Literature Review on Competition, Efficiency and Surcharging in the Retail Payment System

Report prepared for the NZ Commerce Commission

Amelia Fletcher, University of East Anglia, UK

Introduction and Executive Summary

This literature review primarily discusses the economics of "four-party" payment card schemes, as these have been the predominant focus of academic work in relation to retail payment systems.

There has been substantial concern in many jurisdictions that charges to merchants for card acceptance ("multilateral interchange fees") were too high and, in response, regulation has been introduced across many jurisdictions, including New Zealand (NZ).

The academic literature examines the key economic characteristics of these card markets, why these may lead to inefficient market outcomes, and implications for regulation. In this review, we start by describing the nature of four-party card schemes, as well as some unique aspects of the NZ payments space.

Introduction to four-party card schemes

Usage of credit and debit cards has become increasingly prevalent, around the world, as the use of cash and cheques has declined. At the same time, two major providers – Visa and Mastercard – have a very large share of card payments globally (outside China).

This strong position of the two major providers derives partly from high economies of scale, but also from the existence of powerful so-called 'across-group network effects'. These arise because merchants value a scheme more highly if it has high cardholder take-up, while cardholders value a scheme more highly if it can be used across a wide range of merchants. It is well understood that such network effects can lead to concentrated market structures, with high barriers to entry and expansion, and card schemes are no exception.

Both the Visa and Mastercard networks are known as 'four-party' schemes because they involve two types of bank and two types of customer:

- **Issuing banks** have a direct relationship with **cardholders** (consumers).

- **Acquiring banks** have a direct relationship with **merchants** (retailers) that accept the card for payments.

1

For some transactions, the same bank could in practice both be both issuer and acquirer, but in a four-party scheme this need not be the case..

These four-party schemes involve a number of rules which members must adhere to and also a number of fees that are paid between the parties. The fee that has been the key focus of regulatory attention has been the Multilateral Interchange Fee (or MIF), which is charged by the issuing bank to the acquiring bank.

To understand the role of the MIF, a simple example may be helpful. Suppose a cardholder makes a payment of $100 to a merchant. In practice this payment involves several distinct steps.

- The **issuing bank** debits $100 from the cardholder's account (debit cards) or adds it their credit balance (credit cards).

- The **issuing bank** then transfers the money to (or credits) the **acquiring bank** with $100 minus a Multilateral Interchange Fee (MIF) – say 1% (or here, $1)

- The **acquiring bank** then credits the **merchant**'s account with $100 minus a Multilateral Service Charge (MSC) – say 3% (or here, $3)

Thus, in this simple example, the issuing bank retains $1, which is the MIF. The acquiring bank retains $2, which is the difference between the MSC and the MIF. The merchant receives $97 of the original $100 paid. While the MSC is set by each acquiring bank individually, the MIF is agreed between scheme members. The MIF is clearly a key component of all MSCs.

Finally, MasterCard and Visa also charge switch fees (also called network fees) to issuers and acquirers. Interestingly, despite the intense policy debate around the MIF, these switch fees have received little to no attention from academics, perhaps because they are typically low.[1] However, there is clear potential for them to be utilized to replicate the effect of a MIF, and regulators are increasingly focusing on their use.[2]

One key reason why they are important is that, in the above example, absent a MIF, an identical allocation of rents could be achieved by the scheme charging the acquiring bank 1% of the gross consumer payment and then passing this back to card issuers as a negative fee. Any such rent transfer would also have the same economic effects as a formal MIF (as are discussed in this literature review). It is clearly important for regulators to consider the full quantum of cross-payments across the system, and not just the formal MIF.

Note that such four-party systems contrast with 'three-party' schemes such as American Express, which have traditionally dealt directly with merchants and cardholders, and thus only one charge is levied.[3]

---

[1] For example, in Australia In 2020/21, the net scheme fees for domestic card use lay between 0.02% (to issuers for debit cards) and 0.11% (to acquirors for credit cards). See Reserve Bank of Australia (2023 The Shift to Electronic Payments – Some Policy Issues. Speech by Ellis Connolly

[2] For example, UK Payment Systems Regulator (2023) Market Review Into Card Scheme and Processing Fees.

[3] Note that American Express sometimes uses independent banks to issue its cards, and where it does so it is also a 4 party scheme for some cardholders. However, American Express has a far smaller footprint than the two major card schemes. It also exclusively provides credit

The New Zealand payments space

While Visa and Mastercard have a strong position across the globe, the NZ payments space has some unique aspects, and in particular the established position of the EFTPOS payment system. This was originally owned jointly by the four large NZ retail banks but was sold in 2019 and is now owned and operated by Worldline.

The EFTPOS network can be accessed through either an EFTPOS card or a Visa/Mastercard debit card. This is a very low-cost system. Issuers pay a small flat fee per transaction, and there is no merchant service charge. The merchant simply pays a monthly flat fee to connect their card terminal to the network.

The established position and low cost of the EFTPOS network means that it has traditionally accounted for a relatively high share of card transactions in NZ, while cheques have been entirely phased out in NZ and cash use has been low for some time.

However, the EFTPOS system can only be accessed if cards are inserted into or swiped through a terminal. EFTPOS has no contactless or online payment functionality, both of which have become increasingly important in recent years, especially since Covid and the development of mobile phone payments such as Apple/Android Pay.[4]

Moreover, issuer banks' incentives to issue EFTPOS cards, or encourage their use, is low because they can make more money through issuing Visa or Mastercard cards. Account-to-account payments are also slow, limiting the potential of this payment route to provide a competitive alternative to card payments. At the same time, most cardholders 'single home' with one of the two major card schemes, since they tend to receive both debit and credit cards from their main retail bank, each of which in turn has exclusive arrangements with one of the schemes.

As a result, EFTPOS cards have seen declining usage (from around 55% to around 35% of in-person payments[5]), and there has been a growth in the usage of the two major card schemes. It is in this context that the new NZ payment systems regulation has been introduced.

Key concerns relating to the pricing of payment card schemes

The principal policy concern, across multiple jurisdictions, in relation to four-party credit and debit card schemes has been the level of the Multilateral Interchange Fee (or MIF) paid by merchant acquiring banks to issuing banks. Since the MIF is a form of horizontal agreement between rivals, it clearly risks breaching competition law.

---

cards, rather than debit cards, and it has a business model which focuses on high merchant fees and high cardholder rewards. As such, it places limited competitive constraint on the MIF of the two major card schemes.

[4] Terminals are designed such that Visa/Mastercard scheme cards utilise the EFTPOS system if the card is inserted or swiped but utilise the scheme card network if they are used contactlessly. Visa/Mastercards can of course also be used online.

[5] See Commerce Commission New Zealand (2023). *Retail Payment System: Payments Between Bank Accounts*.

The early sanguine view

An early US investigation of the MIF under competition law in 1986 gave it a clean bill of health, viewing it as an efficient mechanism to ensure universal card acceptance.[6] The argument made was that the MIF solves a coordination problem related to an important feature of card schemes; the 'honour all cards' rule. Under this rule, merchants that accept a particular scheme's cards must accept all cards carrying that scheme's logo.

This rule was considered valuable for cardholders, as it provides clarity as to whether their particular card will be accepted. However, it also has implications for pricing. Suppose that issuing banks chose their interchange fee unilaterally. Under the 'honour all cards' rule, merchants cannot pick and choose between card issuers; they have to accept them all, or leave the scheme altogether. This latter choice will depend on the MSC they face under the scheme across all cards in the scheme, which in turn depends on the *average* interchange across issuing banks. Thus, from the point of view of any issuing bank choosing unilaterally, its choice of interchange fee would only have a limited impact on the overall average interchange and thus on card acceptance. This in turn creates a 'hold-up' problem. Each individual issuing bank would unilaterally wish to set its interchange fee far higher than would all of issuing banks acting jointly.

On this basis, it was argued that a jointly set MIF in fact solves this 'hold-up' problem and thereby keeps interchange fees *lower* than would be the case if issuing banks set their own interchange fees, benefiting both scheme members and merchants. For this reason, it was initially not considered to breach competition law.

Later, more critical, views

However, the MIF became viewed more critically from the early 2000s. Rather than unilaterally set interchange fees, it was argued that a zero interchange fee was an alternative counterfactual to a collectively set MIF. In this case the MIF would clearly be raised by the collective price setting.

The real policy concern, though, was not the collective price-setting *per se* but the fact that the jointly set fee was too high. Jurisdictions sought to address this concern through a mix of direct regulation and antitrust.

In 2003, the Reserve Bank of Australia was the earliest to introduce regulation, reducing the MIF from 0.95% to 0.5%.[7] In Europe, the European Commission initially employed competition law to address MIFs, agreeing a settlement with Visa in 2002[8] and sanctioning Mastercard (in relation to cross-border MIFs) in 2007[9]. However, it then changed tack, concluding that competition law was not the most suitable tool for intervention in this context. It instead adopted the EU *Interchange Fee Regulation* in 2015. This regulation

---

[6] *National Bancard Corp v VISA USA Inc* 779 F 2d 592 (11th Cir. 1986).

[7] Reserve Bank of Australia (2001), 'Reform of credit card schemes in Australia'.

[8] *Visa International - Multilateral Interchange Fee* (Case COMP/29.373), Decision of 24 July 2002.

[9] *MasterCard* (Case COMP/34.579), Decision of 19 December 2007; and Case C-382/12 P *MasterCard and Others v Commission*, Judgment of 11 September 2014.

reduced the MIF to 0.2% for debit cards and 0.3% for credit cards.[10] It also limited the allowable scope of any 'honour all cards' rule

Why might the MIF be set too high?

This raises an obvious question. Why might the MIF be set too high?

Before answering this question, it is important to note two important characteristics of payment cards markets. The first is that it can be viewed as a 'transactions' market in that charges are levied in the context of one side of the market (merchants) selling to the other side (consumers). This is important because, in setting prices for consumers, merchants will take account of the costs it faces for accepting cards. Thus, higher fees charged to merchants will tend to feed through to consumers as higher prices.

Neutrality result

This gives rise to an important 'neutrality result' (Gans and King, 2003). If there were perfect card-specific surcharging by merchants, such that high fees to merchants were passed back efficiently to that scheme's cardholders as lower fees (or rewards), then the level of these fees has the potential to be neutral. Any increase in a scheme's fees to merchants would feed into higher prices for its cardholders which would then be discounted to the same cardholders via the rewards. If these perfectly offset each other, as can occur under certain conditions, then the net price paid by cardholders would be unchanged. All parties would then be indifferent to the level of the MIF.

However, this neutrality result depends on several assumptions. Most critically, it relies on perfect surcharging by merchants, which in turn depends on perfect visibility of, and rational reactions to, surcharges on the part of consumers.[11]

In practice, however, this is unrealistic, as will be discussed further in **Section 4**. Although surcharging is no longer banned by the schemes (as it used to be), many retailers are either reluctant to surcharge or, if they do, do not make such surcharges sufficiently clear and prominent to really change consumer behaviour. Key barriers to efficient surcharging by merchants include their concerns about consumer reactions to surcharging, a lack of knowledge amongst merchants about the true MSCs they face, and the ability and incentive of merchants to shroud surcharges so that they have limited effect on consumer choice of payment type.

This is important because, absent perfect card-specific surcharging, the neutrality result breaks down. Instead, any increase in card fees for merchants will tend to feed into to general retail price rises, which are borne not only by that's scheme's cardholders but also by other consumers, who are not scheme members.

---

[10] European Commission (2015), 'Regulation (EU) 2015/751 of the European Parliament and of the Council of 29 April 2015 on interchange fees for card-based payment transactions', 19 May. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015R0751

[11] It also assumes perfect competition amongst merchants, amongst merchant acquirers, and amongst issuers, and the employment of usage fees (rather than fixed fees).

Much of the economic literature on payment cards thus assumes that merchants do not surcharge and considers the implications of this for the setting of the MIF.

An efficient MIF? The 'merchant indifference test'

Given that merchants are assumed to be unable to influence consumers' choice of payment method through surcharging, this will clearly lead to inefficient market outcomes if merchants themselves face very different costs of accepting different payment methods. This can be avoided through a MIF which imposes a cost on merchants for accepting cards which exactly equates to their cost of utilising the alternative payment method (typically assumed to be cash). This is the '*merchant indifference test*' approach to setting an efficient MIF.

Under the very specific assumption that card issuers have zero margins, this test in fact maximises both total user (consumer plus merchant) surplus and total welfare. More generally, though, Rochet and Tirole (2011) show firstly that the MIF which maximises total user surplus typically lies below the MIF which maximises total welfare (since the latter includes scheme participant profits), and secondly that introducing a series of more realistic assumptions (positive/variable issuer margins, network effects, potential or entry) will tend to increase the level of the MIF that maximises total user surplus to above the MIF that meets the merchant indifference test.

On this basis, they argue that the merchant indifference test provides a conservative minimum level of MIF which will (almost) certainly be no higher than the level which would maximise either total user surplus or total welfare.

Reasons why the MIF might lie above the efficient level

There are a number of reasons why the MIF that is observed in an unregulated setting may lie above the efficient level.

First, although there are two card competing card schemes globally (and in NZ), these card schemes typically have **'bottleneck' market power** over merchants. This arises because the vast majority of cardholders are 'single-homing', in that they only use one major card scheme, and thus merchants have little choice but to use both schemes, in order to access the market.

In platform markets with single-homing users on side and multi-homing users on the other, the effect of 'bottleneck' market power on the multi-homing side will tend to drive up fees on that side of the market. Rochet and Tirole (2011) find exactly this result in the context of payment schemes. Indeed, they find that bottleneck market power results in the MIF being set at the same level as would be set by a monopoly card scheme, which in turn may well be substantially above the efficient level. Guthrie and Wright (2008) find that the MIF set under duopoly can even exceed the monopoly level.

Second, given a lack of surcharging, any rise in the MIF of one card scheme will be incorporated into the price across all consumers, and not just cardholders of that particular scheme. Effectively, this means that a scheme's MIF rise confers a **negative externality** on

non-scheme users. Edelman and Wright (2015) how that this limits the competitive constraint on each scheme's choice of MIF, leading to MIFs being set excessively high.

Third, because card acceptance is valued by consumers, merchants that accept cards can set higher prices. Thus merchants '**internalise**' the benefits received by consumers. Wright (2012) shows that this leads to a 'double counting' of cardholder benefits which in turn has the systematic effect of inflating the MIF.

Fourth, Bedre-Defolie and Calvano (2013) highlight that consumers make two decisions (whether to hold cards and then whether to use them) whereas merchants typically only make one (whether to accept them). They note that the **dual decisions of consumers** mean that issuers can extract a greater share of cardholder surplus (their benefit from using cards) than it can of merchant surplus. The card scheme thus has an incentive to maximise cardholder surplus (so it can then extract this back), which in turn involves setting a high MIF that is biased towards cardholders.

Finally, Bourguignon et al (2019) highlight that consumers may have a strong preference for using a card but may have **imperfect information** about each merchant's cash/card policy. This imperfect information can generate 'missed sales', whereby customers are in the shop and eager to buy, but wish to use a card and are discouraged by either a high card surcharge or an outright rejection of the card. The risk of such missed sales induce merchants to feel they 'must take' the card which makes merchant demand for the card scheme more inelastic, further increasing the MIF that is charged.

Implications for regulation

The above literature underpins the case for capping the MIF. However, this in turn raises the question of how this price cap should be set. A complete discussion of this topic is beyond the scope of this review, but the key additional economic literature relevant to this question is discussed in **Section 5**.

The 'tourist test'

The leading proposed test is the '*tourist test*', which is effectively the 'merchant indifferent test' applied to the situation where all consumers are assumed to be tourists (to abstract from any benefits of attracting repeat sales). Under this approach, the MIF would be set equal to the *merchants' average convenience benefit from using a given payment method minus the merchant acquirers' average costs*.

As discussed above, the 'tourist test' is targeted at maximising total user surplus, not total welfare (which also takes account of the profits of scheme participants). Rochet and Tirole (2011) recognise that a social planner – seeking to maximise total welfare – could in principle go further in seeking to address the various cross-market externalities and imbalances of market power. However, they advise against such targeted price regulation, not least because of the large informational requirements it presupposes. Vickers (2005) also cautions against seeking to maximise total welfare, on the basis that it would be equivalent to subsidising a

monopolist in order to reduce its price to cost – an unusual form of government intervention. He thus supports the basic 'tourist test'.

Applying the test

Applying the tourist test is far from straight forward, however. It was developed in the context of a model in which the realistic alternative to cards is assumed to be cash. But while this was always somewhat over-simplistic, it is becoming more so, as cash use is phased out across many countries. This raises a critical question in terms of what counterfactual to employ in calculating merchant benefits from cards.

In many economies, of course, debit cards are now the key alternative to credit cards, but clearly these are also priced by the card schemes and involve MIFs which may themselves be excessive and require regulation. This makes them an unsuitable comparator.

In New Zealand, the EFTPOS card scheme arguably provides a good counterfactual, and the cost of this to merchants per transaction is zero. This does not mean, of course, that 'tourist fee' should be set at zero, as the major card schemes offer additional merchant benefits that should be taken into account. This may be a reasonable approach in the context of contactless payments, the benefit of which should prove possible to estimate. It may be less realistic in the context of online sales, where neither EFTPOS nor cash are reasonable alternatives.

Finally, three alternative approaches to the 'tourist test' are discussed briefly in **Section 5**.

- The first is simply to set the MIF at *zero*, based on the potential for this to be a reasonable counterfactual to a collectively set MIF. This approach risks being too low to incentivise investment and innovation.

- The second is to base the MIF on some measure of costs, such as card issuers variable costs. This approach, which is used in Australia and the US may be pragmatic but not supported by any economic theory.

- The third is to adopt an econometric approach, proposed by Bedre-Defolie et al (2018) and Huynh et al (2022), to quantify the network externalities and identify the main determinants of consumer and merchant decisions. An important distinction between this approach and the 'tourist test' concept is that it explicitly allows not only for 'first time customers' (tourists) but also repeat customers.

Empirical evidence on the impact of MIF regulation

A number of empirical studies have examined the impact of introducing MIF regulation across different jurisdictions. These are discussed in **Section 3**.

Overall, interventions that lowered MIFs to below their privately set level have been found to have positive effects on merchant acceptance, on consumer credit card adoption, the number of payment transactions, and even an increase in overall bank revenues from credit and debit cards.

On the other hand, there are clearly risks associated with MIF regulation. For example, Ardizzi et al (2021) finds that further reductions in the MIF to below that required by the EU Interchange Fee Regulation would likely harm card usage and transactions.

A reduction in the MIF below the privately set level can also have distributional effects. Schuh et al (2010) show that the MIF effectively transfers income from lower income households (who lose out through general increases in retail prices) to higher income households (who gain through reduced card fees and rewards). Albeit Felt at al (2020) find that the relationship may be non-linear, since the very richest households tend to use debit cards, rather than credit cards, and thus gain less from the MIF.

Lower income households could potentially lose out if a lower MIF led to less credit being available, but in fact Agarwal eet al 2023) finds evidence of excessive lending to those 'near prime' consumers who are most likely to be affected by this.

There is less evidence on the impact of MIF regulation on investment and innovation, with theoretical papers offering divergent views on the likely effects. However, Yemail (2022) examines innovation in Latin America, after several interventions across different jurisdictions, including capping MIFs, an describes how this set of interventions helped create an environment which facilitated entry from a range of innovative payment solutions, such as direct bank transfers and digital wallets which allow for payments via links and QR codes.

Report structure

The remainder of this report provides further detail on these various issues, as well as flagging some other related economic literature. **Section 1** provides a basic introduction to the economics of networks. **Section 2** focuses in on card payment schemes and discusses the literature relating to implications for competition. **Section 3** examines the available empirical evidence on implications for efficiency and innovation. **Section 4** examines the economic evidence on merchant surcharging. **Section 5** concludes with a brief discussion of implications for regulation.

1. Competition on and between networks

In this first section, we start by considering the nature of multi-sided markets, such as card schemes, which exhibit strong network effects.[12] The economic literature in this area typically refers to 'platforms', reflecting the growing importance of online platforms as critical multi-sided markets. In economic terms, payment systems such as card schemes can also be seen as 'platforms' and we will utilise this terminology for them too.[13]

The discussion here, while not specific to payment schemes, helps to explain why we observe only two major payment card schemes globally (outside of China), Mastercard and Visa, and why these have proven so hard to challenge. It starts to explain why they may be expected to set their MIFs higher than is optimal, absent regulation. The section concludes with a short discussion about market definition in platform markets.

1.1 Multi-sided markets and network effects

Multi-sided platforms link users on one side of a market with users on another side. They frequently levy fees on one or other side (or both) for this service. They may also impose rules on the conduct of platform users on one (or more) of the sides. For example, the card schemes typically require merchants to 'honour all cards' of a particular type.

Importantly, such platforms typically exhibit 'network effects' These occur where users value the size of the network, in addition to the core service provided.

- **'Within-group' (or 'direct') network effects** involve users on one side of a market valuing a greater number of other users on the *same* side. Think of a social network like Facebook: users value such a network more, the more of their friends that are on it.

- **'Across-group' (or Indirect) network effects** involve users on one side of a market valuing a greater number of users on the *other* side. Think marketplace platforms like Amazon: consumers value such marketplaces more, the more traders they can choose from on them; while traders value them more, the more consumers there are to sell to.

We see both forms of network effect in payment systems. Where payments are between the same type of users, such as account-to-account payment systems, within-group network effects are more relevant. For card systems, by contrast, there are two discrete groups on each side of the platform – cardholders and merchants – and 'across-group' network effects are more relevant. Consumers care about how many merchants take a card; merchants care about how many consumers use the card.

---

[12] The economic literature on platforms is rich and as such only a number of key results are presented here. For more extensive reviews, see Belleflamme and Peitz (2016) and Jullien and Sand-Zantman (2021).

[13] We note that payment platforms differ from many online platforms in that they typically have no 'matching role'. While Amazon 'matches' consumers with suitable products and Facebook 'matches' users with suitable content, payment systems typically execute transactions, but are not involved in the matching process. This means that payment systems are less likely than online platforms to be dependent on collecting data across markets to train their algorithm (albeit such data may be important for fraud prevention and assessing credit-worthiness). We do not therefore focus heavily in this section on the economic literature relating to data-sharing across markets.

Even for cards, though, there may be some direct network effects. For example, if payment systems utilise data to identify fraud, cardholders will get a better service from a scheme that has lots of other cardholders to train its fraud prevention algorithms. Likewise, if credit cards utilise personal transaction and credit usage data to estimate likely credit-worthiness and thus card access, this might be more effectively done, and thus card access better targeted, the greater the overall data across users available to platforms.

An important element in assessing the impact of network effects is whether users on a given side of the platform are 'single-homing' or 'multi-homing':

- **'Single-homing' users** only (or primarily) utilise a single platform.

- **'Multi-homing' users** utilise multiple competing platforms.

In general, if there are single-homing users on one side of a platform, it will be more likely that users on the other side will be multi-homing, as this becomes the only way for them to gain access to the single-homing users.

In practice, there may be a mix of user types of each side of a platforms, but it can nonetheless be useful to consider what most users on each side do. For example, it is likely that some end users of Google Search actually employ multiple search engines, while some advertisers only market their services via Google search. Nonetheless, it is more usual to think of Google Search as having single-homing end users and multi-homing advertisers, as this is the case for the majority of each type of user.

In the case of the two major card schemes, the situation is typically clearer still:

- **Most cardholders are single-homing**, in that they only use one of the two schemes. This partly reflects the fact that many consumers gets their debit and credit cards from their usual bank, and that each bank typically only issues cards from one scheme.

- **Most merchants are multi-homing,** in that they accept both major card schemes. This is the only way they can ensure they can be sure of being able to serve the many single-homing cardholders.

It should be noted that the position of American Express is rather different. Few (if any) cardholders have American Express as their only card. This in turn means that merchants are more readily able to refuse to accept this card scheme, if they consider the fees too high.

In New Zealand, the EFTPOS scheme has traditionally made the situation more complex, as most cardholders have had an EFTPOS card, perhaps alongside a card from Visa or Mastercard, and thus merchants may have been in a better position to accept EFTPOS only, for example by refusing to accept contactless payment.[14]

---

[14] Terminals are designed to utilise the EFTPOS network when cards are inserted or swiped, even where they are Visa or Mastercard cards.

However, the EFTPOS scheme lacks online and contactless functionality, both of which have become ever more important. For such payments, cardholders are again typically single-homing (that is, they have cards from either Visa or Mastercard) and merchants wishing to accept such payments are thus forced to be multi-homing (that is, accept both schemes).

1.2 Implications for market power

Market concentration

It is well understood that network effects can lead to markets 'tipping' towards being highly concentrated. The rationale is straightforward. If users value a platform more, the more users it has, then platforms can gain – and retain – a competitive advantage simply by getting bigger.

At an extreme, in a simple situation where both suppliers and consumers are 'single-homing' (use just one platform), whichever platform can gain the most users may well monopolise the whole market (Caillaud and Jullien, 2003). More generally, such network effects will tend to lead to highly concentrated markets. This helps to explain why we see just two major card schemes globally.

'Bottleneck' market power

In addition, even where there are multiple platforms in a two-sided market, the existence of 'single-homing' users on one side of a platform can confer *'bottleneck'* or *'gateway'* or' market power on that platform in relation to users on the other side (Armstrong, 2006). In the context of card schemes, the fact that cardholders are typically single-homing with one of the major card schemes means that most merchants cannot realistically afford to refuse to accept cards from either scheme, however high their fees become.

That is, irrespective of any competition that exists between the card schemes to win cardholders, they each have strong bottleneck market power vis-à-vis merchants.

The potential for competition for the market

In theory, high levels of concentration need not actually confer significant market power, so long as there is effective *'competition for the market'*. Indeed, such competition can potentially drive profits to zero, even in the context of a monopoly (Caillaud and Jullien, 2003).

However, in practice potential rivals are likely to face significant barriers to entry and expansion, which can limit such competition for the market. The key issue is that it can be hard to enter at small scale, as a small scale entrant may not be attractive to users. But this in turn means that successful entry requires a platform to win users quickly and on a large scale, which can in turn be difficult.

First, there can be important *user coordination* issues. Any given user wishes to adopt a new platform only if the vast majority other users are also going to adopt it. The success or failure of a platform thus depends on the beliefs consumers hold about its own future success or

failure. This in turn creates the risk of 'self-fulfilling' prophecies and 'focal equilibria'. If users all believe a particular platform is the best place to be, then it will be the best place to be.

There may be some potential to shift expectations to allow entry. For example, there may be market circumstances under which a higher quality entrant can signal its higher quality by undercutting a lower quality incumbent and thereby shift user expectations (Halaburda et al, 2020). However, the same paper shows that, where those market conditions do not hold, the same lower quality incumbent will be impossible to shift.

Thus, if expectations cannot be shifted, even efficient entry can be deterred. Indeed, if end users all wait to see what other end users are doing, this can prevent migration to a new platform (Biglaiser et al, 2022). The same paper also finds that multiple migration opportunities increase the perceived benefit of waiting and thus can increase the incumbency advantage still further. 'Status quo' bias or switching costs – even amongst only a subset of users – can also inhibit such competition for the market (Jullien and Sand-Zantman, 2021).

Finally, if one side of the market pays zero price, and if it is hard to price lower than this, then this can also limit the ability of new entrants to compete for the, albeit it may be possible to overcome this to some extent by tying in additional free services (Amelio and Jullien, 2012). Note that this additional barrier is unlikely to be relevant for credit cards, since they can pay 'rewards' to cardholders (a negative price), but it may be relevant for other payment systems.

Such issues do not mean that competition for the market is impossible. For example, Cabral (2011) models a situation where an incumbent monopolist, in seeking to exploit its market position, allows for the development of a niche competitor, and where this latter firm may occasionally displace the dominant firm, the process then starting again.

More generally, though, the higher the barriers to entry and expansion, the less easy will be such competition *for* the market.

The potential for competition <u>in</u> the market

There may be factors that mitigate this tendency towards tipping in multi-sided markets and enable '*competition in the market'* between platforms.

First, *multi-homing* by users can potentially reduce market power. The user coordination issue described above will be reduced if users can try out a second platform, without switching away from the first. Meanwhile, bottleneck market power will be reduced if all users have a choice of platforms through which to access users on the other side of the market.

In the context of the major payment card schemes, there is already multi-homing on the merchant side of the market. This raises the question of whether the bottleneck market power held by these schemes could be reduced by multi-homing on the consumer side, for example by banks issuing cards from multiple schemes, or even by cards being co-badged with multiple schemes.

In this context, it is noteworthy that the EU *Interchange Fee Regulation* (2015) prohibits card schemes from preventing such co-badging.[15] The current EFTPOS arrangements in NZ are arguably a form of co-badging, in that Visa/Mastercard fees also work on the EFTPOS network (and are required to do so, when the card is inserted or swiped). This co-badging has arguably helped to preserve EFTPOS's relatively strong position in the NZ payments market.

However, it is not clear how widely co-badging has been adopted by card issuers more broadly. This may be because there is limited demand for such co-badging from cardholders, given the prevalence of the existing major schemes across merchants. This is again a form of coordination or 'chicken and egg' type issue. Until there is sufficient multi-homing across card schemes by cardholders, merchants will want to accept both major schemes, but cardholders in turn have little incentive to multi-home so long as merchants continue to accept both schemes.

Second, *interoperability* can allow network effects to be shared across firms and thereby facilitate competition within the market. A simple example may be helpful. Suppose it was impossible to make calls across mobile phone networks. Consumers would value the phone network with the most users and the market would tend towards a monopoly. However, with the ability to call across networks, consumers no longer need to be on the same network as anyone they might wish to call, and competition in the market can be sustained.

In the context of cards, such interoperability could potentially be achieved by merchants signing up to a single third party network (such as Paypal or Apple/Android Pay) which can in turn accept any payment method. But note that a dominant platform's incentives to accept such interoperability will tend to be low, precisely because it facilitates competition (Crémer et al, 2000). Indeed, merchants can only currently accept Apple/Android Pay in relation to particular card scheme if they also accept that card scheme directly.

Third, *product differentiation* can allow multiple platforms to co-exist and compete.[16] This is essentially because users then care about more than simply the size of a platform's user base. Users who particularly value one platform will stick with it even if its user base is smaller than that of the rival platform. There is also a link between product differentiation and multi-homing (Armstrong and Wright, 2007). With strong product differentiation on both sides of the market, users may be more inclined to single home with their preferred platform. On the other hand, if there is product differentiation only one side of the market (say, consumers) but not on the other side (say, merchants), then the former will tend to single-home but the latter will tend to multi-home in order to be able to access all consumers.

Rather than differentiating their products, platforms can potentially *differentiate their business models*. For example, if there are some users on both sides who value large networks highly and others who have a low valuation for large networks, equilibrium may involve two platforms. One platform will set prices high on side and low on the other, while the other will do the reverse (Ambrus and Argenziano, 2009). This sort of thinking could potentially explain

---

[15] Article 8(1). See fn. 10.

[16] Indeed, much of the key economic literature relating to competition between platforms (such as Rochet and Tirole (2003) and Armstrong (2006) assumes horizontal product differentiation.

why we observe different pricing strategies by Mastercard and Visa on the one hand and American Express on the other.

Another type of product differentiation can arise where there is a risk of platform *'congestion',* whereby the platform starts to become less attractive – to at least some users – once it has too many members. If a large platform becomes too congested, then those users most affected by congestion may prefer to join a smaller, less congested platform. (Karle et al., 2020). This seems unlikely to be relevant for payment systems (at least unless they get close to a technical capacity constraint and this impacts quality).

One final way in which platforms can differentiate themselves is through *exclusive contracts* with certain 'superstar' users or content. While exclusive contracts can be anti-competitive when imposed by large incumbent platforms, they can also provide a useful route for entry by small entrants. For example, in the gaming industry, empirical evidence shows that exclusive deals between platforms and producers helps small platforms challenge the incumbents (Lee 2013).

In the context of payment card schemes, it could be argued that the use of exclusive contracts between the major schemes and the issuing banks may play a similarly positive role in preserving the current duopoly and preventing the market from tipping still further to effective monopoly. Against this, and as discussed above, such contracts may also help to limit multi-homing by cardholders, which in turn helps to limit competition in the market.

1.3 Implications for pricing

In any two-sided market, a platform's pricing to the two types of users will be affected by the cross-group network effects between them. Positive cross-group effects mean that raising the price on one side of the market doesn't just reduce take-up on that side. It reduces the value of the platform's services and thus shifts demand downwards on the other side of the market too.

In this subsection, we consider implications for platform pricing generally, under both monopoly and oligopolistic competition, without a specific focus on payment schemes (albeit we seek to draw out come key implications for payment schemes as an example). In Section 2, we consider the special case of payment card schemes.

Pricing under monopoly

For a monopoly, if the cross-group effects between the two sides are identical, they may not in fact affect overall pricing, at least in a simple linear model (Belleflamme and Peitz, 2016).

However, suppose that side 1 users (say, merchants) have a *higher* valuation of the interaction with side 2 users (say, consumers) than side 2 users value the interaction with side 1 users. In this case, the platform particularly values participation from side 2 users since this attracts so much extra participation on side 1. The platform will therefore set a lower margin on side 2 (to attract these users) and a higher margin on side 1, relative to what would prevail in the

absence of these cross-group effects. Indeed, if the divergence in cross-group effects is sufficiently high, this can lead to negative margins or even negative prices.

As such, even a monopoly payment card scheme might choose to price high to merchants (who greatly value their interactions with cardholders) and offer rewards (negative prices) to cardholders to increase their participation.

Pricing with oligopolistic competition between platforms

As discussed above, there are situations where we may see multiple platforms co-existing and competing, as we observe for payment card schemes.

With oligopolistic competition between platforms, a key insight is that network effects tend to increase price competition relative to the situation without network effects. Let us return to the situation described above where side 2 users (say, cardholders) exert a larger positive cross-group externality on side 1 users (say, merchants) than *vice versa*. In this case, competition leads to side 2 users being targeted even more aggressively by platforms than would be the case under monopoly (Armstrong, 2006).

There are two key factors driving this result.

- **Value creation:** Each extra user on one side of the market is even more valuable than otherwise, as it brings in additional users on the other side too.

- **Rival harming:** Also, every consumer a platform wins from its rival platform reduces the perceived value of that rival, further enhancing the relative attractiveness of the first platform.

A further key factor affecting competitive outcomes is whether both sides of the market are single-homing, or one side is multi-homing.[17] As discussed above, single-homing users on one side of a platform (say, cardholders) can create 'bottleneck' or 'gateway' market power for that platform vis-à-vis users on other side (say, merchants). In the context of card payment schemes, Vickers (2005) describes this in terms of certain cards being 'must have' from the perspective of merchants.

With single-homing on both sides of the market, then – relative to the situation with no cross-group effects – platform profits are reduced by the sum of the cross-group external effects. Also, the equilibrium price charged on each side is reduced by the cross-group external effect exerted by that side on the other. Thus, if one side's users (say, cardholders) are especially important for driving participation on the other side (say, merchants), the price for the former will be reduced by more (and could even lead to these prices being negative).

By comparison, once one introduces multi-homing on one side, the resulting 'bottleneck' market power will tend to raise prices and revenues on the multi-homing side (in cards, merchants), relative to the situation with single-homing on both sides. These higher revenues

---

[17] As discussed above, in the situation where *both* sides are fully multi-homing, cross-group network effects cease to have a significant effect.

will typically be passed onto the still single-homing side (in cards, cardholders) as lower prices.[18,19] Importantly, this could even lead to merchants facing higher fees under platform competition than they would under monopoly.

Finally, it should be noted that many of the models in this area assume rational users, who can deduce expected participation from users on the other side of the market and act accordingly. But what if users do not readily adjust their expectation of the participation by others so easily? Hagiu and Halaburda (2014) show that this leads to lower elasticity and higher equilibrium prices. This assumption arguably seems realistic and indeed Hurkens and Lopez (2014) study the role of expectations in the context of mobile telephony and show that such passive expectations provide a better model of actual user behaviour.

1.4 Market definition in platform markets

As will be clear from the above discussion, platforms can exhibit two distinct forms of market power.

- First, we may see platform markets tip to being highly concentrated with strong barriers to entry and expansion. In this case, we can think of a platform as having market power in a single **platform market**.

   The position of Google as an online general search platform might be a good example here. Indeed, Google was found to be dominant in this market by the European Commission in its *Google Shopping* decision.[20]

- Second, we may see competition between platforms but still observe these platforms having 'bottleneck' market power in relation to **one side of a platform market**. Most typically, this occurs in relation to multi-homing users who need the platform in order to access its single-homing users on the other side.

   The Google Play app store might be a good example here. App developers have little option but to use this if they wish to access Android phone users. We can thus think of Google as having market power in the market for Google app stores, even though Android phones compete with Apple phones more broadly. This was again the view reached by the European Commission in its *Google Android* decision.[21]

---

[18] In fact, relative to the situation with single-homing on both sides, prices can potentially also increase for users on the side that remains single-homing (cardholders) (Belleflamme and Peitz, 2019). This is because the platform no longer gains any cross-group effect of winning more merchants when it gains more cardholders, as it already serves all the merchants. However, this will only occur in limited situations (formally, if the remaining cross-group effect (from the multi-homing side to the single-homing side) is smaller than the degree of differentiation between platforms (from the merchant perspective). This is relatively unlikely to hold for the major card schemes, given that they offer very similar services from a merchant perspective, while merchants still gain high value from additional card holders.

[19] In many real markets, we may not observe all users on one side multi-homing. But same qualitative conclusions hold if there is some single-homing, with more single-homing driving up the prices charged on the multi-homing side (Rochet and Tirole, 2003; Doganoglu and Wright, 2006; Belleflamme and Peitz, 2019). This is not relevant for the major payment schemes where all merchants where all merchants effectively multi-home. But may be more relevant for considering the position of alternatives such as American Express.

[20] Commission Decision of 27 June 2017, Case 39 740 *Google Search (Shopping)* which has been upheld by the General Court in Case T-612/17 *Google v. Commission*, EU:T:2021:763.

[21] Commission Decision of 18 July 2018, Case 40 099 *Google Android* which has been upheld (in part) by the General Court in Case T-604/18 *Google v. Commission,* ECLI:EU:T:2022:541.

However, the latter of these approaches to market definition has provoked controversy, most notably in the recent US Supreme Court's judgment in *Ohio v American Express Co*.[22] In this case, the Court hold that "two-sided transaction platforms, like the credit card market […] facilitate a single, simultaneous transaction between parties". They therefore determined that Amex's case should be assessed on a single two-sided market, which in turn led to it finding no infringement. The US Supreme Court was heavily influenced in its findings by Filistrucchi et al (2014) who explicitly argue that "in two-sided transaction markets, only one market should be defined."

This judgment raises the question of whether 'transactions markets' (such as payment cards) are so very different from 'non-transactions' markets (such as 'subscription' or 'media' markets) where the monetisation is not directly linked to any transaction value. Many commentators have argued that this is a false dichotomy and have criticised the Supreme Court's decision.[23]

Their arguments are partly on *factual* grounds: credit card schemes don't only provide transactions and they don't only charge on a usage basis. They also carry out credit-worthiness checks, insure against fraud, allow for refunds, and they charge a variety of fixed fees alongside transactions-based fees.

However, they also challenge the *theoretical* basis for making so strong a distinction. The core argument for treating them differently relates to the argument that, in a transactions market, any transactions fees charged to merchants can simply be passed back to consumers by those merchants, and thus there is no real distinction between imposing a fee to consumers directly or imposing it on them indirectly, via charging merchants.

The limits of this 'neutrality' argument are discussed further in Section 2. What is clear is that it is a special case which requires very specific conditions to hold. In practice, merchants are unlikely to be able to fully pass on such increased fees in the way proposed. And in this case, the choice by platforms (including transaction platforms) as to how to price to different groups of users can have significant impacts.

This matters because parties on different sides of a 'transactions' platform may have very different alternatives available to them. For example, the ride-hailing platform Uber meets the definition of a 'transaction platform' in the Filistrucci et al (2014) paper, but its drivers have very different substitution possibilities to passengers. Moreover, some transactions platforms in fact compete with single-sided platforms. For example, the two-sided platform Amazon Marketplace, which again meets the definition for being a 'transaction platform, arguably competes for customers with standard retailers such as Walmart. To ignore this by taking a single market approach would seem misleading.

---

[22] *Ohio v American Express Co*, 585 U.S. (25 June 2018).

[23] See Niels (2019), Hovenkamp (2019). Katz (2019) and Franck and Peitz (2019)

2. Competition in card payment schemes

In this section, we focus in on four-party payment card schemes, highlighting some special economic characteristics of these relative to standard platforms. We discuss the implications for market outcomes, and the efficiency of those outcomes, before turning to proposals for regulating the Multilateral Interchange Fee (MIF).

In doing so, we draw not only on the general economic literature on platforms outlined above, but also on economic literature that is more specifically related to this market context. Much of this literature focuses on the Multilateral Interchange Fee (MIF). Earlier contributions are well-summarised in Rysman and Wright (2014).

2.1 The special case of payment card schemes

The basic economics of four-party payment card schemes was described in the Introduction to this report. There are clearly important economic similarities between these and the more standard platforms discussed in Section 1. In particular, there are strong cross-market network effects between card users and merchants. For example:

- In relation to *consumer card holding/usage being a key driver of merchant card acceptance:*

  - Carbó-Valverde et al (2012) use quarterly bank-level information for the two biggest Spanish card networks over the period 1997-2007. They find that a 1% increase in the number of cardholders of the network for which a specific bank is an acquirer produces a 2.7% increase in merchant demand for the terminals of that bank.

  - Boumie et al (2017) exploit shopping diary data in France and find that consumer preferences in relation to card use are the strongest driver of merchant card acceptance.

  - Arango and Taylor (2008) exploit a Canadian merchant survey and find that the acceptance levels of payment instruments do not reflect merchants' own relative preferences, but rather – within certain limits – they will accommodate consumer preferences.

- In relation to *merchant card acceptance being a key driver of card adoption/use*:

  - Rysman (2007), using Visa data from the US, finds the extent of merchant card acceptance is critically important for consumers' choice of network.

  - Arango et al (2015) use transaction-level data to examine consumers' choices of payment instrument. They find that merchant acceptance is a strong driver of card usage, especially for smaller transaction values.

- In a recent paper, Iorio and Rocco (2022) examine the drivers of cash usage (as opposed to cards) in Italy. They find that lack of merchant card acceptance remains a key driver, with consumers unable to use cards in many places where they would prefer to do so.

However, there are also two important differences between standard platforms and payment schemes.

(i)   **Payment schemes are associations of members.** A key difference relates to the fact that platforms typically face both sides of the market at once and set their own fees to each side unilaterally. By contrast, and as described in the introduction, payment schemes are *associations of members*. The 'platform' part of a four-party payment scheme in fact comprises multiple issuer banks and multiple merchant acquirers (also typically banks). The issuer banks set fees for cardholders unilaterally, while the merchant acquirers set merchant service charges (MSCs) for merchants unilaterally. Finally, there is a MIF charged from issuer banks to merchant acquirers. The MIF could theoretically be set unilaterally or negotiated bilaterally, but is in fact set jointly by scheme members.

(ii)   **Pass-through from merchants to consumers.** Another key difference is that there is typically a degree of pass-through of MSCs by merchants back to consumers as higher prices. While this sort of pass through can potentially be occur for standard platforms too, it is typically ignored in much of the economic literature on platforms. This pass-through has important additional effects to those discussed above, as will be seen below.

Each of these plays a critical role in understanding market outcomes for payment card schemes.

(i)   Implications of card payment schemes being associations of members

The first of these differences is important from an antitrust perspective, as any joint setting of prices between rivals can potentially be viewed as anti-competitive. There are, however, two key reasons why card associations differ from standard cartels, and indeed could act to improve welfare.

- **Solving a 'free rider' problem.** Card schemes typically incorporate some form of 'honour all cards' rule. This rule means that, for a given card type, merchants cannot pick and choose between card issuers. If they accept cards of that type, then they have to accept all issuers or leave the scheme altogether.

   In deciding whether to accept that card type, the merchant will then consider the average MSC they face under the scheme, across all cards in the scheme, and this in turn will depend on the *average* MIF across issuing banks. But this means, from the point of view of any issuing bank acting unilaterally or negotiating bilaterally, its own choice of MIF would only have a limited impact on the overall average MIF and thus on card acceptance.

If any particular issuer increases its interchange fee, it thus captures the benefit of the higher fee, but the harm caused by that increase - lower merchant acceptance of cards - will be shared by all issuers. This creates a free-rider problem. Each individual issuing bank would unilaterally wish to set its MIF very high, even though this is not in the interest of all the relevant issuing banks jointly.

Seen in this context, a collectively set MIF solves this free rider problem and thereby keeps interchange fees *lower* than would be the case otherwise, benefiting both scheme members and merchants (Small and Wright, 2002; Rochet and Tirole, 2002). Indeed, it can reasonably be argued that a common interchange fee may be necessary in order to maintain an 'honour all cards' rule, which in turn provides value to scheme users (Klein et al, 2006).

- **Allowing pricing to reflect network effects.** Building on the thinking relating to platforms in the previous section, it can be argued that pricing in any two-sided market might reasonably diverge from the costs of serving the different sides of the market, in order to reflect the relative balance of cross-market network effects between the two sides. If so, then it is argued that the MIF is a key way of doing this. It is an important revenue source for issuing banks and strongly influences the fees they charge to cardholders. It is also an important cost for merchant acquirers and strongly influences the MSCs they charge to merchants.

  Seen in this context, the MIF plays a critical role in allowing payment card schemes to act more like standard platforms, enabling a pricing structure that reflects network effects.

The first of these two arguments might sound compelling. Indeed, the 1986 US *Nabanco* ruling gave a clean bill of health to a collectively set MIF on this basis.[24] However, more recent thinking has challenged this view, in particular on the basis that the counterfactual to a collectively set MIF need not be higher interchange fees. For example, there could be a zero fee.[25]

The second argument is no less controversial. While the platform pricing described in the previous sub-section may not reflect any anti-competitive intent, this does not mean it is necessarily efficient. Specifically, where there is competition between schemes, both of which have 'bottleneck' market power, prices to multi-homing users (here, merchants) are likely to be inefficiently high.[26] This will be discussed further below.

---

[24] *National Bancard Corp v VISA USA Inc* 779 F 2d 592 (11th Cir. 1986).

[25] In fact, the European Commission came up with a subtly different counterfactual in its 2007 *Mastercard* decision: 'a rule that imposes a prohibition on ex post pricing on the banks in the absence of a bilateral agreement between them'. However, in practice this is tantamount to a MIF of zero (*MasterCard* (Case COMP/34.579), Decision of 19 December 2007). This decision was upheld on appeal (Case T-111/08 *MasterCard and Others v Commission*, Judgment of 24 May 2012; and Case C-382/12 P *MasterCard and Others v Commission*, Judgment of 11 September 2014.)

[26] Indeed, Wright (2012) disputes that the conduct of the card schemes should be viewed as anti-competition or addressed through competition law, but nonetheless argues that the MIF should be regulated, as it is otherwise likely to be set too high.

(ii) Implications of pass-through from merchants to consumers

In certain circumstances, the potential for perfect pass-through from merchants to cardholders, via surcharging, means that market outcomes will in fact be entirely unaffected by the level of MIF (Gans and King, 2003). The logic behind this 'neutrality result' is appealing. On the issuer side, a higher MIF may be fully passed through to cardholders as rewards. On the merchant side, with perfect surcharging of fees to cardholders, the MIF may be fully passed through to cardholders as higher retail prices. The higher retail prices and rewards exactly offset, and leave all parties indifferent, other than merchants if they gain convenience benefits from accepting cards.

However, this neutrality result depends on several critical assumptions (perfect competition amongst merchants[27], amongst merchant acquirers, and amongst issuers, the employment of usage fees (rather than fixed fees), and perfect observability of surcharges by consumers. In practice, however, these conditions are unlikely to hold in many situations, and thus the neutrality result cannot be relied upon. As will be discussed in Section 4, merchants may have incentives to set surcharges that are higher or lower than the MSC they face, to shroud any surcharges (so that cardholders do not react to them), or not to surcharge at all. Moreover, issuers (and indeed merchant acquirers) may set two-part tariffs comprising a fixed membership fee as well as a usage fee.[28]

In practice, much of the literature on payment card schemes assumes an *absence of surcharging* (that is, uniform retail prices, irrespective of payment method).[29] In this case, the implication of pass-through is somewhat more subtle. In particular, it allows for an effect known as 'merchant internalisation' (Rochet and Tirole, 2002; Farrell, 2006).

To the extent that consumers benefit from card-holding, they will be willing to pay a higher price for the merchants' services. This in turn means that merchants benefit not only from any convenience benefit they receive directly themselves from using cards, but also any benefit consumers receive. That is, they internalise benefits from both sides. Their willingness to pay for card acceptance is thus higher than simply their own benefits, but also includes the benefits arising for consumers.

At the same time, with no surcharging, those consumers who do not utilise a particular payment scheme will nonetheless partly fund that scheme, due to shared costs feeding into retail prices. The combination of these various factors can increase the incentives of any one payment scheme to increase its charges to merchants, since not all of the increased costs fall on users of that scheme.

---

[27] Wright (2003) shows that this only holds if merchants are competitive. If they are monopolistic, they will set surcharges that are 'too high'. He argues that a 'no-surcharge rule' may thus improve welfare.

[28] Bedre-Defolie and Calvano (2013) analyse the pricing of a monopoly payment card scheme within with issuers set two-part tariffs. In their model, there is full pass-through of the usage charges, but not the fixed fee.

[29] Although Bourgignon et al (2019) and Edelman and Wright (2015) explicitly consider merchants' surcharging choice.

2.2 Market outcomes and implications for user surplus and welfare

This sub-section builds on these introductory remarks by considering expected market outcomes in payment scheme markets in more detail. It provides a brief history of the key thinking on this topic, before bringing the debate up to date.

Early thinking

In an article that was well ahead of its time, Baxter (1983) provided the first formal modelling of the MIF. He provided a theory of welfare-maximizing interchange fees in which there is perfect competition of issuers and acquirers, and so both obtain zero profits.

In a simple theoretical setting in which there are just two payment alternatives (say cards and cash) that can be used to complete a single transaction, he notes that efficient card transactions arise when consumers use cards if and only if the joint convenience benefits from doing so (i.e., theirs and the merchants') exceed the joint costs to the corresponding issuer and acquirer from providing the service. However, with interchange fees set at zero and absent surcharging, the cardholder will face a fee equal to the issuer's cost and the merchant will face a fee based on the acquirer's cost. This will not, except by coincidence, lead to the efficient usage of cards. An externality will arise.

In this framework, the MIF plays an important role in aligning prices with benefits for both cardholders and merchants, and thereby enhancing welfare. This requires an interchange fee set equal to the merchants' convenience benefit of accepting cards, minus the acquirers' cost of providing its service. The logic behind this 'Baxter Interchange Fee' is intuitive. If merchants are charged an MSC that equates to this MIF plus the acquirers' costs, they will be indifferent ex-post (i.e., at the point of sale) between whether consumers use cards or the relevant alternative (e.g. cash). This principle is known as the 'merchant indifference criterion' (Farrell, 2006).

Baxter's early thinking underpins much of the later literature. While much of that literature analyses the situation with a monopoly payment card scheme (which clearly does not reflect reality and is not discussed in detail here), more recent papers analyse more realistic duopoly situations.

Rochet and Tirole (2011)

The classic paper on market outcomes in duopoly payment card markets is Rochet and Tirole (2011)[30]. Similar outcomes are identified in Guthrie and Wright (2007). Both papers consider the expected level of a MIF that is jointly set by issuers, under various market structures, in the context of uniform retail pricing (no surcharging). They also assume a competitive merchant acquiring market which makes zero profit margin. However, their findings are somewhat different.

---

[30] Note that the same paper is sometimes cited as Rochet and Tirole (2006) as the core of it was written at that time. Also, in an earlier seminal paper, Rochet and Tirole (2002), the same authors had already introduced some of the key ideas within the later paper and had analysed the case of a single card platform with homogenous merchants.

Given the influence of Rochet and Tirole (2011), we consider their findings in some detail here. The paper provides a usefully generalised set-up, which does not make too many restrictive assumptions and is explicit in relation to those that it does make.

The paper first considers the claim by Vickers (2005) that cards are 'must take' for merchants. They highlight that merchants might consider card acceptance rather differently at the initial adoption stage and then at each transaction.

- At the *adoption* stage, merchants care about the additional consumers that they will win from accepting a card. Rochet and Tirole show that there is in fact a limit to the merchant service charge (MSC) that merchants will pay to accept cards. Their 'maximum acceptable MSC' will be the sum of [the convenience benefit the merchant gains from accepting cards] plus [the value cardholders gain from using them]. This is based on the 'merchant internalisation' thinking described above.

- At the *transaction* stage, however, merchants care only about MSC relative to their own convenience benefit from using the card. Thus, if the MSC lies above this convenience benefit, this can feel to merchants as though cards are 'must take', even if the MSC is below their 'maximum acceptable MSC'.

Next, the paper examines the optimal level of MIF in a situation where all merchants are assumed to receive the same convenience benefit from accepting cards.

- The paper starts by re-defining the 'Baxter Interchange Fee' as the '***tourist test threshold***' (a^T). This is given by [the convenience benefit a merchant gains from accepting cards] minus [the costs of merchant acquiring]. The rationale for the terminology is that the so-called 'merchant indifference test' (which effectively underpins the Baxter Interchange Fee') could potentially be affected by a merchant's desire to retain repeat consumers. The 'tourist test threshold' abstracts from this by considering 'merchant indifference' in the context of non-repeat 'tourist' consumers.

- With **constant issuer margins**, the MIF that maximises *total user (consumer plus merchant) surplus* (a^{TUS}) is precisely equal to the 'tourist test' threshold (a^T). This is also the welfare-optimal MIF if issuers have zero margins (as shown by Baxter). However, if issuers have positive but constant margins, the *welfare-optimal* MIF threshold (a^W) will be higher than the 'tourist test threshold' by the amount of this issuer margin.

- With **variable issuer margins**, the situation becomes more complex. It becomes critical whether issuers margins increase (cost amplification) or decrease (cost absorption) with the increases in the cost of issuing. With cost amplification, the MIF that maximises *total user surplus* (a^{TUS}) exceeds the 'tourist test threshold' threshold (a^T). With cost absorption, it is below this threshold.[31] However, in both cases, the *welfare-optimal* MIF (a^W) lies above the 'tourist test' threshold.

---

[31] It should be noted that cost absorption is arguably more realistic in most scenarios, but cost amplification can occur when inverse demand is sufficiently convex or marginal cost curves slope sufficiently downwards (strong and increasing returns to scale).

They then considers the expected MIFs that would arise at equilibrium in different market contexts. In the duopoly situation, all merchants are assumed to be multi-homing. The results below have strong similarities to those identified in Section 1 of this paper.

- Under **monopoly**, the MIF (a^m) will be set *higher* than is socially optimal (a^W), so long as (as is likely) issuer margins (m) are lower than the benefit gained by cardholders from using cards v_B. At this monopoly level, the MIF will extract all of the cardholder benefits from using cards, leaving total user surplus of zero.

- Under **platform duopoly with multi-homing by customers** (ie where customers hold both cards), the MIF (a^{MH}) will be that which maximises total user surplus(a^{TUS}).

- Under **platform duopoly with single-homing customers**, the MIF (a^{SH}) will be set at the monopoly level (a^m). This reflects the 'bottleneck' market power that the platform holds over multi-homing merchants in terms of accessing single-homing customers.

The following figure, from the paper, summarises these findings (for a constant but positive issuer margin, m).
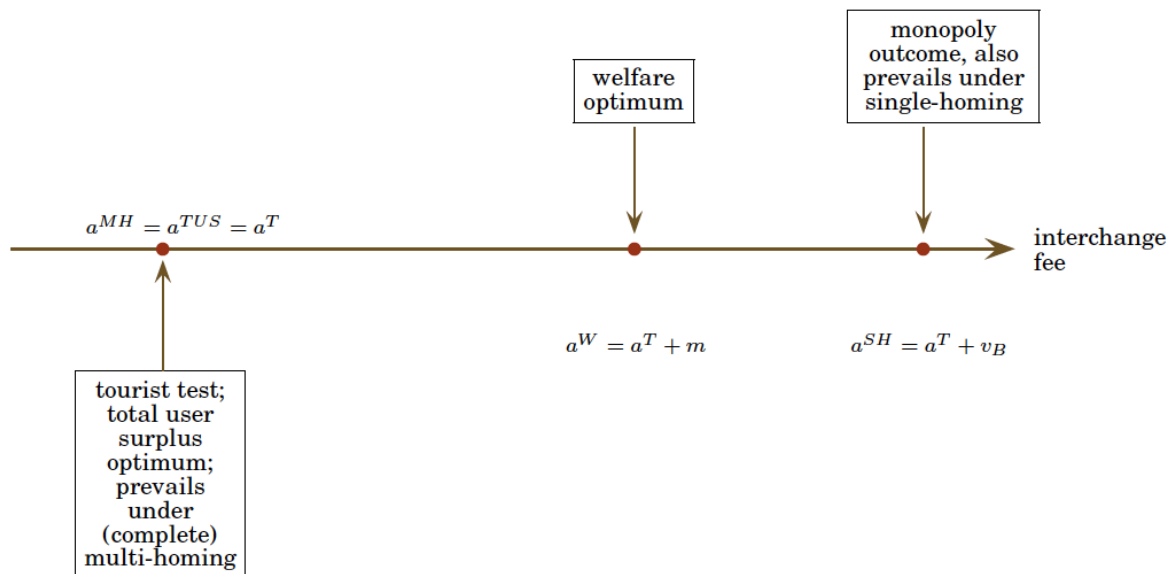
So far, so good. The paper seems to confirm our intuition that a jointly set MIF in a duopoly with single homing cardholders will be higher than that which maximises welfare and significantly higher than that which maximises total user surplus.

However, a limitation of this simple model is that it doesn't in fact allow for network effects arising from increasing the number of merchants who accept cards, as all merchants have the same valuation of card acceptance and thus there is complete market coverage. In practice a higher MIF might be expected to reduce merchant take up by those merchants who receive less of a benefit from card acceptance. And if consumers value a wider merchant network, this may act to limit the level of the profit-maximising MIF.

The paper therefore goes on to consider the implications of *heterogeneous merchants*, who differ in the convenience benefit they receive from card acceptance. It only analyses the monopoly situation with constant issuer margins, but in this case it finds that:

- *Total user surplus* is now maximised by setting the MIF such that it equates to the average convenience benefit of those merchants accepting cards, minus acquirer costs. This is described as an '*average tourist test threshold*' (although note that only the average merchant in fact satisfies the 'merchant indifference test').

- This 'average tourist test' is also *welfare-optimal* when *issuer margins are zero*.

- Under *monopoly*, with constant issuer margins, the MIF will be set higher than this 'average tourist test' if and only if average *cardholder* surplus per card payment exceeds that of *merchants*. This is clearly an empirical question.

The paper also looks at the potential for entry into issuing. Again, this changes the finding of the short run (no entry) model in a significant way. The MIF which maximises *long run* total user surplus may be higher than that which maximises it in the *short run* (albeit lower than the MIF which maximises short run total welfare).

Reflections on Rochet and Tirole (2011)

While theoretical models rarely accord perfectly with reality, we note that in New Zealand, we roughly observe a card scheme duopoly with (mostly) single-homing consumers and multi-homing merchants.[32] We do not see significant evidence of issuer entry. And we would expect to see positive issuer margins, which are either constant or exhibit cost-absorption.

In this scenario, Rochet and Tirole (2011) – and with homogeneous merchants – concludes that we should observe:

$$a^{SH} > a^W > a^T \geq a^{TUS}$$

Thus, we should expect the actual observed MIF in New Zealand, absent regulation, to exceed the welfare-optimal MIF, which should in turn exceed the tourist test, which should in turn exceed (or equal for a constant issuer margin) the MIF that maximises total user surplus.

Moreover, there are five additional factors that could inflate the MIF still further.

- **Merchant heterogeneity in duopoly.** Rochet and Tirole (2011) only consider the impact of merchant heterogeneity in the context of monopoly but Guthrie and Wright (2007) allow for duopoly. With merchant homogeneity, their results are the same as those above. That is, under duopoly with single homing consumers (and thus bottleneck market

---

[32] It should be noted that such single-homing by cardholders is not always prevalent. Indeed, Rysman (2007) finds that the majority of US households hold cards from multiple networks (at that time), with only 36 percent of the households holding cards from just one of the networks (almost always Visa or MasterCard). That said, even in that case, he finds that the results are very different in relation to usage, with 75 percent of households putting more than 97 percent of their spending on a single network.

power), the MIF with multi-homing merchants will be the same as under monopoly card scheme.

With merchant heterogeneity, by contrast, they show that competition between card schemes can increase the chosen MIF in a duopoly with bottleneck market power to above that set by a single monopoly scheme. Indeed, competition can lead card schemes to set interchange fees too high for their own good. In effect, each card scheme sets its interchange fee too high in an attempt to get buyers to switch to holding its card exclusively, an effect which ends up reducing the card holding and card transactions.

- **Negative externality on non-scheme consumers.** Given a lack of surcharging, any rise in the MIF of one card scheme will be incorporated into the price across all consumers, and not just cardholders of that particular scheme. Effectively, this means that a scheme's MIF rise confers a negative externality on non-scheme consumers. Edelman and Wright (2015) how that this limits the competitive constraint on each scheme's choice of MIF, leading to MIFs being set excessively high.

- **'Merchant internalisation'.** Wright (2012) argues the model utilised by Rochet and Tirole (2011) does not in fact fully account for 'merchant internalisation' in the context of no surcharging.

  'Merchant internalisation' occurs because, while cardholders only consider their own benefit from using cards, merchants consider not only their own benefits but also cardholders' benefits (since the latter allow them to extract more surplus from cardholders). Since the MIF will be set partly in relation to the benefits on each side of the market, Wright argues that this 'double counting' of cardholder benefits has the systematic effect of inflating the MIF.

- **The potential for price discrimination on the issuer side.** Abstracting from the issue of surcharging, Bedre-Defolie and Calvano (2013) examine the implications of the fact consumers make two decisions (whether to hold cards and then whether to use them) whereas merchants typically only make one (whether to accept them).

  Their insight is that, because consumers make two decisions, issuers will be able to extract a greater share of cardholder surplus (their benefit from using cards) than it can of merchant surplus. The card scheme thus has an incentive to maximise cardholder surplus (so it can then extract this back), which in turn involves setting a high interchange fee that is biased towards cardholders.

- **Imperfect information:** Consumers may have strong preference for using a card, but do not know, when choosing merchant, which merchants will accept cards or, if they do, whether a surcharge will be levied.

  Bourguignon et al[33] (2019) model this situation. Although consumers are assumed to have rational expectations as to this policy, this imperfect information can generate 'missed

---

[33] It may be relevant to note that one of the co-authors of this paper is Jean Tirole who, with Jean-Charles Rochet, provided the original explanation of the 'must take' nature of cards.

sales'. A missed sale occurs when the customer is in the shop and eager to buy, but has a high cost of paying by cash, and is discouraged by either a high card surcharge or an outright rejection of the card.

Such missed sales induce merchants to feel they 'must take' the card, which makes merchant demand for the card scheme more inelastic , further increasing the MIF that is charged. Note that the importance of missed sales is underscored by the analysis of Bolt et al. (2010), who, using survey data from the Netherlands, document that 5% of consumers reported leaving a merchant's store without purchasing when faced with card refusal or steep card surcharges.

Bourguignon et al (2019) find that this additional 'must take' effect reduces merchant resistance (the card acceptance threshold increases) when cards are especially valuable to consumers relative to cash. They also argue that there is less merchant resistance to credit than to debit cards and thus the card scheme should be charged a higher MSC for credit than for debit cards (as we see in practice).

Additional aspects

In addition, Rochet and Tirole (2011) and related theoretic models focus primarily on just one or two payment card schemes. In doing so, they abstract away from a number of important additional aspects of payment card scheme.

First, they *do not typically distinguish between debit and credit cards*. This may have important implications. Rochet and Wright (2010) highlight that it may be inefficient for consumers to use credit cards for 'ordinary purchases' if credit cards are higher cost than debit cards. This could militate against higher interchange fees being charged for credit cards, since these will be passed back to consumers as rewards and incentivise credit card use.

Second, *they do not consider the impact of rival three-party scheme*s (such as American Express). The implications here are less clear, in theory at least:

- If the interchange fee for four-party cards is regulated, but not that for three-party cards, there is a risk that the latter will charge higher fees to merchants and thus offer rewards to consumers, encouraging them to use those cards too much.

- On the other hand, the incentives on merchants to accept expensive three-party cards may be reduced if they already multi-home on two cheaper four-party card options and most consumers have at least one of these two cards, potentially leading to three-party card schemes also reducing their merchant fees.

In practice, the empirical evidence to date is mixed, but predominantly supports the latter view. In Australia, the introduction of regulation led to the three-party cards doing deals with banks to offer their cards as 'companion cards' alongside the core four-party cards, in order to increase cardholding. Initially, in Australia, this led to a substantial increase in the share of three-party cards. However, the introduction of regulation for these 'companion cards' in 2017 ended this arrangement. Since then, their market shares have reduced again, and so

have their merchant fees.[34] Their share has also decreased (slightly) in the EU, following the introduction of the IFR.[35]

Third, they abstract away from the *multi-dimensional pricing of cards* by issuers. While overall charges to cardholders are likely to rise in response to a lower MIF, this may feed through in a straightforward way.

Morris et al (2022) examine the impact of the EU IFR on fees charged by EU issuing banks by comparing total fees in 2014 and 2018. They find that, contrary to expectations, annual fees for credit and debit cards actually fell following the IFR. Other costs relating to cards did increase to broadly offset the loss in interchange fees, but these were primarily interest charges and late and overdraft fees. These would be expected to have a rather different effect on cardholder adoption and usage than standard fees.[36]

Fees	2014			2018			
	Credit	Debit	Combined	Credit	Debit	Combined	Change
Annual Fees	9501	14620	24121	8741	14036	22777	-5.57%
Interest, late & overdraft	13469	8728	22197	15250	15352	30602	37.87%
International transaction	292	985	1277	382	1180	1562	22.32%
Interchange	4343	7527	11870	2258	4474	6732	-43.29%
Total	27605	31860	59465	26631	35042	61673	3.71%

Source: Morris et al (2022)

Similarly, survey evidence reported by Iranzo et al (2012) suggests that issuers in Spain increased the consumer costs of using credit cards following the reduction in interchange fee, and this went a long way towards offsetting what they lost in interchange fees. But the majority of this (€2.6 billion) took the form of increased interest rates. While credit card fees did increase, this only generated an additional €1.7 billion.

Fourth, they abstract away from the fact that *many issuers are multi-functional retail banks*. These make complex pricing decisions across their various services, and in this context is not clear that the impact of lower interchange fees will fall purely on cardholders. For example, two studies of the 'Durbin Amendment' to the 2010 US Dodd-Frank Act, which regulated debit card interchange fees, found that free banking became rarer.[37] Mukharlyamov and Sarin (2022) find that the proportion of free accounts in affected banks fell from 61% to 28%, and offset around 42% of issuers' lost interchange revenues. Kay et al (2018) estimate an even higher extent of offsetting via wider bank fees of around 90%.[38]

---

[34] Reserve Bank of Australia (2021). *Review of Retail Payments Regulation – Conclusions Paper*.

[35] European Commission (2020), *Study on the Application of the Interchange Fee Regulation: Final Report*.

[36] Moreover, overdraft charges are arguably general bank fees, rather than card-related, and thus should arguably be excluded, reducing the extent of offsetting. See the following point.

[37] Under the 'Durbin Amendment', debit card MIFs, previously averaging 2 percent of transaction value, were capped at $0.22, decreasing collective bank revenues by $5.5 billion annually. The rule only applied to banks with assets of over $10 billion, allowing the study to utilise difference-in-difference methods to compare those banks who were impacted with those who were not.

[38] See also Manuszak and Wosniak (2017).

We note that the overall welfare implications of interventions become more complex in this situation, as it is necessary to consider the wider impact on consumers beyond their card user. Evans et al (2015) examine event-study analysis of stock prices and conclude that consumers lost more in terms of increased banking costs than they gained from lower consumer prices.

Fifth, the economic literature has not so far focused on fact that MasterCard and Visa also charge switch fees (also called network fees) to issuers and acquirers. However, as Rysman and Wright (2014) highlight "the ability to set switch fees to issuers and acquirers would seem to make interchange fees redundant, since positive switch fees could be set to acquirers and negative switch fees (i.e. subsidies) could be set to issuers, to mimic any interchange fees." In practice, as noted in the Introduction, the level of these fees is currently relatively low, relative to the MIF. However, there is increasing regulatory interest in these network fees, and the UK Payment Systems Regulator is currently investigating.[39]

---

[39] UK Payment Systems Regulator (2023) *Market Review Into Card Scheme and Processing Fees*.

3.   Evidence on efficiency and innovation

In this section, we review the available economic literature on efficiency and innovation, in particular associated with the MIF. Much of this literature is based on assessing the impact of introducing MIF regulation across different jurisdictions.

3.1   Allocative efficiency

In many markets, any concerns about excessive prices are equally concerns about excessive profit. However, this is not the case in payment systems. Indeed, a key insight from the platform literature summarised in Section 1 is that, with competition between platforms, network effects tend to increase price competition. So the discussion above gives us no particular reason to sorry that overall profits from cards are too high. The concerns are about pricing structure, not overall level.

The theoretical analysis in Section 2 suggests that MIFs would naturally be set at too high a level to maximise total user surplus, and perhaps even total welfare (which includes scheme profits). In general, total user surplus from cards will tend to be highest when there are high levels of card usage.[40] This will in turn tend to be driven by high levels of merchant card acceptance and consumer cardholding.

An obvious empirical question, therefore, is how these different elements have been affected by the various regulatory interventions we have seen in relation to MIFs over the past decade.

A positive impact from lowering the MIF

In fact, a number of studies show that these measures have either been unharmed or increased, which would tend to support the core theoretical finding that a decrease in MIFs from their privately set level is beneficial. For example:

- Carbó-Valverde et al (2016) evaluate the impact of four regulatory decisions in Spain that reduced interchange fees from 1999 onwards. Using proprietary quarterly payment card data from 1997-2007, these were found to have:

    (i)   A positive impact on merchant acceptance. Specifically, a 10% reduction in the MIF was found to increase merchant acceptance by 1.4%.

    (ii) A strong increase consumer credit card adoption and no significant negative impact on consumer adoption of debit cards[41].

    (iii) A dramatic increase in payment transactions; a 10% reduction in the MIF was found to increase transaction volumes by 1.7%.

---

[40] With the caveat that the negative externality imposed on non-scheme consumers can lead to excessive scheme take-up and usage, as shown in Edelman and Wright (2015).

[41] The authors note that the impact of MIF changes on debit card adoption is likely to be relatively limited, since these are typically bundled with other transaction services, including also being used as ATM cards.

(iv) This impact was found to be so positive that overall bank revenues for credit and debit cards also increased as a result of the MIF reductions.

- Ardizzi and Zangrandi (2018) use institution level data from merchant acquirers to examine the impact of the 2015 Interchange Fee Regulation (IFR) in Italy. This regulation reduced interchange fees by 37%, which led to a 22% in merchant service charges. They estimate that this led to an increase of 8-11% in the number of transactions per POS terminal and explains around 30-40% of the increase in merchant acceptance that occurred between 2015 and 2017[42].

- Of course, any analysis of regulation in a single country risks the results being distorted by other trends, such as the general development of digitisation and mobile banking services. This concern is addressed by the difference-in-difference approach adopted by Ardizzi et al (2021). They draw on data from across 46 countries over a decade (2010-19) during which the EU introduced the IFR. This study finds a significant negative relationship between MIF levels and the number and the growth rate of card-based transactions per capita. It also finds a strong and significant one-off impact of the EU Interchange Fee Regulation immediately after its introduction and considerable propagation effects in the following years.

  Interestingly, this study also finds that further reductions in the fees below the level set by the IFR may lead to an unintended decrease in transactions per capita. This effect is likely to derive from a reduction in card usage due to higher fees charged to the cardholders by those issuers penalized by the 'near-zero interchange fee'. For these reasons, in the context of payment services, setting interchange fees to zero (or negative values) does not seem to be an optimal choice to encourage card payments. Overall, therefore, the study concludes that policy actions aiming at containing, but not eliminating, interchange fees can significantly contribute to the diffusion of electronic payments.[43]

- Finally, a study by De Groen (2020) considers the impact of the IFR on the number of POS terminals. He finds that these increased significantly (a 6.8% increase in 2015 followed by increases of 9.4%, 9.5% and 12.3% in 2016, 2017 and 2018, respectively), suggesting a positive impact of the IFR on merchant card acceptance. (The EU's own review of the IFR also notes the continued increase in merchant card acceptance but is more caveated in terms of the extent of impact of the IFR).[44]

In addition, a key concern relating to high MIFs is that they may artificially inflate general retail prices. Shabgard and Asensio (2023) investigate the impact of the IFR on retail prices using Spanish sectoral data. They find that a 1% reduction in the MIF leads to a long run 0.17% reduction in the retail price index. They also find the MIF to have had negligible impact on payment card usage.

---

[42] An earlier study by Ardizzi (2013) did not examine a regulatory intervention, but rather used data on differences in MIFs between sectors in Italy (in 2010) to examine the implications of MIF level on the ratio of cash to card use. This finds that a 10% fall in the MIF is associated with a 1% reduction in the cash ratio, and thus increased the relative proportion of card use.

[43] A potential criticism of this study is that the data combines credit and debit cards, and there is thus a possibility that these exhibit rather different effects.

[44] See fn. 35.

These could appear perhaps surprising. After all, we might expect reductions in interchange charges to be at least partially offset by higher card fees and reduced rewards for cardholders, even if the extent of this may be limited by the points made above about multi-dimensional pricing of cards by multi-functional issuing banks.

Why, then, do we not seem to see a drop off in cardholders and transactions? The most plausible explanation would seem to be that cardholders care more about merchant acceptance (which should be expected to increase with the reduction of the MIF) than they do about their own fees and rewards.

The empirical evidence on the importance to cardholders of merchant acceptance was set out above. In terms of the role of fees and rewards in cardholder choice, the evidence is more mixed. Overall, it seems to suggest that the existence of a rewards scheme may be important, the level of those rewards rather less so.

- Ching and Hayashi (2010), Simon et al. (2010) and Carbó-Valverde and Linares-Zegarra (2011) all use consumer survey data (from the US, Australia and Spain respectively) to examine the effect of credit card reward programmes on the use of credit cards. They find a small but positive impact. However, these studies could also be criticised for relying on surveys. Also, they do not allow for variations in reward levels.[45]

- Arango et al (2015) use transaction-level data to examine consumers' choices of payment instrument. In line with the previous studies, they find that participation in a rewards programme is a strong driver of credit card usage (at the expense of both debit card usage and cash). However, they also find that the actual level of reward has little to no effect.

Possible distributional effects

The level of the MIF may also have implications for *distributional efficiency* across consumers.

First, the lack of surcharging for card payments can give rise to an important distributional issue. If merchants charge a flat retail price across consumers, then this will reflect their average cost of accepting payments, across all payment methods. If their net cost of cards is higher than their net cost of accepting cash, then this will inherently involve a cross-subsidy from cash to cards, and thus from cash users to cardholders. This will in turn occur if the MSC faced by merchants lies above the convenience benefit that they gain from accepting cards instead of cash.

A couple of studies examine these distributional effects:

- Schuh et al (2010) highlight that any re-distribution from cash users to cardholders is likely to increase inequality, given that cardholders tend on average to be richer than cash users. They estimate that the lowest income households pay an extra $21 per year in retail

---

[45] Ching and Hayashi (2010) and Carbó-Valverde and Linares-Zegarra (2011) also do not fully allow for the fact that card usage has historically been strongly increasing in transaction value (Bounie and François, 2006; Klee, 2008; Cohen and Rysman, 2013). They simply analyse the situation across different sectors which in turn tend to differ in relation to transaction size.

prices, while the richest households receive $750 a year in rewards and reduced card fees from this set-up. They conclude that a reduction in interchange fees would enhance welfare.

- More recently, Felt at al (2020) find a very similar a result for the US, on a more robust analytical basis. However, interestingly, they find an inverted U relationship in Canada. This is because the richest use debit cards, which are the cheapest form of payment overall, for medium to large payments.

Second, the level of the MIF may affect the availability of credit. A key argument for higher MIFs for credit cards is that they can be fed back to cardholders as lower fees and higher rewards, and thus encourage credit card use. However, there are mixed views on the benefits of increasing credit in this way.

- Brannon and Richardson (2022) claim that a lower interchange fee therefore makes credit cards less available to people with poorer credit scores. They argue that this is effectively a form of credit regulation and has very negative effects, which dwarf the lower prices arising from the lower interchange fee.

- By contrast, however, we know that that the behavioural characteristics of consumers can make them overly willing to take out credit (Agarwal and Zhang, 2015). In this case, limiting the availability of credit for people with poorer credit scores may actually be welfare-enhancing.

  Heidhues and Kozsegi (2010) show theoretically how this can in turn lead to a re-distribution from such myopic consumers to more savvy consumers. Agarwal et al (2023) provide supporting empirical evidence, showing that consumers can be unduly influenced by rewards.

  The following figure shows that overall bank profits are in fact substantially higher for cards that offer rewards, and the highest profits are made on 'near prime' consumers (those consumers just below 'prime' status, that have a credit score of 660-720), suggesting that these cardholders are getting the worst deal out of the credit card offering.

(B) Bank Profits
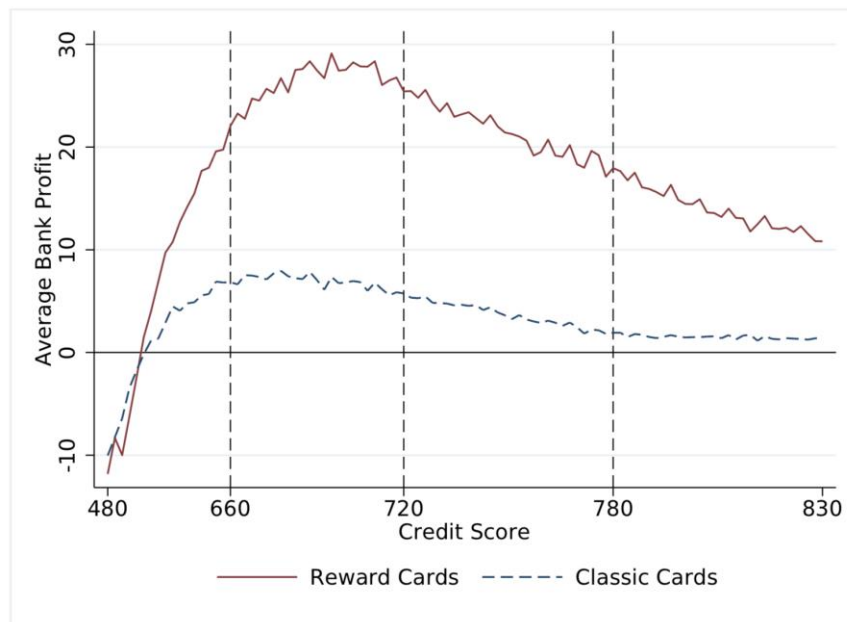
A final linked distributional issue relates to the recent development of 'cashless stores.' These are presumably more likely to occur when the MIF is low, and they can clearly have negative implications for those consumers who prefer to use cash or have no alternative but to do so. Shy (2021) examines this emergence of cashless stores, which he notes has led several cities and states to ban such stores. Using a random utility model, he simulates the effects on consumer welfare caused by a hypothetical complete transition to cashless stores. The simulations show that the harm from this transition to consumers with no credit or debit cards is seven times higher than the harm to consumers who have both cards.

However, this doesn't necessarily mean that the MIF should be raised. Rather it suggests that alternatives to cash may be needed before all brick-and-mortar stores become cashless.

3.2   Investment and innovation

The focus so far in this chapter has been on allocative efficiency. However, productive efficiency and dynamic efficiency are also important. Compared with the literature on pricing (allocative efficiency) in multi-sided markets, however, there is only a limited economic literature on implications for investment and innovation.

Investment in cost reduction

Belleflamme and Peitz (2016)[46] examine incentives to invest in *cost reduction* in the context of competition between two platform, with *single homing users on both sides*. They highlight that, absent any cross-group network effects, the positive 'direct' benefits to either platform from investing in lowering their own costs can be muted by a negative 'strategic' effect. This

---

[46] Their model is based on Belleflamme and Toulemonde (2018).

35

relates to the fact that, at equilibrium, the lower costs will result in both platforms cutting their prices, reducing the profit of both.

Introducing positive cross-user effects can change this situation. However, this can go in two directions. It can exacerbate the strategic effect, potentially to such an extent that investing in cost reduction becomes completely unprofitable. Alternatively, it can ameliorate the strategic effect, potentially to the extent that it becomes positive. In this latter case, platforms would have two reasons to invest in cost reduction, the 'direct' benefits arising, and also 'strategic' price-inflating benefits.

Overall, however, while this paper usefully highlights the complexities of investigating investment incentives in platform markets, its focus on competition with single homing users on both sides means it is arguably of limited relevant to most payment systems.

Investment in service enhancement

Reisinger and Zenger (2019) examine incentives to invest in *service enhancement.* Their model relates specifically to the MIF, albeit modelled as a single monopoly card scheme (with cash as an alternative payment tool). In that context, they find that the 'merchant internalisation' effect highlighted by Wright (2012) tends to enhance incentives to invest. This is essentially because part of the funding for investment, which benefits scheme members, comes from non-scheme members.

Indeed, they find that private investment incentives may be higher than the social incentives. Moreover, so long as service investments are important enough, the welfare-optimal MIF can actually lie above the MIF that the monopoly card scheme would set, a reversal of the standard finding. This is due the investment benefit that a higher MIF can bring. (Perhaps unsurprisingly, given the role played by the 'merchant internalisation' effect, they also find that removing the 'no surcharge rule' may have a negative effect in this context).

They also identify that over-aggressive price regulation of the MIF could harm such investment incentives and therefore be bad for welfare. This is certainly true if the regulated MIF is zero. They also find that, if consumer valuations are (roughly) uniformly distributed, the 'tourist test' interchange fee maximises total user surplus, once service investment is taken into account.

While it highlights important incentive effects, the relevance of Reisinger and Zenger (2019) to reality is limited by its assumption of a monopoly card scheme. Bourreau and Verdier (2019) to some extent fill this gap, analysing investment in service quality in a duopoly framework. They find that, when there are strong cross-group network effects, investment is highest with a zero MIF. However, with more limited cross-group network effect, investment is increased by raising the MIF from zero.

However, this paper effectively allows for surcharging (no 'merchant internalisation') which inherently removes the key driver of investment identified by Reisinger and Zenger. The model also appears to be one of single-homing on both sides, rather than multi-homing, and

thus there is no issue of 'bottleneck' market power arising, which would make the model a more realistic reflection of real card scheme markets.

Innovation

To some extent, we should expect the issue of innovation to be closely linked to the issue of market contestability (and thus barriers to entry and expansion). If a new entrant with an innovative new product find it hard – or even impossible – to enter the market, due to the established incumbent(s) benefiting from strong network effects, then this will clearly tend to inhibit innovation. Not only will the potential entrant be disinclined to invest, but neither will the incumbent(s), given the lack of challenge. The general risks to contestability in digital platform markets were discussed in Section 1.

Innovation can also happen in related or complementary markets. For example, in the digital sphere, there are concerns not only that the largest digital platforms have strong and entrenched market positions in their core markets but also that they are well positioned to extend this into emerging related markets. This reflects a number of 'ecosystem effects', including the fact that they not only exhibit 'within-market' network effects but also 'across-market' network effects (Fletcher, 2020). These are sometimes known as 'ecosystem effects'.

These can facilitate extension of market positions from one platform market into another platform market, a process known as 'platform envelopment' (Eisenmann et al, 2011). Such effects can be exacerbated through practices such as tying (of services together) or 'self-preferencing'.[47]

This raises the question of whether innovation in related markets may be affected by the conduct of the major payment card schemes. However, even if they are, interchange fees may well be only one element of the picture. For example, Yemail (2022) examines innovation in Latin America, after several interventions across different jurisdictions, including capping MIFs, prohibitions on exclusive agreements, and even forced divestiture. The author describes how these interventions helped create an environment which facilitated entry from a range of innovative payment solutions, such as direct bank transfers and digital wallets which allow for payments via links and QR codes.

A key issue in relation to innovation in payment systems is whether increased digitalisation will open up new competition to the existing payment card networks. This could potentially arise through two linked developments, both of which are being observed in financial services markets more generally.

First, the greater use of consumer data can facilitate the development of new services and better provision or monetisation of old ones. Such data analysis may be expected to be a key area of innovation going forward (subject to constraints arising from data protection legislation). For example, Björkegren and Grissen (2020) show how modelling loan repayment

---

[47] In the EU, the use of such practices by the largest digital platforms are the target of new *ex ante* regulation through the EU *Digital Markets Act* 2022.

on the basis of mobile phone data can outperform credit bureau data[48], a finding which is likely to be relevant to the competitive position of credit cards.

Second, and partly as a result, we are seeing entry into the financial sector by pre-existing technology and e-commerce companies. These firms are well-positioned to leverage the data and user relationships they have from their primary business. They may also be able to leverage the data they gain from their financial services activity into their primary business, further increasing their incentives for entry. Zetzsche et al (2018) refer to such firms as 'techfin', rather than 'fintech', and highlight the challenges arising for regulation. Feyen et al (2021) describe the entry we have already seen in this space. De la Mano and Padilla (2018) emphasise that, despite such entry increasing competition and seeming positive in the short term, it risks leading to greater monopolisation over the longer term, as a few digital giants settle into newly entrenched positions.

In the context of payment systems, the biggest new innovations in developed countries are likely to result from digital wallets (such as Paypal), contactless mobile payments (such as Apple Pay and Android Pay). Currently these digital services act as additional intermediators, working on top of the existing payment card schemes. However, they do have the potential to disintermediate over time. A key question for regulators in this context will be how to ensure a level playing field and protect consumers, in the face of potentially very different business models.

---

[48] For example, an individual whose calls to others are returned may have stronger social connections that better allows him to follow through on entrepreneurial opportunities. Likewise, a responsible debtor is more likely to keep his phone topped up to a minimum threshold than one more prone to default.

4. Surcharging

As was discussed above, the extent and clarity of surcharging by merchants is critical for the unregulated setting of the MIF.

We saw that, with perfect surcharging of card fees to cardholders, the effect of the MIF may be entirely neutral, with any increase in the MIF, and thus higher surcharges, being fully offset by increased rewards to cardholders. By contrast, with zero surcharging, an increased MIF will tend to be passed on through higher general retail prices. This means that non-scheme members effectively subsidise any rewards to scheme members, and the MIF will tend to be set too high.

So, what is the situation in practice? In this section we look first at the question of whether merchants surcharge at all, and if not why not. We then consider incentives around the extent and clarity of such surcharging.

4.1 Do merchants surcharge, and if not why not?

The 'no surcharge rule'

In many jurisdictions, the two major payment card schemes impose a 'no surcharge rule'. Some smaller payment systems (notably, American Express) impose an even strong rule, an 'anti-steering rule' which not only prohibits surcharging but also prohibits merchants from doing anything else to steer consumers towards particular payment tools.

As well as the literature already described above, there is a distinct recent literature on this rule (Saxén, 2014; Edelman and Wright, 2015; Carlton and Winter, 2018; Liu et al, 2021; Adache and Tremblay, 2023). This has to some extent been precipitated by the recent US litigation related to the American Express anti-steering rule, which resulted in a highly controversial dismissal by the Supreme Court.[49]

The key message of much of this literature is aligned with the discussion above. The 'no surcharge rule' effectively acts to inflate fees to merchants (since any increase in fees will feed into general retail prices and thus only partially be borne by the scheme's card holders. This in turn leads to excessive take-up of the scheme in question and over-investment in benefits (such as rewards) to cardholders. Overall, it is bad for total surplus and potentially also for total welfare.

Carlton and Winter (2018) highlight the link to the economic literature on vertical Most Favoured Nation (MFN) clauses, which work in a similar way. Saxén (2014) highlights that no surcharge rules can have an additional negative effect, in enabling cards to be viable which confer no social surplus.

However, two papers provide a somewhat more caveated conclusion.

---

[49] *Ohio v. American Express,* 585 U.S. __ [2018].

- Liu et al (2021) argues that, with sufficiently convex demand, no surcharge rules can potentially have a positive 'drawing in' effect in attracting low-valuation cardholders, which can in turn benefit merchants. However, it not obvious in their model why a no surcharge rule would be required in this case; presumably merchants could anyway just decide not to surcharge.

- Bourgignon et al (2019) make the subtle point that surcharging is unlikely to be attractive to merchants if the MSC is lower than the 'tourist test' threshold, since they are then paying less to accept cards than the convenience benefit they receive from doing so. In this context, they argue that there is little point in regulating both the MIF to the 'tourist test' (or below) and also prohibiting the no surcharge rule. More controversially, within their model, they find that if the no surcharge rule is lifted, the imposition of a cap on the MIF will be detrimental to welfare. However, it should be noted that their model assumes a monopoly payment card.

Despite this almost unprecedented degree of coherence in the economic literature on surcharging, competition authorities seem to have found it surprisingly difficult to challenge no surcharging rules. The For example, as well as the American Express case already mentioned, the Canadian competition authorities challenged no surcharge rules, but unsuccessfully, in 2010.[50] At the state level in the US, many states not only allow no surcharge rules but insist on them, enforcing no surcharge rules as a matter of law.

In several jurisdictions, however, no surcharge rules have now been prohibited. In Australia, the rule has been prohibited since 2003. the EU, the 2015 Interchange Fee Regulation prohibits four-party credit cards from engaging in any form of steering. In the US (2010) and Canada (2017), Visa and MasterCard agreed to drop their no surcharge rules in order to settle class actions.

Surcharging is not prohibited by the major card schemes in New Zealand. Indeed, a recent survey commissioned by the NZ Commerce Commission[51] found that around 22% of merchant respondents surcharged for at least some transactions. For those that do, the surcharges typically lie between 2-3%, as is shown below. The Commerce Commission estimates that MSCs for small businesses are around 1.5%, which suggests that surcharges typically exceed these incremental costs of accepting cards. Surcharges do not normally vary across the different card schemes.[52]

---

[50] The US case is *Final Judgment as to Defendants Mastercard International Inc. and Visa Inc*., Civil Action No. CV-10-4496 (E.D.N.Y. Oct. 4, 2010). The Canadian case is *Comm'r of Competition v. Visa Canada Corp. & MasterCard Int'l Inc.,* 2013 Comp. Trib. 10.

[51] New Zealand Commerce Commission (2023) *Retail Payment System: Merchant Research Observations.*.

[52] Except for EFTPOS transactions which are rarely surcharged, reflecting the zero cost to merchants of these transactions

Figure 3.7   Level of surcharges enabled on terminals[31]
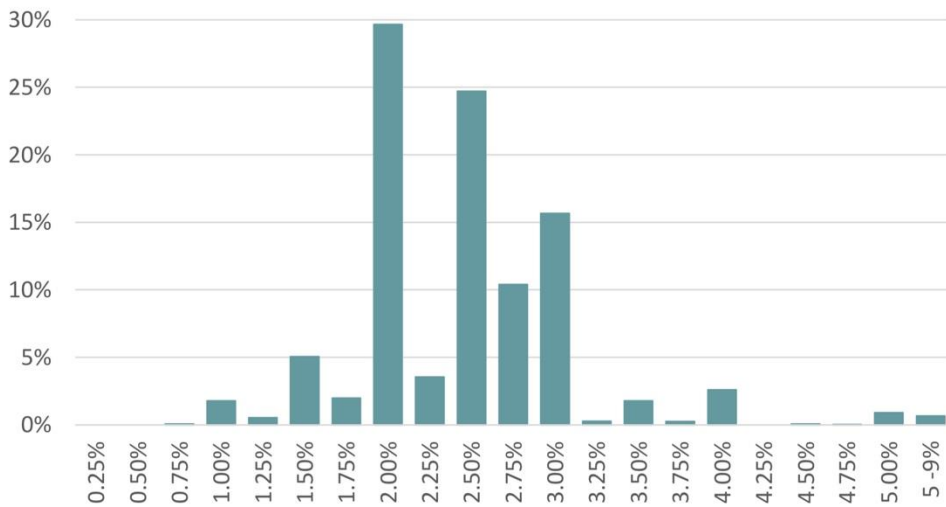
Merchant incentives to limit surcharging

Even where there is no explicit bar on surcharging, however, merchants may well decide not to make use of this freedom. For example:

- Shy and Stavins (2015) look at the impact of the US class settlement. This was implemented in 2012, but even by 2015, they find that most U.S. merchants were rarely taking advantage of their freedom to surcharge (although the prevalence was slightly higher for transactions above $20).

- Shy and Stavins (2015) also look at Australia. By 2010 – 7 years after the introduction of regulation – 20 percent of small merchants and 40 percent of large merchants had started to surcharge. These somewhat higher figures might suggest that expanding surcharging takes time, but also that there may be an inherent limit on the extent of surcharging by merchants.

- Bolt et al (2010) find that in the Netherlands (where there is no 'no surcharging' rule) around 1 in 5 retailers surcharged, but these were typically small retailers, dealing with small transactions.

Why do we only see limited surcharging? A number of reasons have been put forward.

- Shy and Stavins (2015) attribute this finding in part to merchants' fear of alienating consumers, who may not see such surcharging as acceptable. Similarly, Carlson and Weathers (2008) find that partitioned pricing (prices that are distinguished into two parts) can reduce consumer perceptions of trustworthiness for sellers they inherently trust less.

- Schuh et al (2012) evaluate the 2010 Visa and MasterCard settlement in the US. They conclude that merchants could not take advantage of their freedom to surcharge partly because of a lack of full information on credit card merchant fees.

- Merchants may also be concerned that the surcharges would primarily lead to consumers switching payment tool, rather generating additional revenues, thus denying the merchant the convenience benefits of using the tool. For example:

  o Bolt et al (2010) also find that consumers are highly reluctant to pay surcharges and the vast majority switch to cash. They estimate that the removal of surcharging in stores which use it would increase debit card usage, as a share of total payments, by 8 percentage points.

  o In a similar vein, Lam and Ossolinski (2015) found that 60% of consumers are unwilling to pay a surcharge of just 0.1%.

  o Stavins (2018) finds a smaller but still substantial effect. He estimates that using surcharges increases the probability of a cash transaction, by a consumer who otherwise prefers other payment methods, by 19.2%.

  o This may even be true for larger purchases. Schuh et al. (2012) describes an experiment conducted by the Swedish furniture company IKEA, where consumers in the United Kingdom were surcharged for credit card payments, while consumers in the United States were given a discount for debit card payments. Both experiments led to changes in payment behaviour, with a large shift away from credit cards in the UK, and a smaller shift toward debit cards in the US.[53]

The findings of this last research also highlight that consumers may react differently between 'surcharges for one payment tool and 'discounts for using other payment tools, even though these might seem to be formally the same. There are two possible rationales for this finding. The first relates to the information held by the consumer. If we assume that consumers first choose merchant on the basis of price, a cash discount acts as a special offer whereas a card surcharge acts as a form of 'hold up' (Bourgignon et al, 2019). The second is more behavioural. 'Prospect theory' suggests that consumers value gains (here, discounts) differently from losses (here, surcharges) (Kahneman and Tversky, 1979; Frankel, 2007).

Against the above, Wakamori and Welte (2017) suggest that consumers decisions about using cash for small value transactions are relatively unaffected by card acceptance or cost. Moreover, we arguably see more surcharging in practice than these arguments might suggest.

---

[53] This also highlights that consumers may react differently between 'surcharges for card payment' and 'discounts for cash payment', even though these are formally the same. One possible rationale could be informational. If we assume that consumers first choose merchant on the basis of price, a cash discount acts as a special offer whereas a card surcharge acts as a form of 'hold up' (Bourgignon et al, 2019). However, it may also be related to the findings of 'prospect theory' in behavioural economics, which would suggest that consumers value gains (discounts) differently from losses (surcharges) (Kahneman and Tversky, 1979).

4.2   The extent and clarity of surcharging

Much of the discussion above implicitly assumes that surcharges to consumers, if they exist, would be reflective of merchant service charges and would be sufficiently clear, prominent and timely that consumers act in a rational and informed way in relation to them.

However, this may not in practice be the case. First, there is evidence that merchants are not always clear about what their MSCs are. The NZ Commerce Commission's merchant survey[54] found that more than 50% of merchants surveyed did not know what rate they were charged, and – where they thought they did – the vast majority reported rates that were far higher than the publicly available rates or the Commerce Commission's estimate of true rates for SMEs. This lack of knowledge may reflect the fact that merchant services are typically bundled with other business bank services.

Second, even if they are aware of their own costs, merchants may have poor incentives to make surcharges clear and prominent for consumers, which in turn means they may have an incentive to raise them above cost.

If surcharges are shrouded to an extent that consumers are completely unaware of them, then this clearly won't affect their choice of payment tool. There will be limited constraint on their level, and they effectively act as a form of stealth tax. Moreover, the existence of 'attentive' consumers, who are aware of the surcharges and act to avoid them, need not protect the 'inattentive', who are not. Indeed, it could even worsen their situation.

In a seminal paper, Gabaix and Laibson (2006) show how competition can lead to such excessive surcharges from inattentive consumers being competed away as lower retail prices. Merchants do not profit overall from the excessive surcharges, but – acting unilaterally – no merchant can afford to stop levying the surcharge as it would then need to raise retail prices which would in turn make it appear uncompetitive. In this context, attentive consumers (who avoid the surcharges) are then effectively cross-subsidised by the inattentive (who pay them).

Similar results can occur if consumers only become aware of the surcharges late in the purchasing process. This is especially pernicious in an online environment, where there may be no realistic alternative means of payment, such that surcharges are (in effect) compulsory.

The economic literature on drip pricing is relevant here. When the total price payable is 'dripped' over time, behavioural consumers may exhibit endowment effects (whereby they feel committed to their purchase) and loss aversion (in relation to changing their mind). In this context, there is likely to be little competitive constraint on their level (Rasch et al, 2020).

For example, research has consistently shown that when an offer is presented using drip pricing, consumers are more likely to make a purchase and less likely to shop around and compare prices (Ahmetoglu et al., 2014). Consumers encountering drip pricing have been shown to focus predominantly on the lower headline price and underestimate the total price of the offer (Greenleaf et al., 2016). Even when hidden charges are revealed after a few clicks,

---

[54] See footnote 51.

consumers often give insufficient weight to them compared to the more salient headline price (Huck and Wallace, 2015).

In this context, we might expect to see excessive surcharging for payments by merchants, even in competitive markets, at the same time as too little consumer reaction to them. This provides a justification for the requirement within the EU's revised Payment Systems Directive[55] (PSD2) that surcharges may not be any higher than the cost incurred by the merchant in accepting the relevant payment method.

Bourgignon et al (2019) make a similar argument for regulation of payment surcharges, in their context because consumers choose their merchant without knowing that merchant's card acceptance/surcharging policy. Notably, though, they propose a more stringent regulation: a cap that is equal to the MSC minus the convenience benefit that merchants receive from accepting the payment tool.

In a context where the MSC is regulated according to the tourist test threshold, this equates to a requirement of zero surcharging. In fact, this is also consistent with PSD2, which additionally requires that there should be no surcharges levied for any payment tools covered by the 2015 Interchange Fee Regulation. Similar standards can be issued in New Zealand under the NZ Retail Payment System Act 2022.

4.3   Conclusion on surcharging

Overall, there is little evidence that surcharging works in a sufficiently efficient way to either provide an effective constraint on the MIF, or to drive a 'neutrality result' whereby the level of the MIF was not a problem.

Key barriers to efficient surcharging include merchant concerns about consumer reactions to surcharging, a lack of knowledge amongst merchants about the true merchant service charges they face, and the ability and incentive of merchants to shroud surcharges so that they have limited effect on consumer choice of payment type.

---

[55] Directive (EU) 2015/2366.

5. Implications for regulating the MIF

The previous sections have set out the rationale for regulating the MIF. But what level should be set? A complete discussion of this issue would go beyond the academic literature and is outside the scope of this review, but this section includes a few remarks based on the available literature.

5.1 The 'tourist test'

The leading proposed test is the '*tourist test*', which is effectively the 'merchant indifferent test' developed by Baxter (1983) applied to the situation where all consumers are assumed to be tourists (to abstract from any benefits of attracting repeat sales). Under this approach, the MIF would be set equal to the *merchants' average convenience benefit from using a given payment method minus the merchant acquirers' average costs*.

As is shown in Rochet and Tirole (2011), the 'tourist test' is targeted at maximising total user surplus, not total welfare. Maximising total welfare would also takes account of the profits of scheme participants, which would typically imply a higher regulated MIF. Vickers (2005) argues against the latter approach. He argues that to set a regulated MIF so as to maximise total welfare would be equivalent to subsidising a monopolist in order to reduce its price to cost – 'an unusual form of government intervention'. He thus supports the 'tourist test' approach.

Even if seeking to maximise total user surplus, Rochet and Tirole recognise that, in principle, a regulator could go further in seeking to address the various cross-market externalities and imbalances of market power. For example, within their model, the MIF that maximises total user surplus may be:

- *lower* than the average tourist test if issuer margins are variable and decreasing in the cost of issuing (cost absorption); or

- *higher* than the average tourist test if one considers the potential for long run entry, or if issuer margins are variable and increasing in the cost of issuing (cost amplification).

However, they advise against such targeted price regulation, not least because of the large informational requirements it presupposes.

Applying the 'tourist test' to New Zealand

To estimate the convenience benefits of card usage, it is first important to identify the relative alternative. In many jurisdictions, the appropriate counterfactual to card use has traditionally been cash, which is relatively expensive to administer, and the 'tourist test' threshold has typically been considered in that context.[56]

---

[56] For example, the EU's econometric research (based on a survey of merchants) estimated that the long term benefits to merchants were 0.19% to 0.46% for debit cards and 0.19% to 0.47% for credit cards. This informed their eventual choice of cap for the MIF of 0.2% for debit cards and 0.3% for credit cards. By contrast, for the US, Layne-Farrar (2011) investigates the relative merchant costs of debit cards and cash

However, in New Zealand, this may not be the case. The EFTPOS card scheme might seem the more obvious counterfactual, and the cost of this to merchants is zero. This doesn't necessarily mean that New Zealand merchants gain no benefit from using cards, though. The key benefits to merchants of the major card schemes, over the EFTPOS scheme, are (i) the ability to take contactless payments and (ii) the ability to sell online. This raises the obvious question of how valuable these are.

In relation to the benefit from being able to accept contactless payments, the evidence is mixed.

- Bounie and Camara (2020) analyse the French market. They find that accepting contactless payments increases the card sales by 15.3 percent on average compared to merchants who do not accept contactless payments. they also find evidence that accepting contactless payments exerts a positive spillover of about 1.3 percent in the amount of contact card sales, and thus significantly increases the average annual card-sales amount and count for small merchants and new entrepreneurs.

- By contrast Brown et al (2020) examine the Swiss market, exploiting the staggered introduction of contactless debit cards. There the key alternative payment mechanism is cash. They find the impact of contactless cards on cash demand is to be economically small and statistically insignificant. Trütsch (2020) finds the same result for the US market.

The benefit from being able accept cards online is also hard to value and is likely to vary significantly across retailers (and in particular whether they also have a strong offline presence). EFTPOS is no longer the obvious counterfactual in this environment, with American Express or Paypal perhaps being better alternatives. Overall, this could potentially imply that the tourist test would generate different MIFs in an online and offline environment.

The risk of setting differential MIFs between debit and credit cards

In theory, there may be a differential between debit and credit cards in relation to (i) the benefits received by merchants and (ii) the costs incurred by acquirers. As an example of the former, where consumer spending is based on access to credit, credit card acceptance will be more critical for avoiding 'missed sales' than debit card acceptance.

However, Rochet and Wright (2010) highlight that it may be inefficient for consumers to use credit cards for 'ordinary purchases' if credit cards generally are higher cost than debit cards. This could militate against higher interchange fees being charged for credit cards, since these will be passed back to consumers as rewards and incentivise credit card use. Likewise, Mukharlyamov and Sarin (2022) likewise show that interchange fee caps which are higher for debit than credit cards may lead to end users being encouraged to take up the latter, reducing the overall effect of regulation on average interchange fees.

---

across a range of different retail situations, prior to regulatory intervention under the Durbin Amendment. They find that, if anything, debit cards are cheaper to accept than cash, and therefore conclude that there was no economic basis for regulating debit fees.

Thus, a differential approach risks causes cardholder substitution which increases overall costs and undermining the objectives of the regulation. This potential for substitution across card types following regulation is supported by evidence in Morris et al (2022), which finds that:

- in Australia, the fact that the MIF regulation only applied to credit cards has led consumers to switch from using credit cards to using debit cards, with debit transactions increasing at a much faster rate than the volume of credit card transactions.

- Conversely, because the 'Durbin Amendment' of the 2010 US Dodd-Frank Act did not apply to credit cards, this led consumers to shift towards using credit cards for transactional purposes (that is, revolving the credit) in place of debit cards.

The risk of setting differential MIFs between three-party and four-party cards

Since there is no MIF charged within a three-party card scheme, it is not possible to regulate these in a similar way. An alternative would be to regulate merchants fees directly, but this would raise a number of further issues (since there tends to be substantial variation in MSCs).

However, regulating four-party cards alone does create risks, both in terms of regulatory arbitrage and consumer substitution. As Tirole (2011, p.149) notes "*Whatever regulation (or lack thereof) one advocates, neutrality with respect to business organization should be the rule, so as to let the most efficient organizational forms emerge.*" He notes that, in the wake of the mandated decrease in the Australian interchange fee, three of the top four Australian banks signed up agreements to issue American Express or Diners Club cards, as a package alongside their usual four-party card. These "companion cards" enabled consumers to use the higher-earning three-party card where it was accepted.

Morris et al (2022) examine the impact of these 'companion cards' and show a clear substitution effect. By volume of transactions, three-party cards increased from about 10% in 2002 to about 16% in 2013 (a 60% increase). The RBA changed its rules in 2017 to make American Express companion cards subject to the same interchange-fee caps as Visa and Mastercard. Companion cards were quickly discontinued and the market share by volume of three-party cards fell back to 7%.

5.2   Alternative regulatory approaches

While the 'tourist test threshold' arguably provides the upper bound on an appropriate cap, the lower bound must be zero, given the potential for this to be the counterfactual to a collectively set MIF. Some have indeed argued for a zero MIF. However, as mentioned above, a zero fee does risk being too low to incentivise investment and innovation.

An alternative approach to setting a MIF cap would be to base it on some measure of costs. For example, existing regulators (such as the RBA in Australia and the Federal Reserve Board in the U.S.) have used narrowly defined issuers' variable costs to determine the cap on allowed interchange fees. The problem with this approach is that it is not supported by any economic theory.

An interesting recent approach to considering the optimal interchange fee has been proposed by Huynh et al (2022). They construct and estimate a structural two-stage model of the Canadian payments market to quantify the network externalities and identify the main determinants of consumer and merchant decisions. An important distinction between their approach and the 'tourist test' concept is that they explicitly allow for both 'first time' customers (who have limited knowledge about merchants' card acceptance polices) and repeat customers who do. (Recall that the 'tourist test' is so called because it assumes no repeat customers).

Using their model, they can simulate the optimal interchange fee. Assuming that issuers set their margins optimally, they find that this would equate to a MIF of 0.2%. This is a substantial reduction from the actual Canadian MIF in 2014, which was 1.8%. This approach builds on earlier work by Bedre-Defolie et al (2018) who carry out a similar exercise in the context of the Norwegian debit card scheme.

References

Adachi, T. and M. Tremblay (2023). Do No-Surcharge Rules Increase Effective Retail Prices? *Working Paper*, SSRN.

Agarwal, S. and J. Zhang (2015). A Review of Credit Card Literature: Perspectives from Consumers. *Report for the UK Financial Conduct Authority*.

Agarwal, S., A. Presbitero, A. F. Silva and C. Wix (2023). Who Pays for Your Rewards? Redistribution of the Credit Card Market. *IMF Working Paper No*. 2023/054.

Ahmetoglu, G., A. Furnham and P. Fagan (2014). Pricing Practices: A Critical Review of Their Effects on Consumer Perceptions and Behaviour. *Journal of Retailing and Consumer Services*, 21(5), 696–707.

Ambrus, A. and R. Argenziano (2009). Asymmetric Networks in Two-Sided Markets. *American Economic Journal: Microeconomics*, 1(1), 17-52.

Amelio, A. and B. Jullien (2012). Tying and freebies in two-sided markets. *International Journal of Industrial Organization*, 30(5), 436-446.

Arango, C. and V. Taylor (2008). Merchant Acceptance, Costs, And Perceptions Of Retail Payments: A Canadian Survey. *Bank of Canada Discussion Paper,* 2008-12.

Arango, C., K. P. Huynh and L. Sabetti (2015). Consumer Payment Choice: Merchant Card Acceptance Versus Pricing Incentives. *Journal of Banking & Finance*, 55, 130-141.

Ardizzi, G., D. Scalise, and G. Sene (2021). Interchange Fees and Card Payments: A Cross-country Analysis. *Bank of Italy Occasional Paper No. 628*.

Ardizzi, G., and M. Zangrandi (2018). The Impact of the Interchange Fee Regulation on Merchants: Evidence from Italy. *Bank of Italy Occasional Paper No. 434.*

Ardizzi, G. (2013). Card versus Cash: Empirical Investigation of the Impact of Payment Card Interchange Fees on End Users' Choice of Payment Methods. *Journal of Financial Market Infrastructure*, 1(4), 73-105.

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37: 668-691.

Armstrong, M. and J. Wright (2007). Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts. *Economic Theory,* 32, 353-380.

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

Bedre-Defolie, Ö. and E. Calvano (2013). Pricing payment cards, *American Economic Journal-Microeconomics*, 5, 206-231.

Bedre-Defolie, Ö., M. Song and H. Ullrich (2018). Assessing the Impact of Payment Card Fee Regulation. *Working Paper*.

Belleflamme, P. and M. Peitz (2016). Platforms and Network Effects. *Handbook of Game Theory and Industrial Organization*. Edited by L. Corchon and M. Marini, Edward Elgar.

Belleflamme, P. and M. Peitz (2019). Platform competition: Who benefits from multihoming? *International Journal of Industrial Organization*, 64, 1-26.

Belleflamme, P. and E. Toulemonde (2018). Tax Incidence on Competing Two-Sided Platforms. *Journal of Public Economic Theory*, 20, 9–21.

Biglaiser, G., J. Crémer and A. Veiga (2022). Should I stay or should I go? Migrating away from an incumbent platform. *CEPR Press Discussion Paper,* No. 14496.

Björkegren , D. and D. Grissen (2020). Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment. *The World Bank Economic Review*, 34(3), 618–634.

Bolt, W., N. Jonker, and C. van Renselaar (2010). Incentives at the Counter: An Empirical Analysis Of Surcharging Card Payments and Payment Behaviour in the Netherlands. *Journal of Banking & Finance*, 34(8), 1738-1744.

Bourguignon, H. R. Gomes and J. Tirole (2019). Shrouded Transaction Costs: Must-Take Cards, Discounts and Surcharges. *International Journal of Industrial Organization*, 63, 99-144.

Bounie, D. and A. François (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.

Bounie, D., A. François, and L. Van Hove (2017). Consumer Payment Preferences, Network Externalities, and Merchant Card Acceptance: An Empirical Investigation. *Review of Industrial Organization*, *51*(3), 257–290.

Bounie, D. and Y. Camara (2020). Card-Sales Response to Merchant Contactless Payment Acceptance. *Journal of Banking & Finance*, 119.

Bourreau, M. and M. Verdier (2019). Interchange Fees and Innovation in Payment Systems. *Review of Industrial Organization*, *54*(1), 129–158.

Brannon, I. and C. Richardson (2022). The Impact of an Interchange Fee Cap on Credit Card Transactions. *Working Paper*, SSRN.

Briglevics, T., and O. Shy (2014). Why Don't Most Merchants Use Price Discounts to Steer Consumer Payment Choice? *Review of Industrial Organization*, 44(4), 367–392.

Brown, M., N. Hentschel, H. Mettler and H. Stix (2020). Financial Innovation, Payment Choice and Cash Demand-Causal Evidence from the Staggered Introduction of Contactless Debit Cards. University of St. Gallen, School of Finance Research Paper No. 2020/02.

Cabral, L. (2011). Dynamic Price Competition with Network Effects. *The Review of Economic Studies*, 78(1), 83–111.

Caillaud, B. and B. Jullien (2003). Chicken & egg: Competition among intermediation service providers. *The RAND Journal of Economics,* 34, 309–328.

Carbó-Valverde, S., J. M. Liñares-Zegarra, F. Rodríguez-Fernández (2012). Feedback loop effects in payment card markets: Empirical evidence. *Review of Network Economics,* 11(2), 1- 24.

Carbó-Valverde, S., S. Chakravorti and F. R Fernández (2016). The Role of Interchange Fees in Two-Sided Markets: An Empirical Investigation on Payment Cards. *The Review of Economics and Statistics*, 98(2), 367-381.

Carbó-Valverde, S. and J. M. Liñares-Zegarra (2011). How Effective Are Rewards Programs In Promoting Payment Card Usage? Empirical Evidence. Journal of Banking & Finance,
35(12), 3275-3291.

Carlson, J.P. and D. Weathers (2008). Examining Differences in Consumer Reactions to Partitioned Prices with a Variable Number of Price Components. *Journal of Business Research*, 61(7), 2008, 724-731.

Carlton, D. W. and R. A, Winter (2018). Vertical Most-Favored-Nation Restraints and Credit Card No-Surcharge Rules. *Journal of Law and Economics*, 61(2), 215–51.

Cohen, M. A. and M. Rysman (2013). Payment Choice with Consumer Panel Data. *FRB of Boston Working Paper* No. 13-6.

Crémer, J., P. Rey and J. Tirole (2000). Connectivity in the Commercial Internet. *IDEI Working Paper*, No. 87, Institut d'Économie Industrielle (IDEI), Toulouse.

De Groen, W. P. (2020) The Impact of EU Price Rules: Interchange Fee Regulation in Retail Payments. *CEPS-ECRI Working Paper.*

De la Mano, M. and J. Padilla (2018). Big Tech Banking, *Journal of Competition Law & Economics*, 14(4), 494–526.

Di Iorio, A. & G. Rocco (2022). Easier Said Than Done: Why Italians Pay In Cash While Preferring Cashless. *Bank of Italy Occasional Paper No. 731*.

Doganoglu, T. and J. Wright (2006). Multihoming and Compatibility. *International Journal of Industrial Organization*, 24(1), 45-67.

Edelman, B. G. and J. Wright (2015). Price Coherence and Excessive Intermediation *Quarterly Journal of Economics*, 130(3), 1283-1328.

Eisenmann, T., G. Parker and M. Van Alstyne (2011). Platform Envelopment. *Strategic Management Journal*, 32(12), 1270-128.

Evans, D. S., Chang, H. and S. Joyce (2015). The Impact of the U.S. Debit-Card Interchange Fee Regulation On Consumer Welfare. *Journal of Competition Law and Economics*, 11(1), 23-67.

Farrell, J. (2006). Efficiency and competition between payment instruments. *Review of Network Economics*, 5, 26-44.

Felt, M.-H., F. Hayashi, J. Stavins and A. Welte (2020). Distributional Effects of Payment Card Pricing and Merchant Cost Pass-through in the United States and Canada. *Federal Reserve Bank of Kansas City Working Paper* No. 20-18.

Feyen, E., J. Frost, L. Gambacorta, H. Natarajan and M. Saal (2021). Fintech and the Digital Transformation of Financial Services: Implications for Market Structure and Public Policy. *Bank for International Settlements Paper*, No. 117.

Filistrucchi, L., D. Geradin, E. van Damme, and P. Affeldt (2014). Market Definition in Two-Sided Markets: Theory and Practice, *Journal of Competition Law & Economics*, 10(2), 293–339.

Fletcher, A. (2020). *Digital Competition Policy: Are Ecosystems Different?* Note for OECD Hearing on the Competition Economics of Digital Ecosystems. DAF/VCOMP/WD(2020)96.

Franck, J.-U. and M. Peitz (2019). *Market Definition and Market Power in the Platform Economy*. Report for the Centre on Regulation in Europe (CERRE).

Frankel, A. S. (2007). Towards A Competitive Card Payments Marketplace. *Submission to the Reserve Bank of Australia*.

Gabaix , X. and D. Laibson (2006). Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *The Quarterly Journal of Economics*, 121(2), 505–540.

Greenleaf, E. A., E. J. Johnson, V. G. Morwitz and E. Shalev (2016).The Price Does Not Include Additional Taxes, Fees, And Surcharges: A Review Of Research On Partitioned Pricing, *Journal of Consumer Psychology*, 26(1), 105-124.

Guthrie, G. and J. Wright (2007). Competing Payment Schemes. *Journal of Industrial Economics*, LV, 37–67.

Hagiu, A. and H. Hałaburda (2014) Information and two-sided platform profits. *International Journal of Industrial Organization*, 34, 25-35.

Halaburda, H., B. Jullien, and Y. Yehezkel (2020). Dynamic competition with network externalities: how history matters. *The RAND Journal of Economics*, 51(1), 3–31.

Heidhues, P. and B. Kőszegi (2010). Exploiting Naïvete about Self-Control in the Credit Market. *The American Economic Review*, *100*(5), 2279–2303.

Hovenkamp, E. (2019). Platform Antitrust. *Journal of Corporation Law*, 44, 713.

Huck, S. and B. Wallace (2015). The Impact of Price Frames on Consumer Decision Making: Experimental Evidence. In *Experimental evidence* (pp. 1–47).

Hurkens , S. and Á. L. López (2014). Mobile Termination, Network Externalities and Consumer Expectations. *The Economic Journal*, 124(579), 1005–1039.

Huynh, K. P., G. Nicholls and O. Shcherbakov (2022). Equilibrium in Two-Sided Markets for Payments: Consumer Awareness and the Welfare Cost of the Interchange Fee. *Bank of Canada Discussion Paper,* 2022-15.

Iranzo, J., P. Fernández, G. Matías and M. Delgado (2012). The Effects of the Mandatory Decrease of Interchange Fees in Spain, *MPRA Paper* No. 43097.

Jullien, B. and W. Sand-Zantman (2021). The Economics of Platforms: A Theory Guide for Competition Policy. *Information Economics and Policy*, 54.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291.

Karle, H., M. Peitz and M. Reisinger (2020). Segmentation versus Agglomeration: Competition between Platforms with Competitive Sellers. *Journal of Political Economy*, 128(6), 2329-2374.

Katz, M. L. (2019). Platform Economics and Antitrust Enforcement: A Little Knowledge Is A Dangerous Thing. *Journal of Economics and Management Strategy,* 28, 138–152.

Kay, B.S., M. D. Manuszak and C. M. Vojtech (2018). Competition And Complementarities In Retail Banking: Evidence From Debit Card Interchange Regulation. *Journal of Financial Intermediation*, 34, 91-108.

Klee, E. (2008). How People Pay: Evidence from Grocery Store Data. *Journal of Monetary Economics,* 55(3), 526–541.

Klein, B., A. V. Lerner, K. M. Murphy, and L. L. Plache (2006). Competition In Two-Sided Markets: The Antitrust Economics Of Payment Card Interchange Fees. *Antitrust Law Journal*, 73(3), 571–626.

Layne-Farrar, A. (2011). Are Debit Cards Really More Costly For Merchants? Assessing Retailers' Costs And Benefits Of Payment Instrument Acceptance. *Mimeo*, SSRN.

Lee, R. S. (2013). Vertical Integration and Exclusivity in Platform and Two-Sided Markets. American Economic Review, 103 (7), 2960-3000.

Liu, C., F. Niu and A. White (2021). Optional Intermediaries and Pricing Restraints. *Working Paper*, SSRN.

Manuszak, M. D. and K. Wozniak (2017). The Impact of Price Controls in Two-Sided Markets: Evidence from U.S. Debit Card Interchange Fee Regulation. *Federal Reserve Board Working Paper 2017-074*.

Morris, J., T. J. Zywicki and G. Manne (2022). The Effects of Price Controls on Payment-Card Interchange Fees: A Review and Update. *ICLE White Paper 2022*-03-04, George Mason Law & Economics Research Paper No. 22-07.

Mukharlyamov, V. and N. Sarin (2022). Price Regulation in Two-Sided Markets: Empirical Evidence from Debit Cards. *Working Paper*. SSRN

Niels, G. (2019). Transaction Versus Non-Transaction Platforms: A False Dichotomy In Two-Sided Market Definition. *Journal of Competition Law & Economics*, 15(2-3), 327–357.

Rasch, A., M. Thöne and T. Wenzel (2020). Drip Pricing and its Regulation: Experimental Evidence. *Journal of Economic Behavior & Organization*, 176, 353-370.

Reisinger, M. and H. Zenger (2019). Interchange Fee Regulation and Service Investments. *International Journal of Industrial Organization*, 66, 40-77.

Rochet, J.-C. and J. Tirole (2002). Cooperation Among Competitors: Some Economics of Payment Card Associations. *The RAND Journal of Economics*, 33, 549-570.

Rochet, J.-C. and J. Tirole (2003). Platform Competition in Two-Sided Markets. *Journal of the* European Economic Association, Vol. 1, 990–1029.

Rochet, J.-C. and J. Tirole (2011). Must-Take Cards: Merchant Discounts and Avoided Costs. *Journal of the European Economic Association*, 9, 462-495.

Rysman, M. (2007). An empirical analysis of payment card usage. *Journal of Industrial Economics*, 55, 1-36.

Rysman, M. and J. Wright (2014). The Economics of Payment Cards. *Review of Network Economics*, 13(3), 303-353.

Saxén, F. (2014). The No Surcharge Rule and Merchant Competition. *Journal of Industry, Competition and Trade*, 14, 39–66.

Schuh, S., O. Shy, and J. Stavins (2010). Who Gains And Who Loses From Credit Card Payments? Theory And Calibrations. *Public Policy Discussion Paper* No. 10-3, Federal Reserve Bank of Boston.

Schuh, S., O. Shy, J. Stavins and R. Triest (2012). An Economic Analysis of the 2011 Settlement Between the Department of Justice and Credit Card Networks. *Journal of Competition Law and Economics*, 8(1), 107–144.

Shabgard, B and J. Asensio (2023). The Price Effects of Reducing Payment Card Interchange Fees. *Journal of the Spanish Economic Association*, 14, 189–221.

Shy, O. (2021). Cashless Stores and Cash Users. *Journal of Policy Modeling*, 43(3), 622-638.

Shy, O. and J. Stavins (2015). Merchant Steering of Consumer Payment Choice: Evidence from a 2012 Diary Survey, *Journal of Behavioral and Experimental Economics*, 55, 1-9.

Simon, J., K. Smith and T. West (2010). Price Incentives and Consumer Payment Behaviour. *Journal of Banking & Finance*, 34(8), 1759-1772.

Small, J. and J. Wright (2002). The Bilateral Negotiation of Interchange Fees in Payment Schemes. *Working paper*, National University of Singapore.

Stavins, J. (2018). Consumer Preferences for Payment Methods: Role of Discounts and Surcharges, *Journal of Banking & Finance*, 94, 35-53.

Trütsch, T. (2020). The Impact Of Contactless Payment On Cash Usage At An Early Stage Of Diffusion. *Swiss Journal of Economics and Statistics* **156**, 5 (2020).

Vickers, J. (2005). Public Policy and the Invisible Price: Competition Law, Regulation and the Interchange Fee. *Competition Law Journal*, 4, 5-16.

Wakamori, N. and A. Welte (2017). Why Do Shoppers Use Cash? Evidence from Shopping Diary Data, *Journal of Money, Credit and Banking*, 49(1), 115–169.

Wright, J. (2004). The Determinants of Optimal Interchange Fees in Payment Systems. *The Journal of Industrial Economics*, *52*(1), 1–26.

Wright, J. (2012). Why payment card fees are biased against retailers. *The RAND Journal of Economics*, 43, 761-780.

Zetzsche, D. A., R. P. Buckley, D. W. Arner and J. N. Barberis (2017). From FinTech to TechFin: The Regulatory Challenges of Data-Driven Finance. *European Banking Institute Working Paper* 2017- 6.