proxy for expected dividends within the options life adjusted for stock splits. All stock announcements are extracted from TRTH, including dividends, stock splits and other capital adjusting announcements.

4.3.1 Sample attributes

Table 2 provides a summary of the data, with Appendix 1 indicating the allocation of each sample firm into affected or unaffected for the SOLAR and STEEL tariffs. The allocation is based on industry classification, with some manual intervention. As an example, technology, industrial and energy firms are more heavily affected by the SOLAR tariffs compared to pharmaceuticals or financial firms. For the STEEL tariffs, industrials and automobile manufacturers are assumed to be more reliant on steel compared to technology and entertainment firms. Some manual interventions were made in the "Food & Beverage" industry based on the reliance of aluminum in their products, Coca-Cola is a heavy user of aluminum in their beverages, however Mondelez International Inc is not. The number of firms in our sample for each announcement is limited by the number of firms having four or more traded call options on the announcement day with a delta between 0.4 and 0.5. We compare the option volume for the announcement day with the trading day on either side.

Table 2: Sample statistics

This table reports the trading volume of the options around the announcements discussed in this paper. Panel A includes the January 18, 2018 solar and washing machine tariff announcement, and Panel B includes the March 1, 2018 steel and aluminum tariff announcement. The Pre-announcement day, announcement day, and after announcement day sections represent the day before, during and after the announcement day respectively. The paired Wilcoxon rank sum test is used to test the difference in the volumes. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

	Pre-announcement day	Announcement day			Post-announcement day			
Trading volume	Mean	Mean		%Change	Mean		%Change	Number of firms
Panel A: Solar and washing machine tariff announcement								
All	74,710	61,470	**	-18%	72,471	**	18%	78
Unaffected	61,346	41,258		-33%	48,414		17%	48
Affected	96,093	93,809	*	-2%	110,962		18%	30
Panel B: Steel and aluminum announcement								
All	45,956	63,651	***	39%	59,293	**	-7%	77
Unaffected	55,570	76,358	***	37%	71,471	**	-6%	54
Affected	23,384	33,817	***	45%	30,702		-9%	23

Panel A of Table 2 contains the summary SOLAR option volumes and related statistics. Volume is significantly lower on announcement day compared to the previous day at the 5% level, with equally weighted mean volume being 18% lower. Affected firms experience a 2% decrease in volume which is statistically significant at the 10% level. Unaffected firms did not experience a significant change in volume. The preliminary result shows the effect of the SOLAR announcement was concentrated in these affected firms. Post-announcement day for the whole sample is 18% higher on the subsequent date, indicating a continuation of trading on this information.

Panel B of Table 2 contains the summary STEEL announcement option volumes and related statistics. Announcement day volumes for affected and unaffected firms are significantly higher compared to the prior day. The mean increase in volume for affected firms is 45%, with unaffected firms experiencing a 37% increase in volume. For affected firms there is an insignificant decrease in volume for the subsequent day, indicating that although trade is slowing, there is still a significant amount of activity.

4.4. Methodological framework

We adopt the framework of Hietala *et al.* (2003), Bhagat *et al.* (2005), Barraclough *et al.* (2013) and Borochin and Golec (2016) and write the stock price at time t (S_t) as a probability weighted average of the future stock price under the two future states of the tariff:

$$S_t = pS_{t+k}^S + (1-p)S_{t+k}^F \tag{1}$$

Where S_{t+k}^S and S_{t+k}^F is the future stock price if the tariff is implemented or is rejected at $t+k$ respectively. p represents the probability of the successful implementation of the new tariffs.

The value and news effect are based on Barraclough *et al.* (2013) and is calculated as follows:

Value effect	$S_{t+k}^S - S_{t+k}^F$	(2)
News effect	$S_{t+k}^F - S_{t-1}$	(3)

The value effect is equal to the difference in the stock price in the two future states. The news effect is equal to the impact of the additional information generated by the announcement. In order to progress, we need estimates for $S^S_{t+k}, S^F_{t+k}$and p. This can be accomplished through additional information provided by call options, with the value of a single call option being written as:

$$C_t(S_t; X_t) = p\hat{C}_t(S^S_{t+k}, \sigma^S_{t+k}; X) + (1-p)\hat{C}_t(S^F_{t+k}, \sigma^F_{t+k}; X) \quad (4)$$

Where $C_t(S_t; X_t)$ is the price of a call option at t with strike price X. $\hat{C}_t(S^S_{t+k}, \sigma^S_{t+k}; X)$ and $\hat{C}_t(S^F_{t+k}, \sigma^F_{t+k}; X)$ are theoretical option prices at $t+k$ for the two states of tariff outcome, considering a theoretical future stock volatility σ_{t+k}. Using the call option price alongside the stock price leaves five parameters that need to be estimated, including $S^S_{t+k}, S^F_{t+k}, p, \sigma^S_{t+k}$and σ^F_{t+k}. By using the stock price and at least four call options, a system of equations is used, and these parameters are numerically estimated.

We use call options with exercise prices that are in the money (Barraclough *et al.* 2013; Borochin & Golec 2016). Specifically, following Kelly *et al.* (2016), we use call options that have a delta satisfying $0.4 < \Delta < 0.5$. We filter out all options that have no open interest and have no trading volume on the announcement day.

4.5. Results

4.5.1 Solar and washing machine announcement

In this section we show the main results for the analysis conducted on the solar and washing machine tariff using the methodology outlined in Section 4. Table 3 contains the results.

Table 3: Solar panels and washing machines

This table reports the news and value effects of the January 18, 2018 solar component and washing machine tariff announcement. Probability is the probability that the announcement will come into effect. Value effect is the percentage difference between the theoretical price in the state of tariff success and tariff failure, we include an equal weighted and capitalisation weighted form. News effect is the percentage difference between the theoretical price in the state of the tariff not being incorporated and the price one day before announcement (January 17, 2018), we include an equal weighted and capitalisation weighted form. Market cap is the total market cap of all firms in the sample. pval is the p value associated with a t test. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

	No. firms	Probability (success)		pval vs 0.5	Market cap ($bil sum)	Value (equal weighted)		pval	Value ($bil sum)	Value (Cap weighted)		pval	News (cap weighted)		pval	News ($bil sum)	News (cap weighted)		pval	Return
All	78	58.95%	***	0.000	12,542	-4.37%	***	0.000	-546	-4.35%	***	0.000	2.70%		1.000	374	2.98%		1.000	-0.15%
Unaffected	48	57.07%	**	0.030	6,913	-3.40%	***	0.000	-199	-2.88%	***	0.001	1.86%		0.996	130	1.88%		0.993	-0.06%
Affected	30	61.96%	***	0.001	5,630	-5.91%	***	0.000	-347	-6.16%	***	0.000	4.05%		1.000	244	4.34%		1.000	-0.30%
Unaffected - Affected		-4.89%				2.51%	*			3.28%	*		-2.19%	*			-2.46%	*		

From Table 3, the probability of tariff success is approximately 60%. Failure of the tariff implementation depends on the oppositions ability to dissuade President Trump from this course of action. A failure probability of 40% indicates the opposition is unlikely to be successful. The difference between the probability for affected and unaffected firms is insignificant, however we can interpret the higher probability of success for affected firms to potentially indicate a different belief among traders of these firms.

Model estimates on the effect of the SOLAR tariff indicate a significant decrease in the value of sample firms of around -4.37%, with a dollar value of -546 billion dollars. Affected firms experience a greater effect with a loss of -5.91% on average. The tariffs will cause cost increases for firms with green initiatives, and since there aren't any solar producers in the S&P100, there are no expected value winners in our sample. Furthermore, if these tariffs are successful, it would increase the likelihood of further tariffs, changing the probability state weighted value of potentially all firms.

Overall, the news effect for these firms is insignificant, indicating a mixed reaction with limited new information. The average effect is still large however, with the market attributing a positive 2.70% news effect, and 4.05% for affected firms. This result indicates the news is seen as a positive step toward increasing US innovation as under these circumstances US solar innovation is partly shielded from international competition. Insulation of innovation from international competition allows US solar innovation to become valuable, which sends a signal for other US innovative industries. The result not being significant indicates limited new information, as these measures were part of President Trump's election promises.

4.5.2 Steel and aluminum announcement

In this section we show the main results for the analysis conducted on the steel and aluminum tariff using the methodology outlined in Section 4. Table 4 contains the results.

Model estimates for the effect of the steel tariff announcement tell a similar story. There is a significant negative value effect of -3.99%, with the total dollar value being -358 billion dollars. The value effect is lower for the STEEL tariff compared to the SOLAR tariff, which may be due to the firms already experiencing a value loss prior to the STEEL announcement.

Table 4: Steel and aluminum

This table reports the news and value effects of the March 1, 2018 steel and aluminum tariff announcement. Probability is the probability that the announcement will come into effect. Value effect is the percentage difference between the theoretical price in the state of tariff success and tariff failure, we include an equal weighted and capitalisation weighted form. News effect is the percentage difference between the theoretical price in the state of the tariff not being incorporated and the price one day before announcement (February 28, 2018), we include an equal weighted and capitalisation weighted form. Market cap is the total market cap of all firms in the sample. pval is the p value associated with a t test. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively

	No. firms	Probability (success)		pval vs 0.5	Market cap ($bil sum)	Value (equal weighted)		pval	Value ($bil sum)	Value (Cap weighted)		pval	News (cap weighted)		pval	News ($bil sum)	News (cap weighted)		pval	Return
Panel A																				
All	77	57.81%	***	0.003	11,969	-3.99%	***	0.005	-358	-2.99%	***	0.008	0.39%		0.748	-20	-0.17%		0.357	-1.31%
Unaffected	54	57.44%	**	0.018	9,171	-4.43%	**	0.021	-257	-2.80%	**	0.037	0.14%		0.589	-51	-0.56%		0.121	-1.39%
Affected	23	58.68%	*	0.087	2,798	-2.97%	**	0.012	-101	-3.62%	***	0.009	0.96%		0.774	31	1.12%		0.836	-1.14%
Unaffected - Affected		-1.24%				-1.46%				0.82%			-0.82%				-1.68%			
Panel B																				
Heavy steel user	16	58.03%		0.170	2,187	-5.00%	***	0.001	-107	-4.91%	***	0.004	2.52%		0.926	48	2.18%		0.941	-1.18%
Light steel user	7	60.15%		0.361	611	1.69%		0.871	6	0.99%		0.692	-2.59%	***	0.005	-16	-2.65%	**	0.017	-1.05%
Heavy - light		-2.11%				-6.69%	***			-5.90%	***		5.10%	**			4.83%	**		

Interestingly, the value effect is larger for unaffected firms than it is for affected firms. We posit that some affected firms are affected positively, and others are affected negatively depending on whether the steel tariffs are beneficial or detrimental to their business. For example, US steel producers, firms with a low dependence on steel products and firms with flexibility in their steel supply would be less affected than firms with inelastic steel requirements for whom steel products make up a large portion of their expenses. In order to separate out these firms, we investigate and split the firms affected by the steel tariffs into heavy and light steel users, where light steel users are either producers, or are less negatively affected by the steel tariff, and heavy steel users, who are more heavily affected by these tariffs. The segregation is performed based on media releases and articles regarding the effect of the tariff on specific firms. For firms with no media coverage, the effect on a closely related firm is used.

With this segregation, the effects can be more clearly seen with a larger and statistically significant negative effect for heavy steel users, and a positive but insignificant value effect for light steel users. It should be noted that there are only 7 firms in the light steel user group, making statistical tests difficult and unreliable. This positive effect on light steel users may indicate a redistribution of funds into firms that are still exposed to the same macroeconomic drivers but are affected in slightly different ways. This may also indicate a rebalancing within sector specific mutual fund portfolios.

Due to the sample being made up of firms of various sizes, it is instructive to measure the total dollar value effect against the total market cap. We do this in column “Value (Cap weighted)”. In this column the difference between affected and unaffected firms can be more clearly seen. Specifically, the affected firms have a value effect of -3.62%, whereas the unaffected firms have a value effect of -2.80%.

The news effect for the STEEL tariffs is insignificant for all segments except light steel users. This indicates the general lack of cohesion of thought as to the effects of these tariffs.

Table 5: Excluding confounding events

This table reports the news and value effects of the SOLAR and STEEL tariff announcement excluding firms with price sensitive confounding events surrounding the announcement date [-1,+1]. Probability is the probability that the announcement will come into effect. Value effect is the percentage difference between the theoretical price in the state of tariff success and tariff failure, we include an equal weighted and capitalisation weighted form. News effect is the percentage difference between the theoretical price in the state of the tariff not being incorporated and the price one day before, we include an equal weighted and capitalisation weighted form. Market cap is the total market cap of all firms in the sample. pval is the p value associated with a t test. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively

Solar results	No. firms	Probability (success)		pval vs 0.5	Market cap ($bil sum)	Value (equal weighted)		pval	Value ($bil sum)	Value (Cap weighted)		pval	News (equal weighted)		pval	News ($bil sum)	News (cap weighted)		pval	Return
All	67	57.23%	***	0.006	11,173	-4.15%	***	0.000	-483	-4.32%	***	0.000	2.30%		1.000	313	2.80%		1.000	-0.06%
Unaffected	41	56.06%	*	0.098	5,965	-3.38%	***	0.000	-183	-3.06%	***	0.001	1.67%		0.997	112	1.88%		0.989	0.04%
Affected	26	59.08%	**	0.016	5,208	-5.37%	***	0.000	-300	-5.76%	***	0.001	3.29%		1.000	201	3.85%		0.998	-0.22%
Unaffected - Affected	0	-3.02%				1.99%				2.70%			-1.62%	*			-1.97%	*		
Steel results																				
Panel A																				
All	74	57.54%	***	0.005	11,670	-4.14%	***	0.006	-358	-3.06%	***	0.008	0.46%		0.777	-16	-0.14%		0.384	-1.32%
Unaffected	52	57.62%	**	0.020	8,991	-4.52%	**	0.023	-252	-2.81%	**	0.039	0.16%		0.593	-52	-0.57%		0.118	-1.39%
Affected	22	57.36%		0.146	2,679	-3.25%	***	0.008	-105	-3.93%	***	0.007	1.18%		0.813	36	1.34%		0.871	-1.15%
Unaffected - Affected	0	0.26%				-1.26%				1.13%			-1.02%				-1.91%			
Panel B																				
Heavy steel user	16	58.03%		0.170	2,187	-5.00%	***	0.001	-107	-4.91%	***	0.004	2.52%		0.926	48	2.18%		0.941	-1.18%
Light steel user	6	55.58%		0.630	492	1.40%		0.795	2	0.41%		0.569	-2.40%	**	0.014	-12	-2.39%	**	0.042	-1.09%
Heavy - light	0	2.46%				-6.40%	**			-5.32%	**		4.91%	**			4.57%	**		

4.5.3 Robustness check: excluding confounding events

Firm specific events impacting stock and options prices can impact the results of this paper, as such we remove firms with confounding events as the source of price changes can be unclear. We find all earnings, dividend, and capital change announcements in the [-1, +1] region captured in Thomson Reuters Tick History and remove these firms from our sample. In total we remove 11 firms for the SOLAR and three firms for the STEEL announcement from our sample. The results are presented in Table 5.

We find results excluding firms with confounding events to be materially similar when these firms are included. The probability of tariff success is just below 60% with a negative and significant value effect and a positive but insignificant news effect. Overall our findings are robust.

4.5.4 Alternative methodology

A competing methodology in the literature is the intervening events method (Bhagat *et al.* 2005). Bhagat *et al.* (2005) investigates mergers and acquisitions and estimates a prior takeover probability using past takeovers, using competing bids as intervening events. Given the result in the previous section, and the strong opposition to Trump's executive powers, a probability of 60% seems reasonable for this methodology. We use the implementation date as the intervening event, with the implementation date having a tariff success probability of 100%.

Since we have a success price as measured at the implementation date, as well as an appropriate probability estimate, we can use Equation (1) to calculate the price under an unsuccessful state, which is the only missing value in the equation. The results are outlined in Table 6.

Table 6 indicates the value and news effects using the intervening events method. Consistent with the estimation results in Table 3 and Table 4, the value effect is negative and News effect is positive for both announcements. The magnitude using the intervening events method is larger compared to the original estimation, with a -10.5% value effect for SOLAR announcement and -10.1% value effect for the STEEL announcement. The results from splitting light and heavy steel users is consistent with the main results, however the negative value effect of the heavy steel users is smaller than the unaffected firms.

Table 6: Intervening events method

This table reports the news and value effects of the SOLAR and STEEL tariff announcement using the intervening events method. Probability is the probability that the announcement will come into effect. Value effect is the percentage difference between the theoretical price in the state of tariff success and tariff failure, we include an equal weighted and capitalisation weighted form. News effect is the percentage difference between the theoretical price in the state of the tariff not being incorporated and the price one day before, we include an equal weighted and capitalisation weighted form. Market cap is the total market cap of all firms in the sample. pval is the p value associated with a t test. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively

Solar results	No. firms	Probability (success)	Market cap ($bil sum)	Value (equal weighted)		pval	Value ($bil sum)	Value (Cap weighted)		pval	News (equal weighted)		pval	News ($bil sum)	News (cap weighted)		pval	Return
All	78	0.6	12,542	-0.105	***	0.000	-1,411	-0.113	***	0.000	0.062		1.000	841	0.067		1.000	-0.001
Unaffected	48	0.6	6,912	-0.104	***	0.000	-721	-0.104	***	0.000	0.062		1.000	436	0.063		1.000	-0.001
Affected	30	0.6	5,629	-0.108	***	0.001	-690	-0.123	***	0.001	0.061		0.999	404	0.072		0.999	-0.003
Unaffected - Affected				0.004		0.920	-30	0.018		0.920	0.000		0.988	31	-0.009		0.988	0.000
Steel results																		
All	77	0.6	11,968	-0.101	***	0.000	-1,381	-0.115	***	0.000	0.047		1.000	660	0.055		1.000	-0.013
Unaffected	54	0.6	9,170	-0.116	***	0.000	-1,176	-0.128	***	0.000	0.056		1.000	569	0.062		1.000	-0.014
Affected	23	0.6	2,797	-0.063	***	0.001	-205	-0.073	***	0.001	0.026		0.990	90	0.033		0.997	-0.011
Unaffected - Affected				-0.053	**	0.026		-0.055	***	0.007	0.030	**	0.039	478	0.030	**	0.010	-0.003
Panel B																		
Heavy steel user	16	0.6	2,186	-0.078	***	0.000	-167,684,386	-0.077	***	0.002	0.034		0.998	76	0.035		0.996	-0.012
Light steel user	7	0.6	610	-0.030		0.257	-37,621,621	-0.062		0.129	0.007		0.610	14	0.024		0.800	-0.011
Heavy - light				-0.048		0.337		-0.015		0.351	0.027		0.990		0.011		0.997	-0.001

These results confirm the results of our main estimation. Contrary to our main results, the intervening events method depends on suitable probability estimates and intervening events as measured after the initial announcement, with our main analysis only requiring data available on the announcement date.

4.6. Conclusion

In this study we examine the value and news effects associated with the solar and washing machine (SOLAR), and steel and aluminum tariff (STEEL) announcement by President Trump in early 2018. Since the outcome from the announcement is uncertain at the time the announcement was made, traditional event study methodologies are unsuitable as the price on announcement contains a probability parameter measuring the chance of the announcement coming into effect. To overcome this issue we use the recent methodology created by Barraclough *et al.* (2013) and Borochin and Golec (2016) to separate the probability, value and news effects. We find the tariff announcements had a negative value effect of approximately -546 and -358 billion dollars in the case of the SOLAR and STEEL announcement respectively, with directly affected firms experiencing a more negative effect. For both announcements the news effect is positive but not significant, indicating that while news of protectionist tariffs is well received, they are somewhat expected due to Trump's election promises.

Future research can use this same methodology on other potentially value affecting announcements made by President Trump to quantify and document the effect of this unconventional president on the US market.

4.7. Appendix

Firm list and allocation

This table contains the firms and their affected/not affected status as used in this study. STEEL represents whether the firm is affected (1) or not affected (0) by the steel and aluminum tariff announcement. SOLAR represents whether the firm is affected (1) or not affected (0) by the Solar and washing machine tariff announcement. HEAVY represents firms that are affected by the steel tariff, with the split indicating whether a firm is heavily affected (1) or not heavily affected (0), with this classification only applying to firms affected by the STEEL announcement.

	Ticker	Company name	TRBC Business Sector Code Description	STEEL	SOLAR	HEAVY
1	AAPL	Apple Inc	Technology Equipment	0	1	NA
2	ABBV	AbbVie Inc	Pharmaceuticals & Medical Research	0	0	NA
3	ABT	Abbott Laboratories	Healthcare Services	0	0	NA
4	AGN	Allergan Plc	Pharmaceuticals & Medical Research	0	0	NA
5	AIG	American International Group	Insurance	0	0	NA
6	AMGN	Amgen Inc	Pharmaceuticals & Medical Research	0	0	NA
7	AXP	American Express Company	Banking & Investment Services	0	0	NA
8	BA	Boeing Co	Industrial Goods	1	1	1
9	BAC	Bank of America Corp	Banking & Investment Services	0	0	NA
10	BIIB	Biogen Inc	Pharmaceuticals & Medical Research	0	0	NA
11	BLK	Blackrock Inc	Banking & Investment Services	0	0	NA
12	BMY	Bristol-Myers Squibb Co	Pharmaceuticals & Medical Research	0	0	NA
13	C	Citigroup Inc	Banking & Investment Services	0	0	NA
14	CAT	Caterpillar Inc	Industrial Goods	1	1	1
15	CELG	Celgene Corporation	Pharmaceuticals & Medical Research	0	0	NA
16	CHTR	Charter Communications Inc	Cyclical Consumer Services	0	0	NA
17	CL	Colgate-Palmolive Company	Personal & Household Products & Services	0	0	NA
18	CMCSA	Comcast Corporation	Cyclical Consumer Services	0	0	NA
19	COF	Capital One Finance Corp	Banking & Investment Services	0	0	NA
20	COP	Conocophillips	Energy - Fossil Fuels	1	1	0
21	COST	Costco Wholesale Corporation	Retailers	0	0	NA
22	CSCO	Cisco Systems Inc	Technology Equipment	0	1	NA
23	CVS	CVS Health CORP	Healthcare Services	0	0	NA
24	CVX	Chevron Corporation	Energy - Fossil Fuels	1	1	1
25	DIS	Walt Disney Co	Cyclical Consumer Services	0	0	NA
26	EMR	Emerson Electric Co	Industrial Goods	1	1	NA
27	F	Ford Motor Company	Automobiles & Auto Parts	1	0	1
28	FB	Facebook Inc	Software & IT Services	0	1	NA
29	FDX	FedEx Corporation	Transportation	1	1	1
30	FOXA	Fox Corp Class A	Cyclical Consumer Services	0	0	NA
31	GD	General Dynamics Corporation	Industrial Goods	1	1	1
32	GE	General Electric Company	Industrial Conglomerates	1	1	NA
33	GILD	Gilead Sciences Inc	Pharmaceuticals & Medical Research	0	0	NA
34	GM	General Motors Company	Automobiles & Auto Parts	1	0	1
35	GS	Goldman Sachs Group Inc	Banking & Investment Services	0	0	NA

36	HAL	Halliburton Company	Energy - Fossil Fuels	1	1	0
37	HD	Home Depot Inc	Retailers	0	0	NA
38	HON	Honeywell International Inc	Industrial Conglomerates	1	1	0
39	IBM	International Business Machines	Software & IT Services	0	1	NA
40	INTC	Intel Corporation	Technology Equipment	0	1	NA
41	JNJ	Johnson & Johnson	Pharmaceuticals & Medical Research	0	0	NA
42	JPM	JPMorgan Chase & Co	Banking & Investment Services	0	0	NA
43	KMI	Kinder Morgan Inc	Energy - Fossil Fuels	1	1	NA
44	KO	The Coca-Cola Co	Food & Beverages	1	0	1
45	LLY	Eli Lilly & Co	Pharmaceuticals & Medical Research	0	0	NA
46	LMT	Lockheed Martin Corporation	Industrial Goods	1	1	1
47	LOW	Lowe’s Companies Inc	Retailers	0	0	NA
48	MA	Mastercard Inc	Software & IT Services	0	1	NA
49	MCD	Mcdonald's Corp	Cyclical Consumer Services	0	0	NA
50	MDLZ	Mondelez International Inc	Food & Beverages	0	0	NA
51	MDT	Medtronic Plc	Healthcare Services	0	0	NA
52	MET	Metlife Inc	Insurance	0	0	NA
53	MMM	3M Company	Industrial Conglomerates	1	1	0
54	MO	Altria Group Inc	Food & Beverages	1	0	0
55	MRK	Merck & Co inc	Pharmaceuticals & Medical Research	0	0	NA
56	MS	Morgan Stanley	Banking & Investment Services	0	0	NA
57	MSFT	Microsoft Corporation	Software & IT Services	0	1	NA
58	NFLX	Netflix Inc	Software & IT Services	0	1	NA
59	NKE	Nike Inc	Cyclical Consumer Products	0	0	NA
60	NVDA	Nvidia Corporation	Technology Equipment	0	1	NA
61	ORCL	Oracle Corporation	Software & IT Services	0	1	NA
62	OXY	Occidental Petroleum Corporation	Energy - Fossil Fuels	1	1	0
63	PEP	PepsiCo Inc	Food & Beverages	1	0	1
64	PFE	Pfizer Inc	Pharmaceuticals & Medical Research	0	0	NA
65	PG	Procter & Gamble Co	Personal & Household Products & Services	0	0	NA
66	PM	Philip Morris International Inc	Food & Beverages	0	0	NA
67	PYPL	PayPal Holdings Inc	Software & IT Services	0	1	NA
68	QCOM	Qualcomm Inc	Technology Equipment	0	1	NA
69	SBUX	Starbucks Corporation	Cyclical Consumer Services	0	0	NA
70	SLB	Schlumberger Limited	Energy - Fossil Fuels	1	1	0
71	SO	Southern Co	Utilities	1	1	1
72	T	AT&T Inc	Telecommunications Services	1	0	1
73	TGT	Target Corporation	Retailers	0	0	NA
74	TXN	Texas Instruments Incorporated	Technology Equipment	0	1	NA
75	UNH	UnitedHealth Group Inc	Healthcare Services	0	0	NA
76	UNP	Union Pacific Corporation	Transportation	1	1	1
77	UPS	United Parcel Service Inc	Transportation	1	1	1
78	UTX	United Technologies Corporation	Industrial Goods	1	1	NA

79	V	Visa Inc	Software & IT Services	0	1	NA
80	VZ	Verizon Communications Inc	Telecommunications Services	1	0	1
81	WBA	Walgreen Boots Alliance Inc	Food & Drug Retailing	0	0	NA
82	WFC	Wells Fargo & Co	Banking & Investment Services	0	0	NA
83	WMT	Walmart Inc	Food & Drug Retailing	0	0	NA
84	XOM	Exxon Mobil Corporation	Energy - Fossil Fuels	1	1	1

Chapter 5: Concluding remarks

This thesis, in a collection of three essays investigates how financial decisions can be impacted through new data and methodologies. The last few decades have seen astronomical increases in data production and capture, allowing for data driven insights and decision making. This has required new methodologies to accurately take large quantities of data and produce information that can be understood and used for decision making. This thesis investigates three data sources and uses new methodologies to uncover economically meaningful results.

In Chapter 2 we investigate price discovery in the global gold market using high-frequency data for the period 1997 to 2014. The gold market is of interest to both researchers and industry due to its immense size and economic importance.

Our first major finding is that the US futures market contributes more to price discovery compared to the London OTC spot market, despite being ten times smaller. This result indicates that market structure and instrument type is of greater importance than market size and liquidity for market efficiency. Our second key finding reinforces the importance of market structure on the process of price discovery. A change from a floor-based system to the nearly 24-hour, fully electronic low latency platform is associated with a considerable decrease in noisiness of prices in this market, but no increase in the relative speed with which the market reflects new information. Our third key finding is that several factors affect the location of price discovery. Price discovery shares vary substantially at both daily and intraday levels.

Our findings carry important implications for market design. We show that market structure is of greater importance to price discovery than market size and liquidity. Changes to market structure can have opposite effects on the speed at which prices reflect new information and accuracy with which they reflect the information (amount of noise), as is the case in the market structure change that we examine. Market designers can improve the efficiency of their markets by carefully considering the structures they implement, as this can have greater impact than methods focused simply on increasing turnover and participation.

In Chapter 3 we use borrower bank and card transaction data and study the effectiveness of transactional data in predicting the probability of default for owner occupier mortgages. We find borrower financial transactions can be used to develop an accurate indication of financial position and general behavior. These measures can be used to more accurately separate defaulting and non-defaulting borrowers than traditional default drivers.

We further show default risk changes according to a borrower's change in circumstances. If a borrower experiences a change in income, they are more likely to default. Borrowers who are currently renovating, or actively maintaining their home are less likely to default, indicating their outlook on both property value, as well as their ability to service the loan. Borrowers who withdraw large amounts of cash or purchase clothing are more and less likely to default, respectively. Our data and methodology greatly improve the model fit of a traditional variable model and provide a timelier prediction of credit risk.

This study is of interest to regulators, financial institutions, and mortgage holders as it outlines some behaviors that may indicate or lead to financial distress for borrowers. Effectively managing this financial stress can lead to lower default rates, less stress, and lower collection costs for financial institutions. Benefits can also accrue to homeowners if bank intervention can be earlier with more positive long-term outcomes.

In Chapter 4 we examine the value and news effects associated with the solar and washing machine, and steel and aluminum tariff announcement by President Trump in early 2018. Since the outcome from the announcement is uncertain at the time the announcement was made, traditional event study methodologies are unsuitable as the price on announcement contains a probability parameter measuring the chance of the announcement coming into effect. To overcome this issue we use the recent methodology created by Barraclough *et al.* (2013) and Borochin and Golec (2016) to separate the probability, value and news effects. We find the tariff announcements had a negative value effect of approximately 546 and 358 billion dollars in the case of the SOLAR and STEEL announcement respectively, with directly affected firms experiencing a more negative effect. For both announcements the news effect is positive but not significant, indicating that while news of protectionist tariffs is well received, they are somewhat expected due to Trump's election promises.

The results in this chapter are instructive for the effect of future protectionist tariff announcements. Our results suggest, in line with many economists, that in most circumstances tariffs are welfare reducing. It is important to use available data, with correct statistical methods, to assess the potential effect of policies. These effects can be estimated using empirical methodologies, one example of which is utilized in this chapter.

Avenues of future research in the use of data and methods for financial decision making are limitless, and as technology continues to develop, old problems can be quantified, and new problems will arise. With respect to the papers in this thesis, there are several avenues available for future research. The second chapter provides a methodology which can be used to assess the impact of other microstructural changes in financial markets. Financial markets are continually being updated and changed, with the final impact of changes not known prior. The third chapter shows how a new data source can be used to improve an old area of study such as mortgage default modelling. Data similar to that which is used in Chapter 3 can be used to assess the effect of individual behavior on areas other than credit risk. Finally, Chapter 4 shows one example of measuring the impact of new policy decisions, before the policy is implemented. This methodology can be used to assess other policy decisions and can be used to draw conclusions on the announcement date.

References:

Agarwal, S., Green, R.K., Rosenblatt, E., Yao, V., 2015a. Collateral pledge, sunk-cost fallacy and mortgage default. Journal of Financial Intermediation 24, 636-652

Agarwal, S., Pan, J., Qian, W., 2015b. The composition effect of consumption around retirement: Evidence from Singapore. American Economic Review 105, 426-431

Agarwal, S., Qian, W., 2014. Consumption and debt response to unanticipated income shocks: Evidence from a natural experiment in Singapore. American Economic Review 104, 4205-4230

Agarwal, S., Qian, W., 2017. Access to home equity and consumption: Evidence from a policy experiment. Review of Economics and Statistics 99, 40-52

Aggarwal, R., Lucey, B.M., 2007. Psychological barriers in gold prices? Review of Financial Economics 16, 217-230

Akcigit, U., Ates, S.T., Impullitti, G., 2018. Innovation and trade policy in a globalized world. National Bureau of Economic Research

Amromin, G., Chakravorti, S., 2009. Whither Loose Change? The Diminishing Demand for Small-Denomination Currency. Journal of Money, Credit and Banking 41, 315-335

Anand, A., Gatchev, V.A., Madureira, L., Pirinsky, C.A., Underwood, S., 2011. Geographic proximity and price discovery: Evidence from NASDAQ. Journal of Financial Markets 14, 193-226

Aoki, K., Proudman, J., Vlieghe, G., 2004. House prices, consumption, and monetary policy: a financial accelerator approach. Journal of Financial Intermediation 13, 414-435

Arango, C., Huynh, K.P., Sabetti, L., 2015. Consumer payment choice: Merchant card acceptance versus pricing incentives. Journal of Banking & Finance 55, 130-141

Ates, A., Wang, G.H., 2005. Information transmission in electronic versus open-outcry trading systems: An analysis of US equity index futures markets. Journal of Futures Markets: Futures, Options, and Other Derivative Products 25, 679-715

Aydın, D., 2018. Consumption Response to Credit Expansions: Evidence from Experimental Assignment of 45,307 Credit Lines. Washington University working paper

Baillie, R.T., Booth, G.G., Tse, Y., Zabotina, T., 2002. Price discovery and common factor models. Journal of Financial Markets 5, 309-321

Bajo, E., Barbi, M., 2018. Financial illiteracy and mortgage refinancing decisions. Journal of Banking & Finance 94, 279-296

Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research, 159-178

Bank for International Settlements, 2011. Triennial Central Bank Survey. Report on global foreign exchange market activity in 2010. URL http://www.bis.org/press/p101201.htm

Barraclough, K., Robinson, D.T., Smith, T., Whaley, R.E., 2013. Using option prices to infer overpayments and synergies in M&A transactions. The Review of Financial Studies 26, 695-722

Batten, J.A., Ciner, C., Lucey, B.M., 2010. The macroeconomic determinants of volatility in precious metals markets. Resources Policy 35, 65-71

Baur, D.G., Lucey, B.M., 2010. Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. Financial Review 45, 217-229

Baur, D.G., McDermott, T.K., 2010. Is gold a safe haven? International evidence. Journal of Banking & Finance 34, 1886-1898

Bhagat, S., Dong, M., Hirshleifer, D., Noah, R., 2005. Do tender offers create value? New methods and evidence. Journal of Financial Economics 76, 3-60

Bhagwati, J., Ramaswami, V.K., 1963. Domestic distortions, tariffs and the theory of optimum subsidy. Journal of Political Economy 71, 44-50

Bian, X., Lin, Z., Liu, Y., 2018. House price, loan-to-value ratio and credit risk. Journal of Banking & Finance 92, 1-12

Bohl, M.T., Salm, C.A., Schuppli, M., 2011. Price discovery and investor structure in stock index futures. Journal of Futures Markets 31, 282-306
Borochin, P., Golec, J., 2016. Using options to measure the full value-effect of an event: Application to Obamacare. Journal of Financial Economics 120, 169-193
Bounie, D., François, A., 2006. Cash, check or bank card? The effects of transaction characteristics on the use of payment instruments. Telecom Paris Economics and Social Sciences Working Paper No. ESS-06-05 (2006).
Boyd, N., Locke, P., 2014. Price discovery in futures and options markets. Journal of Futures Markets 34, 853-867
Bricker, J., Bucks, B., 2016. Negative home equity, economic insecurity, and household mobility over the Great Recession. Journal of Urban Economics 91, 1-12
Browning, M., Crossley, T.F., 2008. The long-run cost of job loss as measured by consumption changes. Journal of Econometrics 145, 109-120
Browning, M., Crossley, T.F., 2009. Shocks, stocks, and socks: Smoothing consumption over a temporary income loss. Journal of the European Economic Association 7, 1169-1192
Cabrera, J., Wang, T., Yang, J., 2009. Do futures lead price discovery in electronic foreign exchange markets? Journal of Futures Markets: Futures, Options, and Other Derivative Products 29, 137-156
Cai, J., Cheung, Y.L., Wong, M.C., 2001. What moves the gold market? Journal of Futures Markets: Futures, Options, and Other Derivative Products 21, 257-278
Caminschi, A., Heaney, R., 2014. Fixing a leaky fixing: Short-term market reactions to the London PM gold price fixing. Journal of Futures Markets 34, 1003-1039
Campbell, J.Y., Cocco, J.F., 2015. A Model of Mortgage Default. The Journal of Finance 70, 1495-1554
Campbell, T.S., Dietrich, J.K., 1983. The determinants of default on insured conventional residential mortgage loans. The Journal of Finance 38, 1569-1581
Capie, F., Mills, T.C., Wood, G., 2005. Gold as a hedge against the dollar. Journal of International Financial Markets, Institutions and Money 15, 343-352
Carmichael, B., Coën, A., 2018. Real estate as a common risk factor in bank stock returns. Journal of Banking & Finance 94, 118-130
Chan, S., Haughwout, A., Hayashi, A., Van der Klaauw, W., 2016. Determinants of mortgage default and consumer credit use: the effects of foreclosure laws and foreclosure delays. Journal of Money, Credit and Banking 48, 393-413
Chen, H., Huynh, K.P., Shy, O., 2019. Cash versus card: Payment discontinuities and the burden of holding coins. Journal of Banking & Finance 99, 192-201
Choi, H.-S., Hong, H., Scheinkman, J., 2014. Speculating on home improvements. Journal of Financial Economics 111, 609-624
Christie–David, R., Chaudhry, M., Koch, T.W., 2000. Do macroeconomics news releases affect gold and silver prices? Journal of Economics and Business 52, 405-421
CME Group, 2014a. Gold futures contracts specs. URL http://www.cmegroup.com/trading/metals/precious/gold_contract_specifications.html
CME Group, 2014b. Gold futures daily settlement procedure. URL http://www.cmegroup.com/trading/metals/files/daily-settlement-procedure-gold-futures.pdf
CME Group, 2014c. International Incentive Program (IIP) questions and answers. URL http://www.cmegroup.com/company/membership/files/InternationalIncentiveProgramQA.pdf
CME Group, 2014d. Types of membership. URL http://www.cmegroup.com/company/membership/types-of-membership-nymex.html
CME Group, 2014e. The world's leading electronic platform. CME GLOBEX. URL http://www.cmegroup.com/globex/files/globexbrochure.pdf
CME Group Media Room, 2014. NYMEX to offer COMEX metals side by side, miNYs beginning on December 3. URL http://cmegroup.mediaroom.com/index.php?s=43anditem=954
Cox, D.R., Snell, E.J., 1989. Analysis of Binary Data, Second Edition. Taylor & Francis.

Crook, J., Banasik, J., 2012. Forecasting and explaining aggregate consumer credit delinquency behaviour. International Journal of Forecasting 28, 145-160
Crook, J., Bellotti, T., 2010. Time varying and dynamic models for default risk in consumer loans. Journal of the Royal Statistical Society: Series A (Statistics in Society) 173, 283-305
David, B., Abel, F., Patrick, W., 2016. Debit card and demand for cash. Journal of Banking & Finance 73, 55-66
David, H., Dorn, D., Hanson, G.H., 2013. The China syndrome: Local labor market effects of import competition in the United States. American Economic Review 103, 2121-68
DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics, 837-845
Demyanyk, Y., Van Hemert, O., 2009. Understanding the subprime mortgage crisis. The Review of Financial Studies 24, 1848-1880
Dinopoulos, E., Segerstrom, P., 1999. The dynamic effects of contingent tariffs. Journal of International Economics 47, 191-222
Dolatabadi, S., Nielsen, M.Ø., Xu, K., 2015. A fractionally cointegrated VAR analysis of price discovery in commodity futures markets. Journal of Futures Markets 35, 339-356
Elder, J., Miao, H., Ramchander, S., 2012. Impact of macroeconomic news on metal futures. Journal of Banking & Finance 36, 51-65
Elul, R., Souleles, N.S., Chomsisengphet, S., Glennon, D., Hunt, R., 2010. What "triggers" mortgage default? American Economic Review 100, 490-494
Fama, E.F., Fisher, L., Jensen, M.C., Roll, R., 1969. The adjustment of stock prices to new information. International Economic Review 10, 1-21
Felbermayr, G., Jung, B., Larch, M., 2015. The welfare consequences of import tariffs: A quantitative perspective. Journal of International Economics 97, 295-309
Foote, C.L., Gerardi, K., Willen, P.S., 2008. Negative equity and foreclosure: Theory and evidence. Journal of Urban Economics 64, 234-245
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, 20-33
Fuangkasem, R., Chunhachinda, P., Nathaphan, S., 2014. Information Transmission Among World Major Gold Futures Markets: Evidence From High Frequency Synchronous Trading Data. Journal of US-China Public Administration 11, 255-269
Gan, J., 2010. Housing wealth and consumption growth: Evidence from a large panel of households. Review of Financial Studies 23, 2229-2267
Gerardi, K., Herkenhoff, K.F., Ohanian, L.E., Willen, P.S., 2017. Can't pay or won't pay? unemployment, negative equity, and strategic default. The Review of Financial Studies 31, 1098-1131
GFMS Ltd., 2012. Gold survey 2012. GFMS Ltd., London
Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. Journal of Business & Economic Statistics 13, 27-35
Goodman, L.M., 2011. The Asylum: The Renegades who Hijacked the World's Oil Market. Harper Collins.
Grossman, G.M., Helpman, E., 1990. Comparative Advantage and Long-Run Growth. American Economic Review 80, 796-815
Guiso, L., Sapienza, P., Zingales, L., 2013. The determinants of attitudes toward strategic default on mortgages. The Journal of Finance 68, 1473-1515
Han, J., Linnenluecke, M.K., Pan, Z.T., Smith, T., 2019. The wealth effects of the announcement of the Australian carbon pricing scheme. Pacific-Basin Finance Journal 53, 399-409
Hasbrouck, J., 1995. One security, many markets: Determining the contributions to price discovery. The Journal of Finance 50, 1175-1199
Hautsch, N., Hess, D., Veredas, D., 2011. The impact of macroeconomic news on quote adjustments, noise, and informational volatility. Journal of Banking & Finance 35, 2733-2746
Helms, A.C., 2003. Understanding gentrification: an empirical analysis of the determinants of urban housing renovation. Journal of Urban Economics 54, 474-498

Herkenhoff, K., 2012. Job loss, defaults, and policy implications. working paper
Hietala, P., Kaplan, S.N., Robinson, D.T., 2003. What Is the Price of Hubris? Using Takeover Battles to Infer Overpayments and Synergies. Financial Management 32, 5
Hillier, D., Draper, P., Faff, R., 2006. Do precious metals shine? An investment perspective. Financial Analysts Journal 62, 98-106
Kartashova, K., Tomlin, B., 2017. House prices, consumption and the role of non-Mortgage debt. Journal of Banking & Finance 83, 121-134
Kelly, B., Pástor, Ľ., Veronesi, P., 2016. The price of political uncertainty: Theory and evidence from the option market. The Journal of Finance 71, 2417-2480
King, G., Zeng, L., 2001. Logistic regression in rare events data. Political Analysis 9, 137-163
LBMA, 2014. Market tools. URL http://www.lbma.org.uk/market-tools
Lin, H.-N., Chiang, S.-M., Chen, K.-H., 2008. The dynamic relationships between gold futures markets: evidence from COMEX and TOCOM. Applied Financial Economics Letters 4, 19-24
Lin, J.Y., Wang, X., 2018. Trump economics and China–US trade imbalances. Journal of Policy Modeling 40, 579-600
Linnenluecke, M.K., Chen, X., Ling, X., Smith, T., Zhu, Y., 2016. Emerging trends in Asia-Pacific finance research: A review of recent influential publications and a research agenda. Pacific-Basin Finance Journal 36, 66-76
Lucey, B.M., Larkin, C., O'Connor, F.A., 2013. London or New York: where and when does the gold price originate? Applied Economics Letters 20, 813-817
Miller, S., 2015. Information and default in consumer credit markets: Evidence from a natural experiment. Journal of Financial Intermediation 24, 45-70
Mocetti, S., Viviano, E., 2017. Looking behind mortgage delinquencies. Journal of Banking & Finance 75, 53-63
Montgomery, C., 1992. Explaining home improvement in the context of household investment in residential housing. Journal of Urban Economics 32, 326-350
Morrison, K., 2006. COMEX considers parallel trading. In: Financial Times
Murray, S., 2011. Loco London liquidity survey. The Alchemist 63, 9-10
Norden, L., Weber, M., 2010. Credit line usage, checking account activity, and default risk of bank borrowers. Review of Financial Studies 23, 3665-3699
O'Callaghan, M., 1991. The Structure and Operation of the World Gold Market. IMF, New York
Osang, T., Pereira, A., 1996. Import tariffs and growth in a small open economy. Journal of Public Economics 60, 45-71
Puri, M., Rocholl, J., Steffen, S., 2017. What do a million observations have to say about loan defaults? Opening the black box of relationships. Journal of Financial Intermediation 31, 1-15
Putniņš, T.J., 2013. What do price discovery metrics really measure? Journal of Empirical Finance 23, 68-83
Qi, M., Yang, X., 2009. Loss given default of high loan-to-value residential mortgages. Journal of Banking & Finance 33, 788-799
Riddiough, T.J., 1992. Equilibrium mortgage default pricing with non-optimal borrower behavior.
Rivera-Batiz, L.A., Romer, P.M., 1991. International trade with endogenous technological change. European Economic Review 35, 971-1001
Rosenberg, J.V., Traub, L.G., 2009. Price discovery in the foreign currency futures and spot market. The Journal of Derivatives 17, 7-25
Sherman, E.J., 1982. Gold: a conservative, prudent diversifier. The Journal of Portfolio Management 8, 21-27
Stavins, J., 2018. Consumer preferences for payment methods: Role of discounts and surcharges. Journal of Banking & Finance 94, 35-53
VISA, 2017. VISA merchant data standards manual. URL usa.visa.com
WTO, 2019. URL https://www.wto.org/english/tratop_e/tariffs_e/tariffs_e.htm

Xu, X.E., Fung, H.-G., 2005. Cross-market linkages between US and Japanese precious metals futures trading. Journal of International Financial Markets, Institutions and Money 15, 107-124

Yan, B., Zivot, E., 2010. A structural analysis of price discovery measures. Journal of Financial Markets 13, 1-19