AUSTRALIAN INTERCHANGE FEE REGULATION

a regulation in search of market failure

About the Authors

Sinclair Davidson is Professor of Institutional Economics in the School of Economics, Finance and Marketing at RMIT University, and a Senior Fellow at the Institute of Public Affairs. He is a regular contributor to public debate. His opinion pieces have been published in The Age, The Australian, Australian Financial Review, Sydney Morning Herald, and Wall Street Journal Asia. Sinclair has published in academic journals such as the *European Journal of Political Economy*, *Journal of Economic Behavior and Organization*, and the *Cato Journal*.

Jason Potts is Professor of Economics in the School of Economics, Finance and Marketing at RMIT University, an Australian Research Council Future Fellow, and also an Adjunct Fellow at the Institute of Public Affairs. He has written five books and over 70 articles on the theory of economic evolution. His work focuses on how entrepreneurship and innovation drive economic growth and development. Jason has published in academic journals such as *Journal of Economic Behavior and Organisation*, *Journal of Institutional Economics*, and *Economic Affairs*.

Executive Summary

The Reserve Bank of Australia has been a world leader in interchange fee regulation. In this paper we suggest that this regulatory intervention has been based on wishful thinking at best and represents a failure to understand the actual working of the market economy.

In short, the Reserve Bank of Australia engaged in an extensive regulatory intervention based on poor theory, and no empirical evidence. Theory has not provided an unambiguous indication of market failure, and there is no empirical evidence to support the notion of monopoly pricing – other than a vague notion that interchange fees were "excessive". What the Reserve Bank identified as being "externality" any fair minded observer would label "gains from trade".

We argue that interchange fees are the outcome of an efficient bargaining process given that banks and consumers, and banks and merchants form long term relationships with each other. For as long as there is competition in the banking sector and competition in the retail sector, the interchange fee itself is subject to competitive pressure.

There is no market failure and no economic justification for government intervention. The $13 billion "saving" to merchants that the Reserve Bank identifies following its regulatory reform is simply a redistribution away from consumers (and banks) towards merchants. The Reserve Bank assumes that the saving has been passed onto consumers, but cannot provide any evidence to support that hypothesis.

It is not at all clear that consumers have benefited from interchange fee regulation. To the contrary is likely that consumers are worse off – while merchant fees have declined, so too have the benefits of using credits while the costs (including the interest rate premium over the cash) have increased.

1. Introduction

Ronald Coase famously argued that "if an economist finds something – a business practice of one sort or other – that he does not understand, he looks for a monopoly explanation".[1] So it is with credit card interchange fees. As we will demonstrate intellectual confusion has lead to the phenomenon of interchange fees being misdiagnosed as being a monopoly problem leading to inappropriate policy intervention. Following George Stigler's path breaking analysis of the US Security and Exchange Commission he claimed that financial regulation was "founded upon prejudice and … reforms are directed by wishfulness".[2] In our opinion, Australian regulation of interchange fees should be placed into the same category: reforms initiated by ignorance and anti-bank prejudice.

A 2000 joint study by the Reserve Bank and Australian Consumer and Competition Commission concluded *inter alia*:[3, 4]

II Credit card interchange fees are significantly above levels suggested by cost-based methodologies and contribute to margins of revenues over average costs of around 39 per cent for card issuers. …

IV 'No surcharge' rules in credit card schemes prevent purchasers from confronting the cost of this payment instrument vis-à-vis lower cost payment instruments such as debit cards. It means that other consumers subsidise credit cardholders and financial institutions which are card scheme members. An alternative arrangement would have merchants exercising discretion to charge customers prices that are net of the cost of the payment instrument, and add a surcharge to cover that cost.

V Competition in credit card issuing and acquiring is limited by restrictions on access to credit card schemes. Excluding all institutions other than authorised deposit-takers from access to acquiring, in particular, is difficult to justify on risk grounds.

…

Interchange fees are set by card issuers and acquirers at 'one step removed' from the cardholders and merchants who ultimately bear these fees through transaction charges or through the general cost of goods and services. Users therefore do not have a direct influence on the pricing of card payment services but must rely on their financial institutions to represent their interests. As a consequence, the price signals and competitive responses that would be expected to put pressure on margins in card payment networks have not worked effectively. These difficulties are reinforced by restrictions on access to the card networks, both explicit and informal, and by the 'no surcharge' rules in credit card schemes.

The regulatory concerns then relate to excessive pricing, price fixing, abuse of market power, the creation of barriers to entry, increased consumers prices generally, and excessive use of credit cards relative to alternate payment methods. The fact that end-users do not observe the interchange fee

[1] Coase, 1972 [1988], pg. 67.
[2] Stigler, 1964, pg. 142.
[3] Reserve Bank and Australian Consumer and Competition Commission, 2000, pg. 73 – 74.
[4] Hereinafter RBA – ACCC.

makes it opaque, and less prone to competitive pressure. All these arguments suggest that regulatory intervention can easily correct these apparent market flaws and result in improve economic performance.

As a result of these concerns and the apparent ease at which corrective action could be undertaken Australia embarked on a program of regulatory intervention. In this paper, we argue that the regulatory concerns were over-sold and rely on a faulty understanding of the underlying economic principles. There is no case for intervention.

The remainder of the paper is set out as follows. In section 2 we explain what an interchange fee is. In section three we critique the Australian arguments for regulatory intervention and show data as to consequences of that intervention. In section 4 we provide alternative, non-monopoly but efficiency enhancing, explanations for interchange fees.

2. What is an interchange fee?

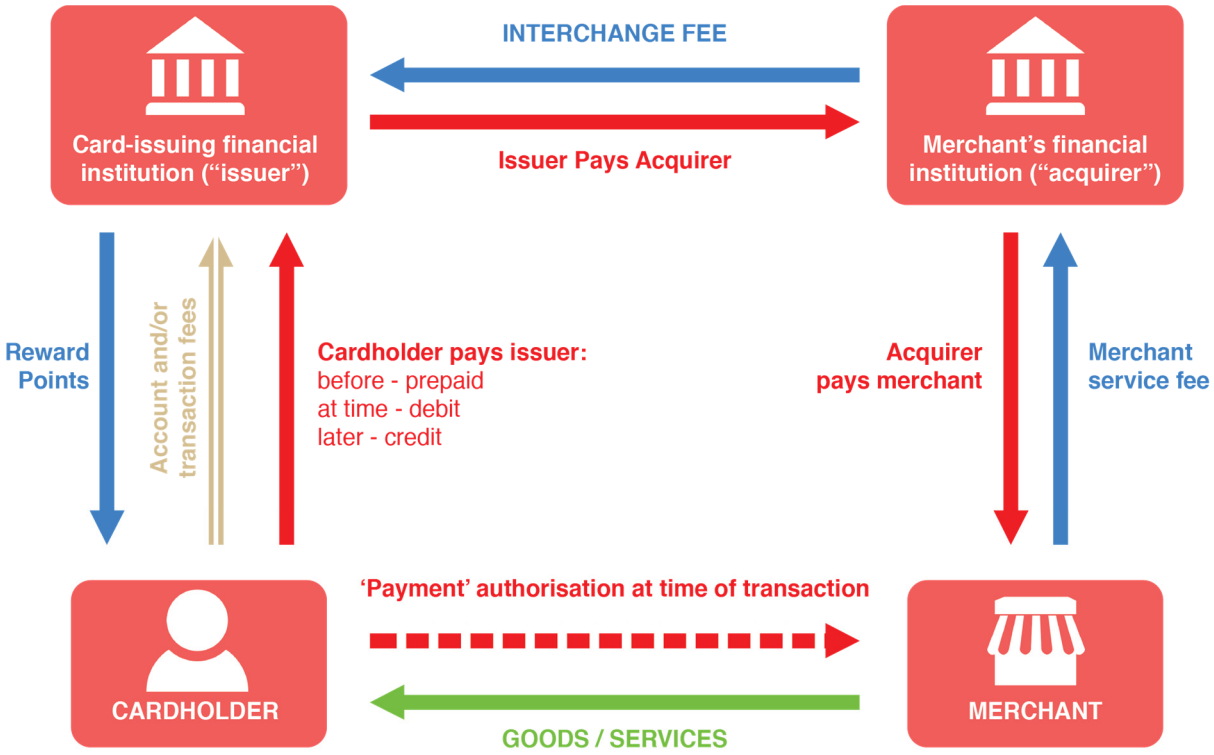
Interchange fees are fees that banks charge each other as a result of their respective clients entering into a credit card transaction. Figure 1 below shows how the Reserve Bank of Australia depicts an interchange fee. The figure shows a stylised (four-party system) example of transactions involving a credit card.

The consumer (cardholder) purchases goods and services from a merchant and pays for the goods and service using a credit card. Underpinning that particular transaction is two prior transactions and a long-term relationship. The first prior transaction is between the consumer and their own financial institution whereby they acquire a credit card and pay a fee for the credit card use. As part of that transaction the consumer may or may not earn reward points as a function of the credit card usage. The second prior transaction is between the merchant and their financial institution whereby the merchant pays a fee to their financial institution in order to process credit card payments. The long-term relationship is between the two financial institutions that provide financial services to the consumer and merchant.

When the merchant sells goods and services to the consumer, the consumer authorises his financial institution to pay a sum of money to the merchant. The merchant passes the authorisation to his financial institution which then collects the money from the consumer's financial institution and pays the merchant. Finally the consumer's financial institution gets paid once the consumer pays off their outstanding credit card balance.[5]

[5] What is missing from the Reserve Bank explanation is that the consumer's bank as extended credit to the consumer while immediately paying the merchant's bank. The risk of non-payment is borne by the consumer's bank.

Figure 1: RBA depiction of an interchange fee



Source: RBA 2015, pg. 6

The interchange fee is a fee paid by the merchant's bank to the consumer's bank.

Neoclassical economists describe this type of arrangement as being a "two-sided" market. Two-sided markets consist of two sets of end-users who have their needs met simultaneously. In this case the credit card example the two sets of users include consumers who use the credit (card holders) and merchants who accept the card. The card itself is useless if either consumers will not use the card, or merchants will not accept the card in payment. Credit card companies, or associations, have a joint maximisation problem: maximising the number of consumers who will use the card and maximising the number of merchants that will accept the card. The incentives facing consumers and merchants being somewhat different Hayashi and Weiner argue that the interchange fee "an instrument that networks can use to achieve a desired balance of cardholder usage versus merchant acceptance across the two sides of the market … In other words, interchange fees are a mechanism that can be used to transfer revenues from one side of the market to the other to generate the desired level of card activity."

There are two issues of importance.

- The direction the interchange fee flows in.
- The magnitude of the interchange fee.

In most credit card systems the interchange fee flows from the merchant side of the transaction towards the consumer side of the transaction.[6] This implies that in some economies consumers require more of an inducement to hold and use credit cards than merchants need to accept those cards. To argue that this relationship is somehow inefficient is to argue that consumers have monopoly power over merchants. While it is true that merchants are subject to consumer sovereignty few economists, or policy makers, would argue that consumers have monopoly power over merchants, or if they did that this monopoly power should be restrained.

There is a rich *academic theoretical* literature that considers the magnitude of the interchange fee. In their 2006 survey paper, Hayashi and Weiner categorise the theoretical literature into one of four categories.

1. Assumptions about the (credit card) networks. Are the networks themselves competitive, or monopolies?
2. Assumptions about financial institutions. Are financial institutions competitive or monopolies?
3. Assumptions about consumers and merchants. Do merchants have monopoly power? Do consumers have single cards or multiple cards?
4. Other factors that might be important. What network rules are in place? No-surcharge rules? Honour all card rules?

Recall that the regulatory concern relating to credit card interchange fees is that the fees themselves were opaque, excessive, and encouraged excessive usage of credit cards relative to other payments mechanisms.

With a rich theoretical literature, including contributions from the 2014 economics laureate Jean Tirole, we might expect that clear unambiguous theoretical results could inform real world observations and shed light on the need, if any, for regulatory intervention. That, however, is not the case. For example, Katz (2001) reports that monopolistic networks with no-surcharge rules and reward points will result in excessive credit card use. That result appears to be consistent with the regulatory concerns. But credit card networks are not monopolistic. Studies that assume competitive networks have conflicting results. Rochet and Tirole (2002) show that if networks are competitive and consumers hold more than one card that interchange fees are not affected. But if consumers do not hold more than one card that merchants reduce acceptance of cards and interchange fees fall. In the same paper, however, they also show that even if networks are monopolistic as long as financial institutions are competitive (in issuing credit cards) that interchange fees will fall. Several other papers show similar mixed results. Interchange fees may either be higher or lower depending on the assumptions made in the analysis.

Importantly for our purposes, changing assumptions about network rules such as the no-surcharge rule or honour all cards rule has differing results. Again interchange fees could be higher or lower depends on a host of other factors or assumptions being made in the analysis.

[6] This is not always the case. In some markets the interchange fee has gone from the consumer side of the transaction to the merchant side, and in some economies the interchange fee is zero.

After an extensive survey of the literature Hayashi and Weiner conclude:[7]

What one comes away with after surveying this rich theoretical literature is an appreciation for the many factors that may affect interchange fees. Even a single factor may impact interchange fees differently, depending on other factors. Determining the actual impact of such variables is, in the end, an empirical question.

What that implies is that the theoretical results are not robust to changes in the underlying assumptions in the modelling. The 1990 economics laureate Merton Miller has claimed that there is nothing more practical than good theory. By that benchmark the theoretical analysis of interchange fees is simply not good theory as it give no practical guidance to what we might expect to observe in the real world.

In a 2003 paper Rochet and Tirole had come to the same conclusion, summarising the theoretical academic literature as follows:[8]

On the contrary, recent academic work concurs to establishing that there is no systematic bias in the IFs selected by cooperative networks: there is no reason to think that privately optimal IFs are higher or lower than socially optimal ones. Misunderstanding the economics of the problem and imposing cost-based regulation could impose substantial distortions in the industry.

They are even more damning that Hayashi and Weiner. Rochet and Tirole claim, quite correctly as we will argue below, that the very nature of the economic problem at hand has been misunderstood.

3. The Australian literature

Rochet and Tirole establish the basis for public intervention in markets as being a two-fold process:[9]

> The standard approach to public intervention in industries involves two steps:
>
> (1) the theoretical identification of a serious market failure and the validation of its empirical relevance,
>
> (2) the identification of the least distortionary way of addressing the market failure and a check that the remedy will not be worse than the illness.

As we have shown above, the very first step of that process has not been achieved. There is no theoretical basis for regulation of interchange fees. Rochet and Tirole are clear – the problem is a misunderstanding of the economics. In this section we highlight those misunderstandings in the Australian literature.

[7] Hayashi and Weiner, 2006, pg. 88.
[8] Rochet and Tirole, 2003, pg. 71.
[9] Rochet and Tirole, 2003, pg. 70.

The Australian literature on interchange fees consists of a joint report by the RBA – ACCC, a series of papers by Joshua Gans and Stephen King[10], and a more recent 2015 Reserve Bank of Australia Issues paper. In this section, we mostly focus our attention on the work undertaken by the Reserve Bank.

The RBA – ACCC report provides a description of credit card networks as per figure 1. It then describes how networks provide benefits to users (both consumers and merchants) as they increase in size i.e. more consumers hold a particular card and/or more merchants accept that particular card. Rather than considering an increase in network size as an increase in the size of the market and therefore any benefits flowing from that increase as being the gains from trade, the RBA – ACCC report instead views the benefits as being an externality.[11] This, in our opinion, constitutes a methodological error. Gains from trade constitute a benefit of the market mechanism, while externalities arise from market failure.

In this particular case the argument is that a network can generate positive externalities for users (suggesting that it should increase in size), but negative externalities for non-users (suggesting that networks can become too big). This possibility occurs if and when the merchant has monopoly power and can pass their service fees (including the interchange fee, see figure 2 below) onto consumers. At this point the interchange fee could be increased and result in greater private benefits to cardholders but higher prices to non-card holders. Given a somewhat non-standard definition of efficiency, "A payment network is said to operate efficiently if the net benefits it provides to society are being maximised", the RBA – ACCC study is able to argue that credit card networks may be too large in Australia. [12] Definitions of efficiency would normally suggest that an institution or process was meeting stated objectives at least possible cost. The argument here results in the proposition that increased competition to expand the network could result in increasing prices if merchants have some monopoly power.

The problem being exacerbated, the RBA – ACCC claim, by the fact that cardholders and merchants "are not involved in determining the interchange fee".[13] As we argue below, that statement is not strictly speaking true. It is correct to say that the interchange fee is not established in a spot market, but to argue that cardholders and merchants are not involved in overall price determination in a network is simply incorrect.

Nonetheless in the early 2000s Australia embarked on a series of regulatory interventions. The Reserve Bank of Australia announced its intention to introduce a series of reforms in August 2002. See table 1 for a time-line of reforms.[14]

Hayashi and Weiner are blunt in their assessment of the literature and regulation in Australia: "None of the models appears to closely fir the Australian market over a large number of parameters".[15] In other words there is no theoretical basis to support the introduction of regulation in Australia.

---

[10] Gans and King, 2001, 2002, 2003a, 2003b, 2003c.
[11] RBA – ACCC, 2000, pg. 24.
[12] RBA – ACCC, 2000, pg. 27.
[13] RBA – ACCC, 2000, pg. 28.
[14] In this paper we are primarily interested in credit card interchange fees, but include other reforms for completeness.
[15] Hayashi and Weiner, 2006, pg. 100.

Table 1: A time line of payment reforms

Date	Reform
October 200	Joint RBA – ACCC study published
December 2001	RBA consultation document released
August 2002	Intention to reform announced
January 2003	No Surcharge Rule eliminated
July 2003	Interchange fees capped
January 2004	Access regime modified
February 2004	Debit card reform (Visa)
September 2004	Debit card reform (MasterCard)
April 2006	Debit card reforms announced
July 2006	Debit card reforms implemented
November 2006	Common cost-based Interchange fee Benchmark introduced
January 2007	Honour all card rule abolished

Source: Authors, RBA 2015

Two Australian academics, Joshua Gans (now at Toronto University) and Stephen King (now at Monash University) have published a series of theoretical papers looking at interchange fees and regulatory concerns in credit card markets. It is fair to say that their views, while in favour of regulation, are nuanced. Overall their view is that the no-surcharge rule should be eliminated and as a result the interchange fee would become irrelevant. There is no need then to both eliminate the no-surcharge rule and regulate interchange fees.

The Gans and King analysis is predicated on resolving what they refer to as being "the inefficiency". They define an efficient transaction as follows:[16]

If a credit card transaction was efficient then it would probably be implemented if the customer and merchant as joint consumers and the issuer and acquirer as joint suppliers all negotiated over that transaction.

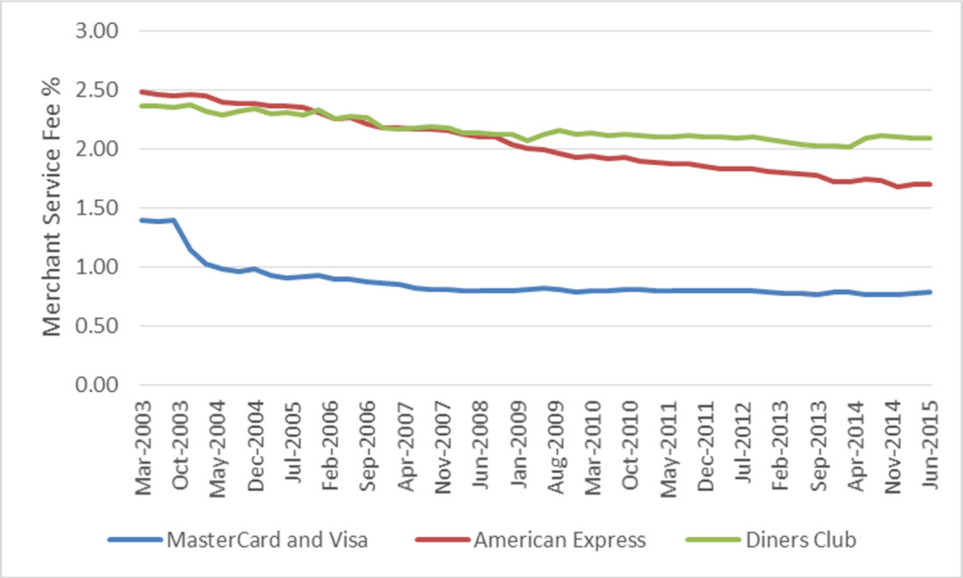
They refer to this description of a transaction as being Coasian bargaining after the economics laureate Ronald Coase.[17] They are making, at least, two errors at this point. First they are characterising only spot market transactions as possibly being efficient. Second they are ignoring the efficiency gains that can come about by entering into long-term relationships. We discuss this in greater detail in the next section. For our purposes here it is important to note that *the* inefficiency that Gans and King analyse is *an assumption based on a methodological error*.

They then canvass three possible "solutions" to their "inefficiency". The first solution involves horizontal integration – the two financial institutions merge into one (converting a four party credit card system into a three party credit card system). This is how American Express and Diner's Club are organised. In practice, however, the costs associated with those two providers tend to be higher than those

[16] Gans and King, 2001, pg. 99.
[17] This, of course, is a (common) mischaracterisation of Coase 1960.

associated with four-party systems (see exhibit 1). Alternatively a no-surcharge rule could resolve the inefficiency, or the existence of interchange fees could resolve the inefficiency.

Exhibit 1: Merchant Service Fees across four-party systems and three-party systems



Source: RBA Statistics

Gans and King are of the opinion that in the absence of a no-surcharge rule that interchange fees are competitively neutral.[18] The RBA – ACCC was concerned that excessively high interchange fees would distort consumer preferences towards excessive usage of credit cards relative to other payment mechanisms. A consequence of this possibility is that cash paying consumers pay too much for their goods and services and effectively "cross-subsidise" credit card paying consumers. Rather than have regulators set prices, Gans and King prefer regulators to eliminate the no-surcharge rule allowing merchants to charge differential prices (if the market will bear a price differential) depending on payment mechanism. They sum up:[19]

In the absence of a no surcharge rule, cooperative setting of interchange fees cannot have any anticompetitive effect.

Even in the presence of a no surcharge rule, the setting of interchange fees only creates competitive concerns if there is inadequate retail level competition.

Overall Gans and King consider the no-surcharge rule and the interchange fee as substitutes and argue that eliminating the no-surcharge rule makes regulating the interchange fee redundant. Overall, they doubted that the RBA interventions would result in many benefits.[20]

[18] Gans and King 2003a.
[19] Gans and King, 2003a, pg. 39.
[20] Gans and King, 2003c, pg. 472.

In summary, our analysis casts doubt on the benefits that will be created by the RBA's credit card reforms. While allowing surcharging makes sense, it is not certain that the regulated approach to interchange fees adopted by the RBA will lead to lower costs of transacting.

While we believe the Gans and King analyses are methodologically flawed it is interesting to note that they argue the interchange is competitively neutral. Of course, the RBA does not agree with assessment.

The RBA 2015 issues paper seems to suggest that its regulatory interventions are been successful. It restates unproven regulatory concerns as having been fact. For example,[21]

Competition between the schemes had, if anything, created upward – not downward – pressure on these fees. The higher the interchange fee paid to card issuers, the greater their incentive to issue the cards of a scheme and the larger the subsidies that can be paid to cardholders to encourage use of those cards. At least up to some limit, merchants appear unable to resist the high merchant service fees that result, typically finding it difficult to decline acceptance of cards given the risk of losing sales.

Whether or not competition resulted in increased interchange fees and increased merchant service fees (resulting in downward pressure on merchant profit margins) is an empirical question. If the evidence to validate that view exists, it is not in the public domain. It is true that interchange fee regulation did lead to a decline in merchant services fees, but as the RBA admits:[22]

*It is **impossible** – given the imprecision in any econometric model of consumer price inflation – **to measure exactly** how these reductions in merchant service fees have flowed through into prices for consumers.*

The RBA do report, however, that the reduction in merchant service fees since the regulatory intervention has been some $13 billion. They assume that those "savings" have been passed onto consumers claiming, "it seems reasonable to assume that they have mostly flowed through to lower retail prices for consumers".[23] Yet the RBA provides no reason why it would not be equally reasonable to assume that the $13 billion flows mostly to the merchants' profit margins. Indeed profit is something that is curiously missing from the entire RBA analysis.

We are told, for example, "competition in well-established payment card networks can lead to the perverse result of increasing the price of payment services to merchants (and thereby leading to higher retail prices for consumers)".[24] It simply never occurs to the RBA that, alternatively, increased costs to merchants could result in reduced profit margins. Much the same as the economic incidence of taxation is determined by the market, so too the economic incidence of costs is determined by the market.

It is important to note that the $13 billion is not a saving to the economy. It is simply a redistribution. If that money had been paid in interchange fees it would have been shared between consumers, in the

[21] RBA, 2015, pg. 4.
[22] RBA, 2015, pg. 23 (emphasis added).
[23] RBA, 2015, pg. 23.
[24] RBA, 2015, pg. 7.

form of reduced fees and loyalty programs, and their financial institutions. At best the RBA argument is that the $13 billion is being shared by merchants and consumers.

In addition, the RBA appears to be ignorant of standard business practices such as the "cash discount". It writes, "the consumer typically decides which means of payment is tendered and used in a transaction".[25] Yet merchants and consumers often bargain over price and over payment method. The cash discount is a very common mechanism to induce consumers to switch payment method. This is an astonishing oversight for the RBA given that it assumes the alternative payment mechanism to credit cards is a cash payment.

In summary, the RBA engaged in an extensive regulatory intervention based on poor theory and no empirical evidence. Theory has not provided an unambiguous indication of market failure, and there is no empirical evidence to support the notion of monopoly pricing – other than a vague notion that interchange fees are "excessive".

While we have other criticisms of the RBA approach – for example, we suspect the regulatory interventions were protectionist measures designed to support the local eftpos system – those arguments are beyond the current paper.

In March 2006, the Melbourne Business School hosted a Payment Systems conference discussing the interchange fee regulations in Australia.[26] Jean-Charles Rochet (of Rochet and Tirole fame) presented at that conference and made a number of predictions:[27]

First predicted consequences of a reduction in interchange fees:

- *increase in cardholders fees,*
- *decrease in merchants fees,*
- *reduction of the profit of issuers,*
- *increase in the profit of acquirers.*

Reduction in interchange fees likely to decrease the share of card payments (maybe after a delay). Ambiguous impact on consumer demand and consumer surplus:

- *Merchants may* ***decrease retail prices*** *(small?)*
- ***Transaction costs for consumers increase*** *(less convenient to use cards)*

Most important consequences of a reduction in interchange fees are medium to long term:

- *Issuing is likely to become more concentrated and less efficient*
- *Issuers may be tempted to bypass the regulation of interchange fees (socially inefficient)*

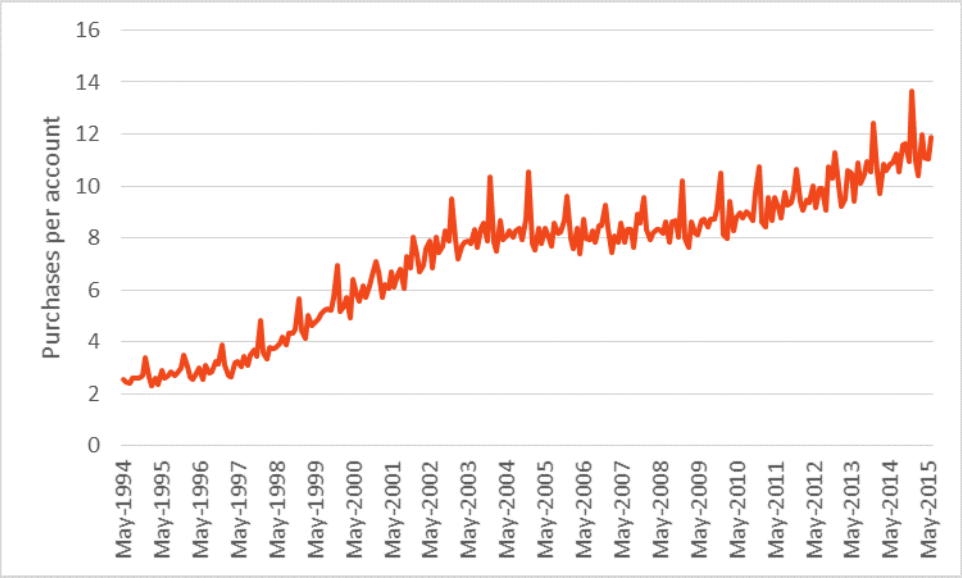
While it is not possible to test all of these predictions – it is possible to test some of them. What is particularly noteworthy, however, is that Rochet clearly identifies that profitability can and will be

[25] RBA, 2015, pg. 8.
[26] Papers available at http://web.archive.org/web/20060613224511/http://www.mbs.edu/payments_system/
[27] Emphasis original.

impacted by regulatory change, yet the RBA fails to discuss that issue. Rather the RBA focusses on consumer price changes, something that Rochet suggests will be small.

It is clear from the data that there was some impact in the credit card market following the RBA's regulatory intervention. We show that consistent with Rochet's predictions the advantages of using credit cards declined and the benefits associated with using credit cards declined. In Exhibit 2 we calculate the average number of transactions per credit card account using RBA data.[28]

Exhibit 2: Transactions per card



Source: RBA Statistics, Author calculations.

There is a very clear turning point in the data following the RBA's initial regulatory interventions. The growth in credit card transactions plateaus for nearly six years. Clearly the advantages associated with using credit cards declined.

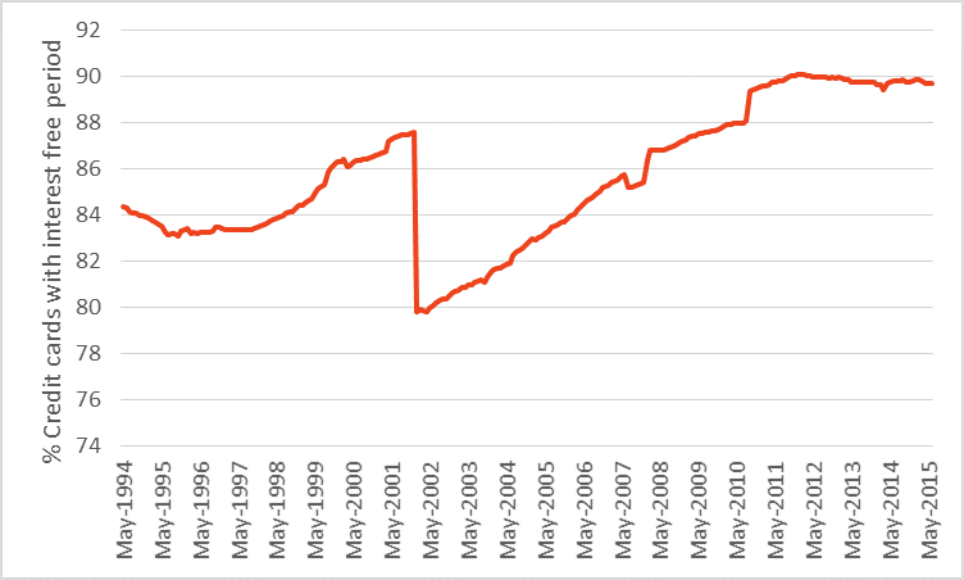
Similarly the benefits of using credit cards declined too. In Exhibit 3 we show the proportion of cards that had an interest free period.[29] Looking at the exhibit, the result is quite stark. A sudden decline from 86.7% in December 2001 to 79.8% in January 2002 is a massive change. While those dates do not quite line up with the actual regulatory timeline set out in table 1, it does immediately follow the publication of an RBA consultation document into the Australian credit card market. If we were to assume that financial institutions and consumers correctly anticipated the RBAs intentions then it is plausible to imagine that they would modify their behaviour before the regulatory intervention.

At the same time Rochet had predicted that issuing would become more concentrated. The RBA provides market share data for credit card schemes but indicates that one of the original three schemes

[28] Number of credit and charge card purchase transactions divided by Number of credit and charge card accounts.
[29] Number of personal credit card accounts with an interest-free period divided by Number of credit and charge card accounts.

that it initially regulated, Bankcard, closed in January 2007. At the same the domestic payments scheme eftpos has lost market share too.

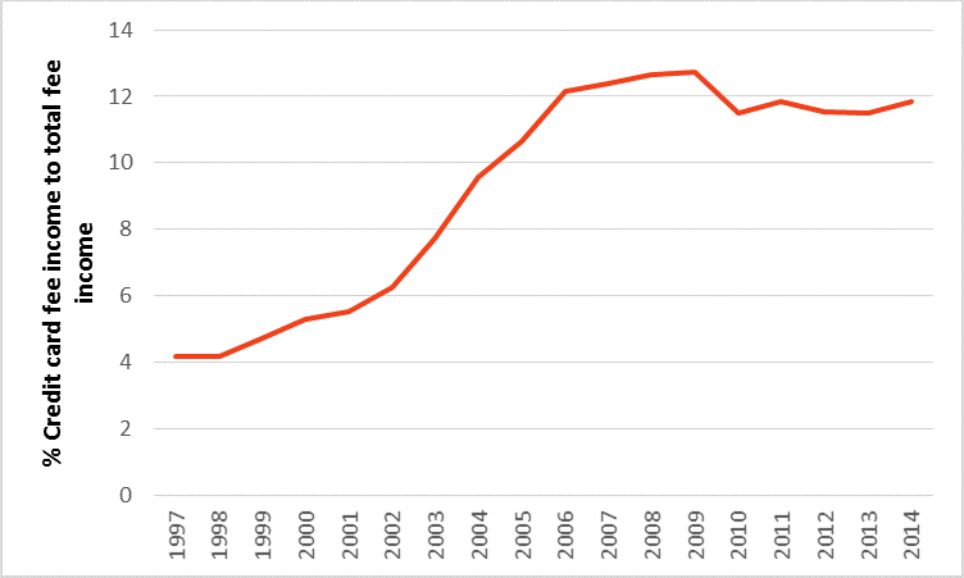
Exhibit 3: Proportion of credit card accounts with an interest free period



Source: RBA Statistics, Author calculations.

In Exhibit 4 we show the proportion of bank fee income from credit cards as a percentage of total bank fee income. It is clear over the period the RBA was introducing its regulations that fee income from credit cards accelerated as percentage of total bank fee income.

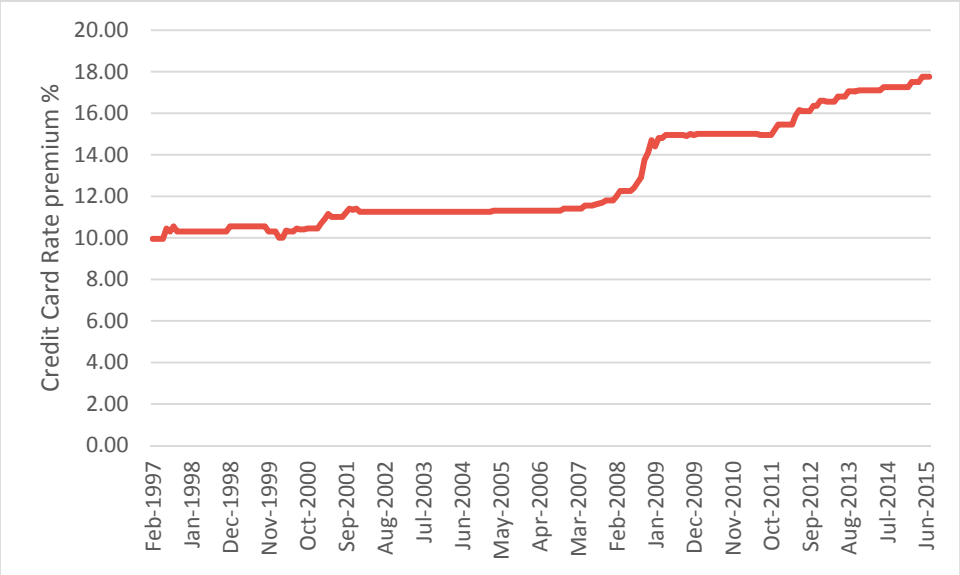
Exhibit 4: Credit card fee income to total fee income



Source: RBA Statistics, Author calculations.

Finally we show the credit card (standard) rate premium over the cash rate in Exhibit 5. Between December 2000 and December 2001 there is a 95 basis point increase in the credit card interest rate premium over the cash rate. In the context of the subsequent global financial crisis and risk-rerating that has occurred over the past few years, that increase is small. Nonetheless it is clear that interest rates charged by financial institutions moved in anticipation of regulatory change.

Exhibit 5: Credit card premium over Cash Rate



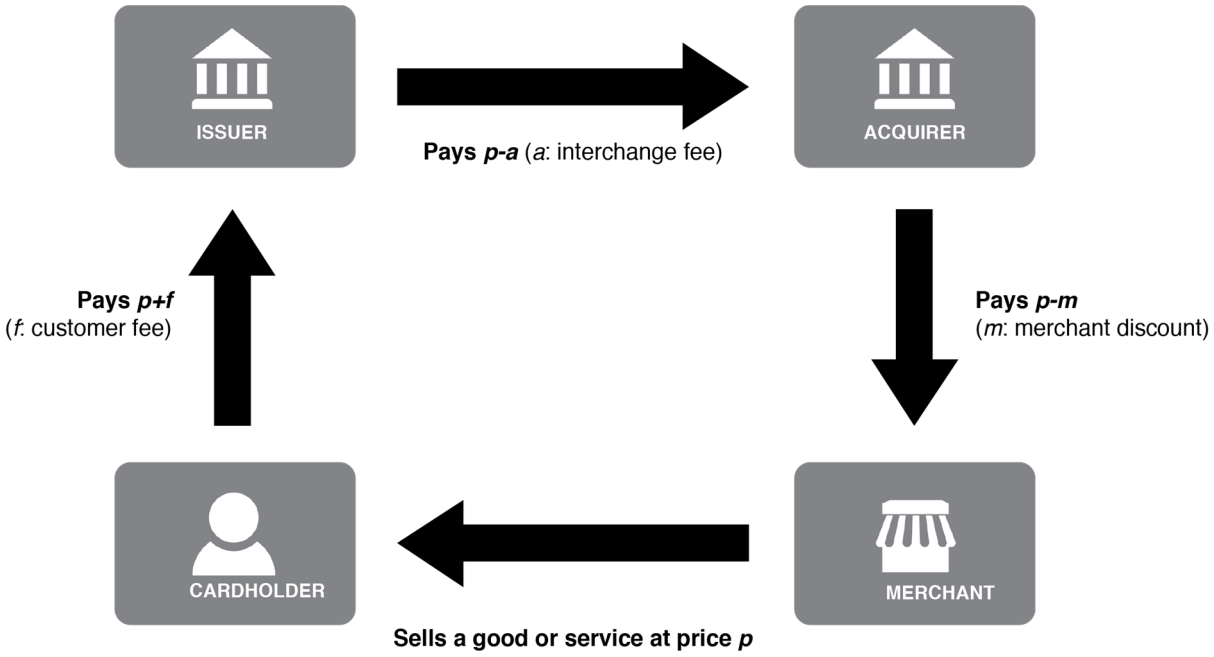
Source: RBA Statistics, Author calculations.

Consistent with Rochet's predictions, the RBA regulatory intervention has resulted in consumers paying more for their credit cards in the form of interest and increasing the fee income of banks while the benefits of the cards declined. The usage of credit cards relatively declined. All that for the $13 billion saving to merchants that the RBA identifies – yet the RBA is uncertain as to what actually happened to that money. They assume that it was passed onto consumers, but cannot know for sure. In addition, they are unable to point to any actual decreases in consumer prices following their intervention.

4. Alternative perspectives

We believe that the Reserve Bank has failed to understand the problem at hand. To see the issue more clearly consider not their exposition of the interchange fee as shown in figure, but rather Rochet and Tirole's exposition that we reproduce in figure 2.

Figure 2: Rochet and Tirole depiction of an interchange fee



Source: Rochet and Tirole (2003: 74)

This depiction shows the net cash flows in the various relationships. Again the consumer (cardholder) buys goods and services from the merchant. The consumer then pays the price (p) and a net fee to his financial institution. The consumer's financial institution then pays the price (p) less the interchange fee (a) to the merchant's financial institution who then pays the merchant the price (p) less their own net fee. This depiction of the issue makes very plain that if both financial institutions are to remain profitable that m > a. The merchant pays the interchange fee. Of course, this is not surprising. The interchange fee exists to rebalance the relationships within the two-sided market. In a competitive market for financial services, the interchange fee would be used to reduce the net consumer fee for credit cards. It is also unsurprising then that retail associations have led the charge against interchange fees. After all it is cost of doing business to them and reduces the profitability of their businesses. The subsequent regulation of the market is then well explained by the 1981 economics laureate George Stigler's theory of regulatory capture.

However, the basic issue is not one of monopoly exploitation, which has thus far been the guiding regulatory impulse that Stigler criticises, but rather is one of efficient contracting in the shadow of what 2009 economics laureate Oliver Williamson (1973) called the Fundamental Transformation that occurs in consequence of transactions that require both parties to make idiosyncratic investments – transforming ex ante competition into an ex post bilateral monopoly – that can subsequently give rise to opportunism.

The credit payments system is not and cannot ever be an interlinked series of anonymous spot markets exchanging financial commodities because the information asymmetries and moral hazards inherent in these exchanges require the parties to the transactions to make idiosyncratic investments (also known as asset specificity) that bind them into a bilateral monopoly – i.e. the fundamental transformation – in

which quasi-rents[30] are only secured through mechanisms to inhibit opportunism by aligning incentives to long term relational contracting.

The interchange fee, we argue, has evolved as an efficient governance mechanism to achieve this outcome without requiring horizontal integration – i.e. collapsing the four party payments system into a three-party payments system, and the associated losses of technical and information efficiency and competition that would imply. Banks need to make transaction specific investments in acquiring information about the properties of customers and merchants, the value of which – the quasi-rent – is realised through a long term relation.

4.1. Argument 1: The interchange fee represents an efficient institutional mechanism, not monopoly exploitation

Alternatives to collective setting of interchange fees, varying from bilateral negotiation to government-regulated cost-based fees, all have serious drawbacks in terms of generating excessive transactions costs, failing to internalize external benefits and costs, and distorting incentives.

Chang and Evans (2000: 461)

The existence of the interchange fee at what appears to be both a fixed and high level has been criticized by competition regulators because of its seeming departure from what would be expected in a perfectly competitive market. Among competition authority regulators, this is widely taken to be prima facie evidence of collusive price fixing and monopoly exploitation.

In an institutionally frictionless world of zero transaction costs, perfect rationality, perfect information, and zero uncertainty, any such fixed fee structure collectively agreed upon by competitors that seemed to generate permanent uncontestable flows of what would appear to be (natural) monopoly rents would certainly appear to be evidence of collusive monopoly exploitation. In this version of the story, the monopoly aspect of these rents are attributed to high entry costs owing to strong network effects on payments platforms.

In consequence, banking and competition regulators around the world have sought price caps on bank interchange fees (Schmalensee 2002). In Australia, this was reduced from 0.95% to 0.55% in 2003 (Europe Economics 2014: 27-32). These regulatory imposed fee caps are allegedly justified because they restrain anti-competitive behaviour and therefore benefit consumers.

Not only is there no evidence for this supposed regulatory benefit (ATA & IAEP 2015), but we argue that the economic theory behind it is also flawed. What it neglects is the adapted efficiency of the contractual and governance structure of the economic organization of payments systems and consumer finance.

[30] Klein et al (1978), pgs 289 – 307.

The argument we make (expanding on the work of Chang and Evans 2000) is that the interchange fee, as it has emerged and developed around the world over many decades, is an efficient governance outcome in a largely private ordering of mostly long term relational contracting between consumers, issuing banks, acquiring banks and merchants, all operating in the context of uncertainty, opportunism and asset specificity (Williamson 1985).

There are two specific aspects that we seek to highlight, both of which point to the fact that these are non-standard exchanges, and that the particular institutional and contractual features of the overall economic organization that depart from an Arrow-Debreu zero-transaction cost and complete markets model – i.e. the interchange fee – most likely reflects efficient contractual governance adaptations to these particular aspects of the exchange situation.

(1) The four-party exchange involves different types of contractual relationships, only one of which (between customer and merchant) is typically a spot-market transaction. The other three that have banks at one or more ends are typically long-term relational contracts. These involve complex contractual agreements that trade-off risks from uncertainty, opportunism, and asset specificity. The conditions of the spot market will be considerably shaped by the agreements made in the other three long run relational contract markets.

(2) The default payments model is assumed to be cash, which is assumed to be costless as a two-party-exchange between consumer and merchant. The four-party credit exchange relation is assumed to be more costly because of the additional services offered in the interbank payments and processing network that benefit both consumers (by extending finance) and merchants (by facilitating payments, screening credit-worthiness, covering credit risk). Both consumers and merchants benefit from these services and are willing to pay for these services. However, cash is also costly to both consumers and merchants (carry cost, risk, opportunity cost) and thus both will be willing to pay to use an alternative payments technology that mitigates these costs. Yet in a pure exchange spot market, merchants will only accept cash because to accept credit requires them to assume the costs of screening or of a long-term relationship that exposes them to consumer opportunism. However, by leveraging off the long-term relations established in the interbank payments networks, merchants can become indifferent at some fee margin between cash and credit transactions in the spot market, thus maximizing the overall transaction value by accepting all bids.

Our central argument then, as informed by transaction cost economics and the New Institutional theory of the firm (Williamson 2002), is that the various structures of fees that we observe in the long-term relation contracts that banks intermediate are most likely to represent an efficient bargaining outcome to arrive at stable long term relational contracts, given the various risks associated with opportunism and asset specificity, and are therefore not prima facie evidence of monopoly rent extraction.

The spot market between consumer and merchant is likely to be efficient when effective governance institutions in the long-term credit networks and payments systems emerge. These are facilitated by the inter-banking system, at the core of which is the interchange fee.

In consequence, regulatory attempts to treat these fees as if they were the result of collusive rent-extraction by seeking to constrain them within a price ceiling can risk harming an otherwise efficient system of institutional adaptation through long-run relational contracting to specific governance problems associated with uncertainty and transactions costs in the supply of consumer finance and payments systems (Balto 2000, Chang and Evans 2000).

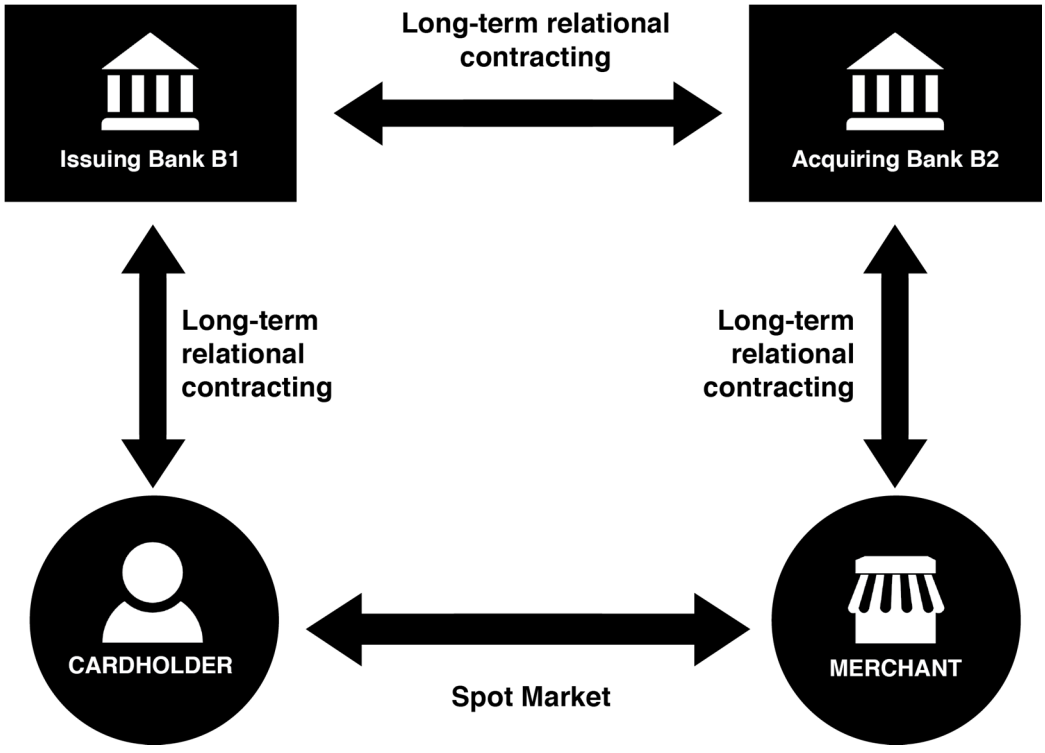
Models of the four-party and two-party payments systems

In a simple model of economic coordination, all exchanges take place in spot markets between firms (which in this model are hierarchical organizations whose boundaries are determined by the technology of production). In such a world, payments networks and consumer finance would be modelled as a natural monopoly (because of scale economies and network effects) such that the most efficient form of economic organization would be a single monopoly firm – call it The Bank. All consumers and all merchants would be customers of The Bank. The Bank would levy a fee across consumers and merchants, but the incidence of which would ultimately fall on consumers either directly or through higher prices as a function of the substitution margin with cash. An interchange fee would simply be an internal aspect of the firm's cost accounting. The total price The Bank charges would likely be regulated.

But under competition in retail payments networks, consumer banking and finance, and merchant banking we expect there will be multiple banks and that the boundaries of banks and financial services firms will depend upon specialization, competences and capabilities, often tied to specific assets (including reputational assets and context specific knowledge). This will be governed in large part by long term relational contracts between agents and firms, such as between customers and a bank, both consumers and merchants, and between firms within the banking and payments network. Indeed, for the most part the only spot contracts in this system of economic coordination are the exchanges of goods and services for money between consumers and merchants.

Figure 3 re-imagines the credit card network from a contractual governance perspective. Our central argument in this report is that figures 1 and 2 (above) have dominated discussion and analysis without sufficient consideration of the implications of figure 3.

Figure 3: Interchange fee in a governance framework



Source: Davidson & Potts 2015

Theoretical foundations: efficiency, not monopoly

The efficient organization of economic activity entails matching governance structures with these transactional attributes [uncertainty, frequency of exchange, asset specificity] in a discriminating way.

Oliver Williamson (1979: 261)

Economics laureate Oliver Williamson won his prize in large part for his classic work *The Economic Institutions of Capitalism*. Building on the work of Ronald Coase, Williamson developed the transaction cost-based field of New Institutional Economics, at the heart of which was a clear distinction between the monopoly branch and the efficiency branch of microeconomic analysis. As Williamson (1985: 23) explains:

The monopoly approaches ascribes departures from the classical norm to monopoly purpose. The efficiency approaches hold that departures serve economizing purpose instead.

Williamson explained how economic agents will seek to 'organize transactions so as to economise on bounded rationality while simultaneously safeguarding them against the hazards if opportunism' (ibid: 32). Williamson's point is that sometimes forms of economic organization that may look like collusive or monopolistic behaviour when examined in terms of resource allocations are actually forms of economizing when analysed from the institutional perspective of transactions.

We argue that the dominant regulatory view of payments networks and interchange fees is through the lens of the monopoly view of economic organization (Carlton and Frankel 1995). This view focuses on resource flows and rents (as in figure 1), and within that seeks to identify the exercise of monopoly power. The monopoly view of bank interchange fees is based around an applied price theory approach in which barriers to entry give rise to leverage and price discrimination, resulting in rent capture. The implied correction to this outcome is to restrict the ability to exploit the rents through a legislative price ceiling – i.e. fixing a maximum interchange fee.

But this same situation looks rather different when the unit of analysis is the transaction (as in figure 3). The notion of a transaction includes both exchanges and contracts. Economic organization can occur in a spot-market (exchange) with neither future promises nor responsibility, or through long-term relational contracting, where parties make investments of which the profitability and utility depends on the other parties subsequent behaviour (Alchian and Woodward 1988: 66). Transaction cost economics predicts that where there are transaction specific assets, trading regularities will emerge that support and signal continuity intentions (Rochet and Tirole 2000), thus expanding trade from a unilateral spot-market relation to a bilateral ongoing relational contract.

From the transactions cost perspective, observed departures from the classical model may therefore reflect economizing behaviour in conducting ongoing transactions, and in the context of risk of opportunism and bilateral investment may already be ex post efficient forms of organization of economic activity. In consequence, if these adapted institutions and contracts are efficient forms of economic organization, then regulatory intervention will harm efficiency. Consider why this might be so.

Long term contracting and spot markets in credit and payments systems

Figure 3 indicates that of the four types of transactions relations between consumers (C), issuing banks (B1), acquiring banks (B2), and merchants (M), three of those relations (C-B1; B1-B2; B2-M) will usually be governed by long-term relational contracting, and with only C-M being a spot market transaction. Why is this?

First, why are they not all spot contracts? Specifically, why are C-B1 and B2-M typically long-run relational contracts rather than spot contracts?

One, they are engaged in multiple repeated transactions, and minimizing transactions costs associated with processing scale economies are achieved through bundling transactions through a single supplier. This incentivizes B1 to form a long-term contract with C.

Two, there is asymmetric information about creditworthiness of C that accumulates through repeated transactions, and which then enables a cumulatively better offer to be made to C as their true risk is cumulatively revealed, which then incentivizes C (if their 'true type' is low risk) to form a long-term contract with B1. This moral hazard problem of constraining C to good behaviour is enforced with threat of expulsion from the contract by B1, which would then take them back to a higher rate with a new issuing bank that had not accumulated information about the credit properties of C.

This in turn works as an effective screening mechanism by B1 on C, because only a high quality C will accept the conditions of a long-term contract, which will be valuable to C and profitable to B1, only if C can be effectively constrained from opportunistic behaviour.

Three, the same arguments apply between B2 and M, where B2 accumulates information about the transaction volume of M and their propensity to accept fraudulent sales (which require chargebacks). This information is a specialized asset that is profitable to B2 (and B1) if they can constrain opportunism by M (and C). The long-term relational contract, and the credible threat of expulsion from that contract, is an efficient governance mechanism to organize economic coordination in the context of the threat of opportunism and information asymmetry.

Four, incomplete relational contracts enable many specific contingencies to be dealt with by negotiation between the parties under the threat of exit, with the ensuing costs that imposes. These are a private ordering that may have final recourse to courts, but will often be most efficiently handled through direct bargaining under credible commitments and threats through the various hostages (threat of default versus threat to harm credit score) that each side has offered the other (Williamson 1983).

Five, long-term contracts may arise because of differential risk preferences between consumers, merchants and banks, which banks being systematically risk neutral and consumers and merchants being risk adverse.

Second, why is B1-B2 a relational contract, rather than either a spot exchange or horizontally integrated within a single firm (see Williamson 1985: ch6)?

A single bank – integrating B1 and B2 within a single firm – might be technologically efficient, but would be informationally inefficient, would be exposed to greater risk of shirking behaviour because of information impactedness and costly monitoring, and would be exposed to opportunism in internal pricing transfers. Because retail consumers and merchants are highly heterogeneous and

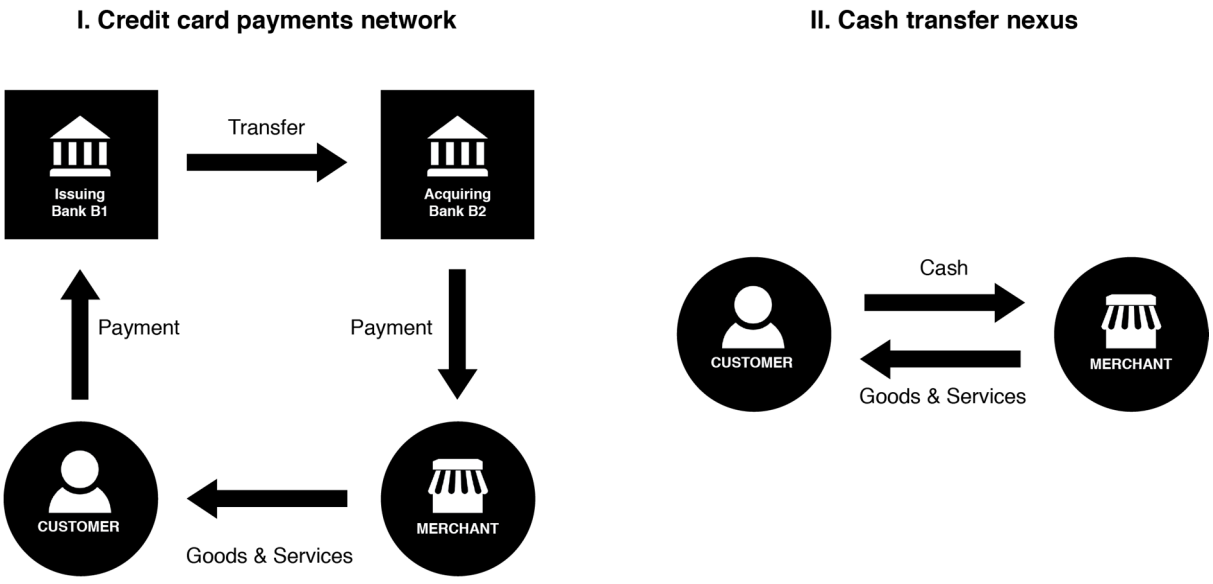
geographically distributed, specialized skills and investments are required in assessing quality (i.e. true type) and in delivering services. Banks will therefore tend to specialise under competition in order to economise on information. Long-run relational contracts then reconnect this into a payments network under high-powered incentives. In general this can be observed in the relative market success of open payments networks over closed payments networks.

Long-term relation contracting is efficient because banks take different sides of many transactions, giving rise to threat of exploitation through non-cooperative play. However, opportunism is disciplined only by threat of retaliation. Furthermore, repeated transactions enable learning and synchronising of processes and transaction routines in order to generate an efficient payments system, all without loss of high-powered incentives if the transactions were integrated into a single firm.

4.2. Argument 2: Equilibrium in choice of payments: cash versus credit cards

An important point follows from these considerations of the payments system in terms of transaction costs and the institutions that develop to efficiently govern these. In essence, these will be some margin of equivalence between alternative governance institutions, which we represent in figure 4 below with the credit card payments network on the left and the cash transfer nexus on the right.

Figure 4: Payment Networks and Cash Transfer



Source: Davidson & Potts 2015

First, the added complexity and physical and organizational resources involved in the card payments network, which are approximated by the flow of fees that consumers and merchants pay to the banks, will in equilibrium be competitively disciplined by the threat of exit to the cash transfer nexus (on the right in Figure 4 above). What is crucial to understand is that the cash transfer nexus is not the default setting of free, against which to compare the costs and fees of the card payments network. There are significant costs associated with the use of cash, for both customer and merchant, and both will be willing to pay some margin to avoid those costs. For the consumer, the costs are the carry costs and risks of using cash. For the merchant, these are the same costs in processing cash, but also in the

reducing in sales due to financing constraints by the consumer. The merchant will be willing to pay some margin to enable the consumer to access credit.

Second, the merchant is not indifferent between cash and credit because of asymmetric information and adverse selection. For the customer, in their relation with Bank1 and Merchant, the equivalence between cash and credit depends on the benefit of liquidity plus the carrying cost of liquidity (cost of carrying cash, cost of credit cards). In equilibrium, the cost of carrying and using cash will equal the maximum credit fee charge. However, this assumes that the customer is of a type: 'creditworthy and solvent', and that this is known to the merchant and the bank. Yet there is no reason to suppose the merchant knows this, or can acquire this information at low cost.

An equivalent argument occurs on the merchant side of the equation. In a long-term relationship between Customer and Merchant there would emerge an equivalence between cash and credit, plus the transaction cost that would be self-enforcing in long run equilibrium of a repeated game only if the exchange relation was at least a one-sided monopoly. But in a competitive spot market the logic is different because the consumer choosing credit over cash is not just facing a transaction cost decision but also signalling information about their 'true type' as a credit risk [i.e. good or bad]. A consumer choosing credit in the C-M transaction risks signalling that they expect not to pay (that they are a bad type), which drives an adverse selection/moral hazard spiral that will drive credit out of the spot market, leaving only cash. This will result in a lower equilibrium level of transactions because good credit use in the spot market (i.e. 'good' customers, for whom the cash carrying cost greater than the credit fee cost) suffers a 'lemons' problem (Akerlof 1971). B1, however, has a long term relation with C, and thus can effectively underwrite that use of credit in the spot market.

5. Summary and Analytic Conclusions

Interchange fees are not a problem of monopoly exploitation, but rather an efficient solution to an unavoidable bilateral monopoly that arises because banks need to form long term relations with customers and merchants – what are in effect irreversible investments that pay off only if the relationship continues – and which are therefore vulnerable to opportunism.

We make two specific theoretical claims that explain why regulatory intervention to cap the interchange fee will harm consumer welfare. Both claims hinge on recognizing that the governance structure of the card payments system is composed of long run relational contracts, the threat of exit from which disciplines short run opportunism in the system.

First, the interchange fee equilibrates the issuing (B1) and acquiring (B2) sides of payment cards systems. A fee setting association of banks is not evidence of collusive monopoly, but of minimizing transactions costs across the network in achieving economic coordination between all transacting parties. Constraints placed on internal bargaining and side-payments – i.e. an interchange fee ceiling – cause less efficient outcomes, resulting in higher fees to consumers and an unnecessary loss of social welfare.

A further implication is that interchange fees also enable an efficient network governance structure based around relational contracting that avoids horizontal integration between issuing and acquiring banks, maintaining incentive intensity and minimizing administrative monitoring burden arising from information impactedness.

Second, the relevant theoretical comparison between the four-party card payments system and the simple two-party cash nexus exchange must recognize that cash is also costly to consumer and merchant and that both parties will be willing to pay some margin to use a superior payments technology. This can be seen clearly when we consider why merchants do not usually offer credit payments to customers – or are risk averse in doing so – but banks can be risk neutral in this offering, namely because they are in a long term relational contract with the customer, and can effectively punish opportunism. Both consumers and merchants are willing to pay to avoid cash transactions by agreeing to enter long term contacting relations with banks.

Bibliography

Alchian A. and S. Woodward. 1988. The firm is dead; long live the firm: a review of Oliver Williamson's The Economic Institutions of Capitalism. Journal of Economic Literature, 26: 65-79.

Balto, D. 2000. The Problem of Interchange Fees: Costs Without Benefits? European Competition Law Review, 21, pp. 215-224.

Baxter, W. 1983. Bank Interchange of Transactional Paper: Legal and Economic Perspectives' Journal of Law and Economics 26(3): 541-588

Carlton, D. and A. Frankel. 1995. The Antitrust Economics of Credit Card Networks. Antitrust Law Journal, 63: 643-668.

Chang, H. and D. Evans. 2000. The Competitive Effects of the Collective Setting of Interchange Fees by Payment Card Systems. Antitrust Bulletin, 45: 641-77.

Coase, R. 1937. The Nature of the Firm. *Economica*, 4: 386-405.

Coase, R. 1960. The problem of social cost. *Journal of Law and Economics*, 3: 1 – 44.

Coase, R. 1972. *Industrial Organization: A proposal for research*. In R. Coase. 1988. The firm, the market and the law. University of Chicago Press.

Europe Economics 2014. The economic impact of interchange fee regulation. London, UK.

Gans, J. and S. King. 2001. The role of interchange fees in credit card associations: Competitive analysis and regulatory issues. *Australian Business Law Review*, 29 (2): 94 – 122.

Gans, J. and S. King. 2002. Regulating Credit Cards in Australia: A Submission to the Reserve Bank of Australia. Available from http://www.rba.gov.au/payments-system/reforms/cc-schemes/consult-doc-responses/core-120302.pdf

Gans, J. and S. King. 2003a. The neutrality of interchange fees in payment systems. *Topics in economic analysis & policy*. 3(1): online.

Gans, J. and S. King. 2003b. Approaches to regulating interchange fees in payment systems. *Review of Network Economics*, 2(2): 125 – 145.

Gans, J. and S. King. 2003c. A theoretical analysis of credit card reform in Australia. *The Economic Record*. 79(247): 462 – 472.

Hayashi, F. and S. Weiner. 2006. Interchange fees in Australia, the UK, and the United States: Matching theory and practice. *Economic Review*. 91(3): 75 – 112.

Katz, M. 2001. Reform of credit cards schemes in Australia. RBA Commissioned Paper. http://www.rba.gov.au/payments-system/reforms/cc-schemes/ii-commissioned-report/

Klein, B., R. Crawford, and A. Alchian. 1978. Vertical Integration, Appropriable Rents and the Competitive Contracting Process. Journal of Law and Economics XXI (1978), 297-326.

Reserve Bank of Australia and Australian Competition and Consumer Commission. 2000. Debit and Credit card schemes In Australia: A study on interchange fees and access. October 2000.

Reserve Bank of Australia. 2015. Review of card payments regulation: Issues paper. March 2015.

Rochet, J. 2006. The Consequences of Reducing Interchange Fees. Presentation to 2006 Payments System Conference, March 2006, Available at http://web.archive.org/web/20060613224511/http://www.mbs.edu/payments_system/

Rochet, J. and J. Tirole. 2000. Cooperation Among Competitors: The Economics of Payment Card Associations, mimeo, University of Toulouse.

Rochet, J. and J. Tirole. 2003. An economic analysis of the determination of interchange fees in payment card systems. Review of Network Economics. 2(2): 69 – 79.

Schmalensee, R. 2002. Payments systems and interchange fees. Journal of Industrial Economics

Stigler, G. 1964. Public regulation of the securities market. Journal of Business, 37(2): 117-142.

Williamson, O. 1973. Markets and hierarchies: some elementary considerations. American Economic Review, 63(2): 316-25.

Williamson, O. 1975. Markets and Hierarchies: Analysis and Antitrust Implications, New York: Free Press.

Williamson, O. 1979 Transaction cost economics: the governance of contractual relations. Journal of Law and Economics, 22(2): 233-61.

Williamson. O. 1983. Credible commitments: Using hostages to support exchange. American Economic Review, 73(4): 519-40.

Williamson, O. 1985. The Economic Institutions of Capitalism, New York: Free Press.

Williamson, O. 1996. The Mechanisms of Governance, New York, Oxford University Press.

Williamson, O. 2002. The theory of the firm as governance structure: from choice to contract. Journal of Economic Perspectives, 16(3): 171-95.