Debit Card Interchange Fee Regulation: Some Assessments and Considerations

Zhu Wang

I n the summer of 2011, the Federal Reserve Board of Governors issued a final rule governing debit card interchange fees. This regulation, named Regulation II (Debit Card Interchange Fees and Routing), was required by the Durbin Amendment to the Dodd-Frank Act. The regulation, which went into effect on October 1, 2011, limits the maximum permissible interchange fee that a covered issuer can collect from merchants for a debit card transaction.

The Durbin Amendment and the resulting regulation were created to resolve the long-time conflicts between card issuers and merchants regarding payment card interchange fees. The interchange fee is the amount that a merchant has to pay the cardholder's bank (the so-called issuer) through the merchant acquiring bank (the so-called acquirer) when a card payment is processed. Merchants have criticized that card networks (such as Visa and MasterCard) and their issuing banks have used market power to set excessively high interchange fees, which drive up merchants' costs of accepting card payments. Card networks and issuers disagree, countering that interchange fees have been properly set to serve the needs of all parties in the card system, including funding better consumer reward programs that could also benefit merchants.

By capping debit card interchange fees, the regulation has generated significant impact on the U.S. payments industry since its implementation. The most visible impact is the drop of multibillion-dollar

annual revenues for card issuers in terms of the interchange fees that they collect from merchants. Meanwhile, the regulation has yielded other intended and unintended consequences. In this article, we review the regulation's impact from both positive and normative perspectives. We first look into the empirical evidence of the regulation's first-year effects on different players in the debit card market, namely issuers, merchants, and consumers. We then provide a simple two-sided market model, based on the work of Rochet and Tirole (2011), to assess the regulation's implications on payments efficiency. The model sheds light on important policy questions, for example, whether the debit card market performs inefficiently without regulation and whether the Durbin regulation can improve market outcome. Finally, we extend the model to explain the regulation's unintended consequence on small-ticket merchants and discuss an alternative regulatory approach.

The article is organized as follows. Section 1 provides the background of payment card markets and the interchange fee regulation. Section 2 reviews the empirical evidence on the regulation's impact on different players in the debit card market. Section 3 lays out a simple model of the payment card market and discusses the regulation's implication on payments efficiency. We then extend the model to address the regulation's unintended consequence on small-ticket merchants. Finally, Section 4 provides concluding remarks.
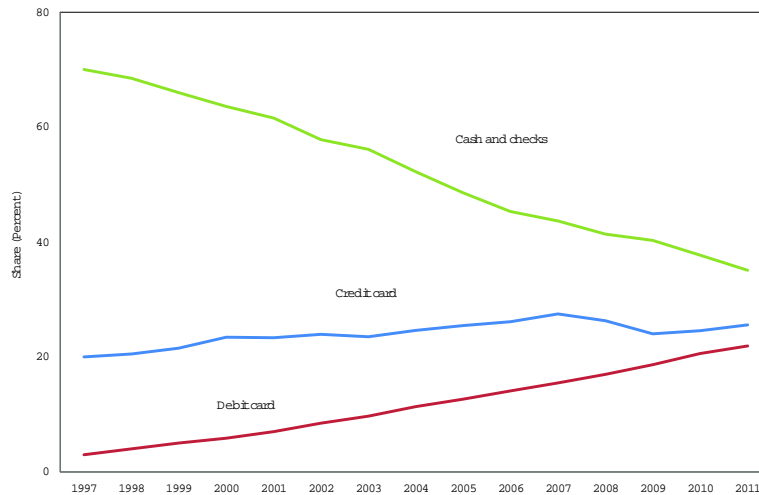
1. INDUSTRY BACKGROUND

As payments migrate from paper to electronic forms, credit and debit cards have become an increasingly important part of the U.S. payments system. Recent data show that the payment share of credit and debit cards in personal consumption expenditures rose from 23 percent in 1997 to 48 percent in 2011, while the share of cash and checks dropped from 70 percent to 35 percent (Figure 1).[1] In 2011, debit cards were used in 49 billion transactions for a total value of $1.8 trillion, and credit cards were used in 26 billion transactions for a total value of $2.1 trillion.

Along with this development has come controversy. Merchants are critical of the fees that they pay to accept cards. These fees are often referred to as the "merchant discounts," which are composed mainly of interchange fees paid by merchants to card issuing banks through merchant acquiring banks. Merchants believe that the card networks

---

[1] The data are drawn from various issues of the *Nilson Report*. Payment shares not shown in Figure 1 include the automated clearing house and some other miscellaneous types.

Figure 1 Payment Shares of U.S. Personal Consumption Expenditures

and issuing banks have wielded their market power to set excessively high interchange fees. The card networks and issuers counter that these interchange fees are necessary for covering issuers' costs as well as providing rewards to cardholders, which may also benefit merchants by making consumers more willing to use the cards.

Market Overview

To understand the interchange fee controversy, some familiarity with the payment card markets is helpful. Credit and debit cards are two of the most popular general-purpose payment cards in the United States.[2] Credit cards typically provide credit or float to cardholders, while debit cards directly draw from the cardholder's bank account right after each transaction. Debit card payments are authorized either by the cardholder's signature or by a personal identification number (PIN). The

---

[2] Pre-paid cards are another type of general-purpose card, but their market size is much smaller compared with credit and debit cards. In 2011, the transaction value of pre-paid cards accounted for 2 percent of U.S. personal consumption expenditures (Data source: *Nilson Report*).

former is called signature debit and the latter is called PIN debit. In terms of transaction volume, signature debit accounts for 60 percent of debit transactions, while PIN debit accounts for 40 percent.

Visa and MasterCard are the two major credit card networks in the United States. They provide card services through member financial institutions and account for 85 percent of the U.S. consumer credit card market.[3] Visa and MasterCard are also the primary providers of debit card services. The two networks split the signature debit market, with Visa holding 75 percent of the market share and MasterCard holding 25 percent.[4] In contrast, PIN debit transactions are routed over the PIN debit networks. Currently, there are 14 PIN debit networks in the United States. Interlink, Star, Pulse, and NYCE are the top four networks, together holding 90 percent of the PIN debit market. The largest PIN network, Interlink, is operated by Visa.

Visa, MasterCard, and PIN debit networks are commonly referred to as four-party schemes because four parties are involved in each transaction in addition to the network whose brand appears on the card. These parties include: (1) the cardholder who makes the purchase; (2) the merchant who makes the sale and accepts the card payment; (3) the financial institution that issues the card and makes the payment on behalf of the cardholder (the so-called issuer); and (4) the financial institution that collects the payment on behalf of the merchant (the so-called acquirer).

In a four-party card scheme, interchange fees are collectively set by the card network on behalf of their member issuers. For a simple example of how interchange functions, imagine a consumer making a $50 purchase with a payment card. For that $50 item, the merchant would get approximately $49. The remaining $1, known as the merchant discount, gets divided up. About $0.80 would go to the card issuing bank as the interchange fee, and $0.20 would go to the merchant acquiring bank (the retailer's account provider). Interchange fees serve as a key element of the four-party scheme business model and generate significant revenues for card issuers. In 2009, U.S. card issuers made approximately $48 billion revenue in interchange fees, with debit interchange revenues being $17 billion and credit interchange revenues being $31 billion.[5]

---

[3] American Express and Discover are the other two credit card networks holding the remaining market shares. They handle most card issuing and merchant acquiring by themselves, and are called "three-party" systems. For a "three-party" system, interchange fees are internal transfers.

[4] Discover has recently entered the signature debit market, but its market share is small.
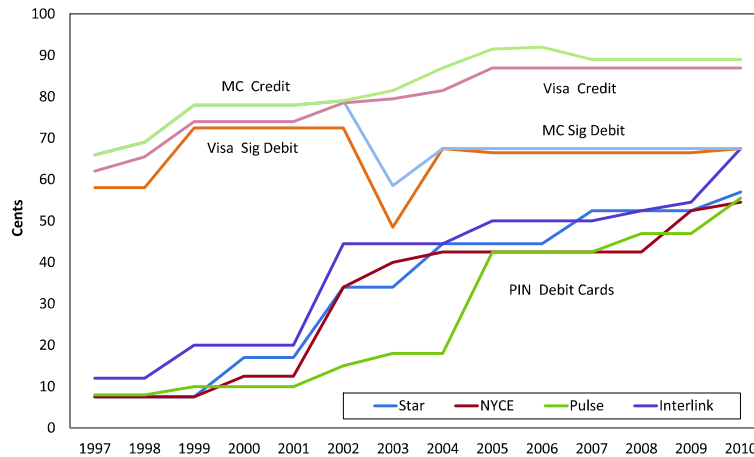
[5] See Levitin (2010).

Figure 2 Interchange Fees for a $50 Transaction

Figure 2 plots the interchange fee for a $50 non-supermarket transaction for Visa and MasterCard credit cards, signature debit cards, as well as the top four PIN debit cards in the United States.[6] As the figure shows, credit and PIN debit interchange fees have been rising since the late 1990s, while signature debit interchange fees came down in 2003 before rising again soon after.[7] Over the years, the gap of interchange fees between PIN debit and signature debit has also narrowed substantially.

Interchange Battles

Merchants criticize the interchange fees for being excessively high. They point out that the high and rising interchange fees deviate from cost basis and are in sharp contrast to the falling card processing and fraud costs during the same period.[8] In recent years, merchant groups

---

[6] Data source: *American Banker* (various issues).

[7] The temporary drop of signature debit interchange fees was due to the settlement of the Wal-Mart case, which allowed merchants who accept Visa or MasterCard credit cards to not have to accept their signature debit cards.

[8] Payment cards is primarily an information-processing industry. As the information technology progresses, the relative prices of computers, communications, and

launched a series of litigation against what they claim is anticompetitive behavior by the card networks and their issuers. Some of the lawsuits have been aimed directly at interchange fees, including both credit and debit cards. For example, a group of class-action suits filed by merchants against Visa and MasterCard in 2005 alleged that the networks violated antitrust laws by engaging in price fixing. As a result, Visa and MasterCard recently agreed to a $7.25 billion settlement with U.S. retailers, which could be the largest antitrust settlement in U.S. history.[9] Other merchant lawsuits have focused not on interchange fees *per se*, but on alleged anticompetitive practices. A prime example is the lawsuit filed by Wal-Mart and other merchants in 1997 against the networks' honor-all-cards rule, which required a merchant accepting a network's credit cards to also accept its signature debit cards. The Wal-Mart case was settled in 2003. As a result, Visa and MasterCard agreed to unbundle credit cards and signature debit cards, and also temporarily lowered their interchange fees on signature debit cards (Figure 2).

The interchange fee controversy has also attracted great attention from policymakers, who are concerned that interchange fees inflate the cost of card acceptance without leading to proven efficiency.[10] In the two years leading up to the passage of the Durbin Amendment, three separate bills restricting interchange fees were introduced in Congress: a House version of the Credit Card Fair Fee Act of 2009, a Senate version of the same act, and the Credit Card Interchange Fees Act of 2009.[11] Before any of these bills could be brought to a vote, the Dodd-Frank Act was passed and signed into law in July 2010. A provision of the Dodd-Frank Act, known as the Durbin Amendment, mandates a regulation aimed at debit card interchange fees and increasing competition in the payment processing industry.

---

software have been declining rapidly, which should have driven down the card processing costs. Meanwhile, industry statistics show that card fraud rates also have been declining steadily. For the U.S. credit card industry as a whole, the net fraud losses as a percent of total transaction volume has dropped from roughly 16 basis points in 1992 to about 7 basis points in 2009. Data source: *Nilson Report* (various issues).

[9] Visa, MasterCard, and their major issuers reached the settlement agreement with merchants in July 2012. The settlement is currently pending final court approval.

[10] Worldwide, more than 20 countries and areas have started regulating or investigating interchange fees. Primary examples include Australia, Canada, the European Union, France, Spain, and the United Kingdom (Bradford and Hayashi 2008).

[11] None of the bills called for direct regulation of interchange fees, and all three applied to interchange fees for both credit and debit cards (Hung 2009).

Durbin Amendment and Regulation

The Durbin Amendment of the Dodd-Frank Act directs the Federal Reserve Board to regulate debit card interchange fees "reasonable and proportional to the cost incurred by the issuer with respect to the transaction." The Federal Reserve Board subsequently issued the final rule on debit cards in July 2011, effective on October 1, 2011.

The Federal Reserve Board ruling establishes a cap on the debit interchange fees that financial institutions with more than \$10 billion in assets can charge to merchants through merchant acquirers. The permissible fees were set based on the Federal Reserve Board's evaluation of issuers' costs associated with debit card processing, clearance, and settlement. The resulting interchange cap is composed of the following: a base fee of 21 cents per transaction to cover the issuer's processing costs, a five basis point adjustment to cover potential fraud losses, and an additional 1 cent per transaction to cover fraud prevention costs if the issuer is eligible. This cap applies to both signature and PIN debit transactions.

In addition, the regulation sets rules that prohibit certain restraints imposed by card networks on merchants. First, networks can no longer prohibit merchants from offering customers discounts for using debit cards versus credit cards. This gives merchants a way to steer consumers toward using less expensive payment means.[12] Second, issuers must put at least two unaffiliated networks on each debit card and are prohibited from inhibiting a merchant's ability to direct the routing of debit card transactions. This gives merchants more freedom for routing debit transactions through less costly networks. Third, networks can no longer forbid merchants from setting minimum values for credit card payments. Going forward, merchants are allowed to establish such minimum values as long as the minimum does not exceed \$10.

2. EMPIRICAL IMPACT

A direct impact of the debit card regulation is the redistribution of interchange revenues from issuers to merchants. According to a Federal Reserve study, the average debit card transaction in 2009 was approximately \$40. Post regulation, the maximum interchange fee applicable to a typical debit card transaction is capped at 24 cents (21 cents + (\$40 × .05%) + 1 cent), which is about half of its pre-regulation

---

[12] Since the passage of the Cash Discount Act in 1981, merchants have been allowed to offer their customers discounts for paying with cash or checks. However, the card networks have continued to prohibit merchants from offering customers discounts for using one type of card rather than another.

industry average level. As a result, issuers were expected to lose multi-billion dollar annual revenues in terms of the interchange fees that they collect from merchants. In this section, we look into the empirical evidence of the regulation's first-year effects on different players in the debit card market.

Impact on Issuers

The regulation reduces debit card interchange fees by about half and also introduces more competition by abolishing certain network restrictions. As a result, issuers face a big drop in their interchange revenues. Meanwhile, the regulation allows small issuers to be exempt from the interchange fee cap—those with less than $10 billion in assets.[13]

To assess the regulation's impact on covered and exempt issuers, we conduct a study on a subsample of card issuers, which includes all the commercial banks that report their interchange revenues in the quarterly Call Report. Our sample includes 7,049 commercial banks between the first quarter of 2009 and the third quarter of 2012. Among those, we identify 102 covered issuers and 6,969 exempt issuers. The status of exemption is based on whether the bank asset value exceeds the $10 billion threshold as of prior year end.[14]

We first compare the interchange revenues of all covered and exempt banks right before and right after the regulation, as shown in Figure 3 with solid lines. Covered banks had a substantial loss of interchange revenues during the period. Between the third quarter and fourth quarter of 2011 (i.e., the immediate quarter before and after the regulation took effect), covered banks' interchange revenues dropped by $2.1 billion (or 29 percent), equivalent to an $8.5 billion drop annually. In contrast, exempt banks' quarterly interchange revenues did not fall during the same period, instead rising by $11.8 million (or 2 percent).
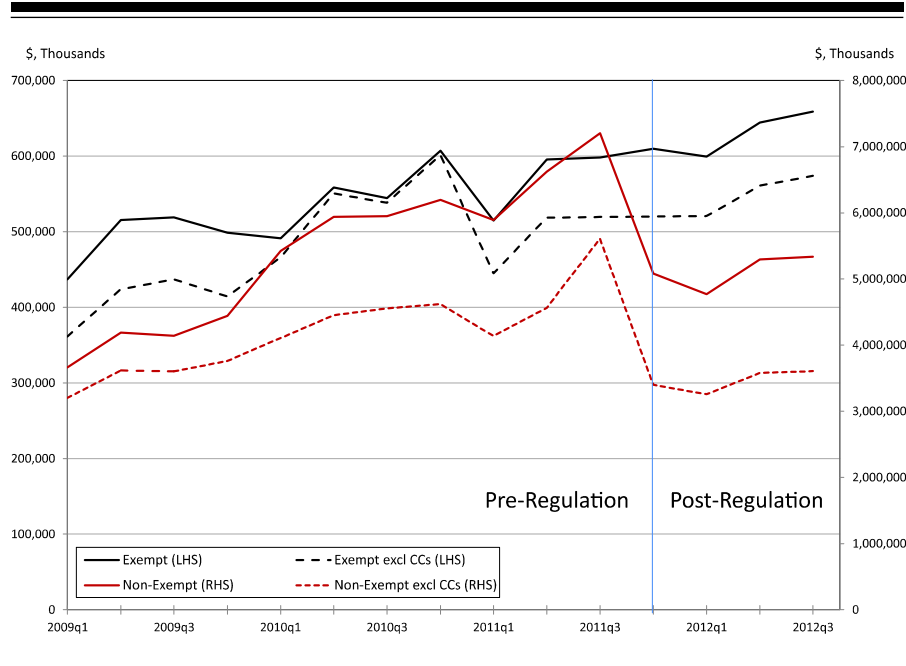
We also compare the interchange revenues one year before and one year after the regulation to control for potential seasonality. The result is similar: Covered banks' annual interchange revenues dropped by $5.4 billion (or 21 percent), while exempt banks' annual interchange revenues increased by $198 million (or 9 percent).

For an alternative check, we construct counterfactual interchange revenues for one year after the regulation (the fourth quarter of 2011

---

[13] This exemption is applied at the holding company level, to ensure that large issuers cannot evade the regulations by establishing subsidiaries under the size limit.

[14] Note that a bank's exemption status may change as its asset size changes, so the sum of non-exempt banks and exempt banks may exceed the total number of banks in the sample.

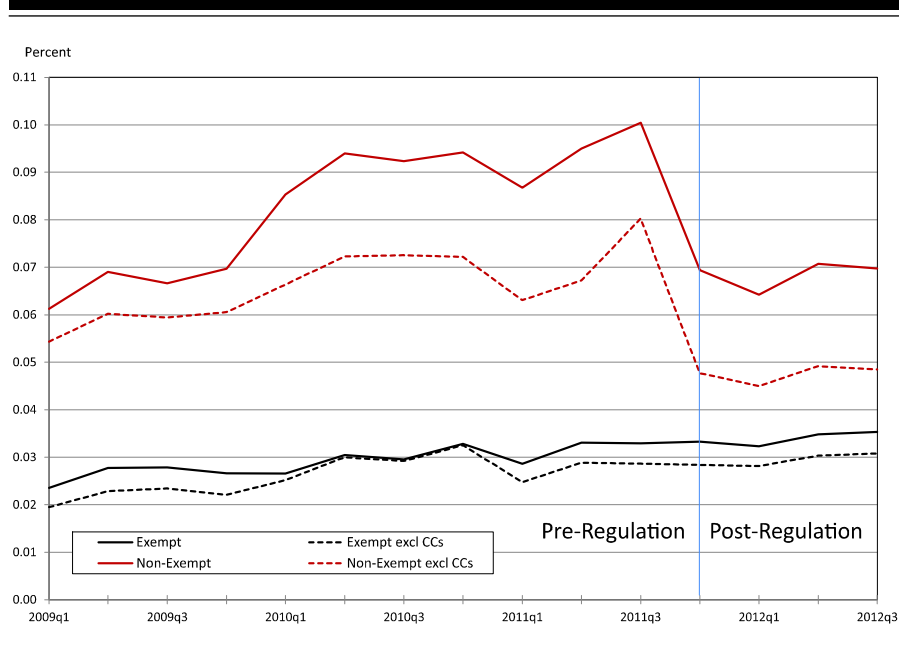**Figure 3 Aggregate Interchange Fee Revenues**

through the third quarter of 2012), assuming that the regulation did not take effect and the annual interchange revenues kept a constant growth rate since two years ago. The finding shows that the annual interchange revenues for covered banks dropped by $10.4 billion (or 34 percent) compared with the counterfactual. In contrast, exempt banks' interchange revenues only dropped by $47 million (or 2 percent).

A limitation of the Call Report data is that they do not separate interchange revenues between debit and credit cards. Therefore, when we conduct the above exercises, we implicitly assume that the changes in interchange revenues were primarily driven by the debit card transactions (but not credit card transactions). In order to focus more on debit interchange fees, we then re-ran the above exercises by excluding mono-lined credit card banks.[15] The pattern, shown in Figure 3 with dashed lines, turns out to be similar. In terms of actual interchange revenues one year before and after the regulation, covered banks'

---

[15] Mono-lined credit card banks are defined as commercial banks with a minimum of 50 percent of assets in consumer lending and 90 percent of consumer lending in the form of revolving credit. See the "Report to the Congress on the Profitability of Credit Card Operations of Depository Institutions," Board of Governors of the Federal Reserve System, 2011.

Figure 4 Aggregate Interchange Revenues as a Percent of Deposits

annual interchange revenues dropped by \$5.1 billion (or 27 percent), while exempt banks' annual interchange revenues increased by \$90.9 million (or 4 percent). In terms of the counterfactual comparison, covered banks' annual interchange revenues dropped by \$7.4 billion (or 35 percent), while exempt banks' annual interchange revenues only dropped by \$31.1 million (or 1 percent).

We also replicated the above exercises by comparing the interchange-revenue-to-bank-deposits ratio. By focusing on the ratio to deposits, we may control for the potential effect of changing bank sizes on interchange revenues. Again, as shown in Figure 4, the results are very similar.

Overall, the empirical evidence suggests that the debit regulation has largely achieved its objective of reducing the interchange revenues for large issuers, while exempt small issuers so far have been well protected.[16]

---

[16] Further monitoring is needed to evaluate the regulation's long-run impact on issuers. There are three concerns that the exempt small issuers might be adversely affected by the regulation. First, networks may voluntarily lower the interchange rates for small issuers to level the playing field between large and small issuers. Second,

Impact on Merchants

Merchants as a whole have greatly benefited from the reduced interchange fees under the regulation. Presumably, the loss of issuers' interchange revenues would be the gain of the merchants. However, the distribution of the gain appears uneven among merchants. In fact, the regulation has yielded an unintended consequence: Interchange fees rose for small-ticket merchants.

Prior to the regulation, Visa, MasterCard, and most PIN networks offered discounted debit interchange fees to small-ticket transactions as a way to encourage card acceptance by merchants specializing in those transactions. For example, Visa and MasterCard used to set the small-ticket signature debit interchange rate at 1.55 percent of the transaction value plus 4 cents for sales of $15 and below. As a result, a debit card would only charge a 7 cents interchange fee for a $2 sale or 11 cents for a $5 sale. However, in response to the regulation, card networks eliminated the small-ticket discounts, and all transactions (except those on cards issued by exempt issuers) have to pay the maximum cap amount set by the regulation (i.e., 21 cents plus 0.05 percent of the transaction value).[17] For merchants selling small-ticket items, this means that the cost of accepting the same debit card doubled or even tripled after the regulation.

The rising interchange fee on small-ticket sales could affect a large number of transactions. According to the 2010 Federal Reserve Payments Study, in 2009 debit cards were used for 4.9 billion transactions below $5, and 10.8 billion transactions between $5–$15. The former accounts for 8.3 percent of all payment card transactions (including credit, debit, and prepaid cards), and the latter accounts for 18.3 percent. Since merchants may have different compositions of transaction sizes, they could be affected differently by the changes of interchange fees.[18] However, merchants who specialize in small-ticket transactions would be most adversely affected.[19]

---

merchants may offer preferential treatment to cards issued by large issuers that carry lower interchange rates. Third, the regulation requires each debit card be connected to at least two unaffiliated networks and merchants have the freedom to choose the lower-cost routing. This provision took effect after April 2012 and small issuers are not exempt from it.

[17] E.g., in the case of signature debit, any sales below $11 now face a higher interchange rate.

[18] Shy (2012) used the data from the Boston Fed's 2010 and 2011 Diary of Consumer Payment Choice to identify the types of merchants who are likely to pay higher and lower interchange fees under the debit regulation.

[19] E.g., Visa classifies merchant sectors specializing in small-ticket sales, which include local commuter transport, taxicabs and limousine, fast food restaurants, coffee shops, parking lots and garages, motion picture theaters, video rental stores, cashless

In response, many small-ticket merchants have tried to offset their higher rates by raising prices, encouraging customers to pay with alternative payment means, or dropping card payments altogether.[20] In the meantime, a lawsuit was filed in November 2011 in federal court by three of the retail industry's largest trade associations and two retail companies against the Federal Reserve's debit interchange regulation. The lawsuit alleges that the Fed has set the interchange cap too high by including costs that were barred by the law, and "forcing small businesses to pay three times as much to the big banks on small purchases was clearly not the intent of the law and is further evidence that the Fed got it wrong."[21]

The unintended consequence on small-ticket merchants calls for a further examination on the regulation, which we will provide in Section 3.

Impact on Consumers

The regulation's impact on consumers is less clear. On the one hand, merchants argue that with a lower interchange fee, they would be able to offer lower retail prices to consumers. On the other hand, issuers argue that they will have to reduce card rewards and raise banking service fees to consumers in order to make up for the lost interchange revenues.

At this point, little empirical evidence has been reported on the change of merchant prices due to the debit interchange regulation. After all, even if the reduced interchange fees have resulted in lower retail prices, the magnitude would be quite small so it is not easy to measure. Meanwhile, several studies report that consumers now face higher banking and card service fees. A recent Pulse debit issuer study shows that 50 percent of regulated debit card issuers with a reward program ended their programs in 2011, and another 18 percent planned to do so in 2012.[22] The Bankrate's 2012 Checking Survey shows that the average monthly fee of noninterest checking accounts rose by 25 percent

---

vending machines and kiosks, bus lines, tolls and bridge fees, news dealers, laundries, dry cleaners, quick copy, car wash and service stations, etc.

[20] See "Debit-Fee Cap Has Nasty Side Effect," *Wall Street Journal*, December 8, 2011.

[21] Source: "Merchants' Lawsuit Says Fed Failed to Follow Law on Swipe Fee Reform," *Business Wire*, November 22, 2011.

[22] The 2012 Debit Issuer Study, commissioned by Pulse, is based on research with 57 banks and credit unions that collectively represent approximately 87 million debit cards and 47,000 ATMs. Research was conducted in April and May of 2012, and the data provided by issuers is for 2011. The sample is nationally representative, with issuers segmented into "regulated" ($\geq$ \$10 billion in assets) and "exempt" ($<$ \$10 billion in assets) to report on the impact of the interchange provision of Regulation II.

compared with last year, and the minimum balance for free-checking services rose by 23 percent.[23] According to the report, the rising bank fees are largely due to banks' response to recent regulations including the debit interchange cap. In addition, several major banks including Bank of America, Wells Fargo, and Chase attempted to charge a monthly debit card fee to their customers in response to the interchange regulation, but they eventually backed down due to customer outrage.[24]

3.   THEORETICAL CONSIDERATIONS

The debit card regulation was created to reduce the interchange fee by capping the fee at the card issuers' marginal cost. To understand the welfare implications of the regulation, we turn to a theoretical analysis in this section.
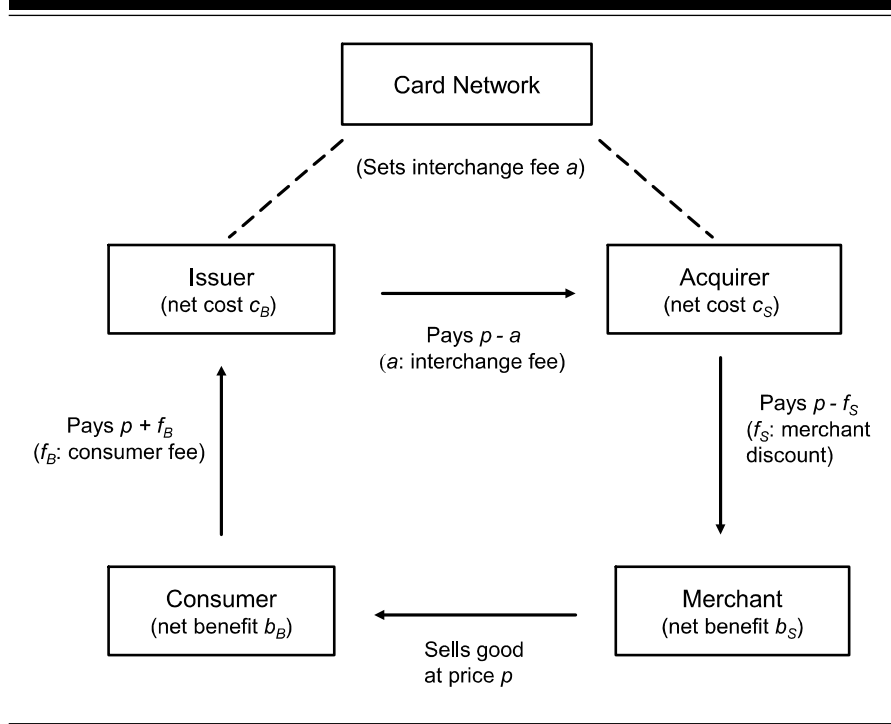
First, we lay out a simple model based on the work of Rochet and Tirole (2011). The model conceptualizes payment cards as a two-sided market, that is, two end-user groups (i.e., merchants and consumers) who jointly use the card services.[25] The interchange fee serves as a transfer between merchants and consumers to balance their joint demand for using cards. Under the assumption of homogenous merchants, the model shows that (1) market-determined interchange fees tend to exceed the socially optimal level, so reducing interchange fees may improve the payments efficiency; (2) however, capping interchange fees based on issuers' marginal cost does not necessarily restore the social optimum; and (3) the theory suggests an interchange fee regulation based on the merchant transaction benefit of accepting cards.

While the simple two-sided market model sheds light on key policy issues related to the interchange fee regulation, it does not address the regulation's unintended consequence on small-ticket merchants. To fill the gap, we then introduce an extension of the model by considering card demand externalities across heterogenous merchant sectors, based on the work of Wang (forthcoming). The findings suggest that an

---

[23] Bankrate surveyed banks in the top 25 U.S. cities to find the average fees associated with checking accounts in their annual Checking Account Survey, which was conducted in July and August 2012.

[24] Source: "Banks Adding Debit Card Fees," *The New York Times*, September 29, 2011.

[25] In recent years, a sizeable body of literature, called "two-sided market theory," has been developed to evaluate payment card market competition and pricing issues. For instance, Baxter (1983), Rochet and Tirole (2002, 2006, 2011), Schmalensee (2002), Wright (2003, 2004, 2012), Armstrong (2006), Rysman (2007, 2009), Prager et al. (2009), Wang (2010, forthcoming), Weyl (2010), Shy and Wang (2011), and McAndrews and Wang (2012).

Figure 5 A Payment Card System

```
                    ┌──────────────┐
                    │ Card Network │
                    └──────────────┘
                  (Sets interchange fee a)

    ┌──────────────┐                    ┌──────────────┐
    │   Issuer     │ ───────────────▶   │   Acquirer   │
    │ (net cost cB)│    Pays p - a      │ (net cost cS)│
    └──────────────┘  (a: interchange   └──────────────┘
                          fee)
                                        Pays p - fS
   Pays p + fB                          (fS: merchant
   (fB: consumer fee)                    discount)

    ┌──────────────┐                    ┌──────────────┐
    │   Consumer   │ ◀───────────────   │   Merchant   │
    │(net benefit bB)   Sells good      │(net benefit bS)
    └──────────────┘   at price p       └──────────────┘
```

alternative regulation, capping the weighted average interchange fee, instead of the maximum interchange fee, may restore the social optimum and avoid the unintended consequence on small-ticket merchants.

A Simple Model

We first lay out a model with homogenous merchants, which is a simplified version of Rochet and Tirole (2011). The model considers a payment card system that is composed of five types of players: consumers, merchants, acquirers, issuers, and the card network, as illustrated in Figure 5.

Consumers

There is a continuum of consumers who purchase goods from competitive merchants selling a homogenous good. Consumers have inelastic demand and each buy one unit of the good. Consumers need to decide which store to patronize. They know the stores' price and card acceptance policy before making the choice. Once in the store they then

select a payment method (a card or an alternative payment method such as cash), provided that the retailer indeed offers a choice among payment means. We assume price coherence such that retailers find it too costly to charge different prices for purchases made by different payment means.[26] Whenever a transaction between a consumer (buyer) and a retailer (seller) is settled by card, the buyer pays a fee f_B to her card issuing bank (issuer) and the seller pays a merchant discount f_S to her merchant acquiring bank (acquirer). We allow f_B to be negative, in which case the cardholder receives a card reward. There are no annual fees and all consumers have a card.

The consumer's convenience benefit of paying by card relative to using cash is a random variable b_B drawn from a cumulative distribution function H on the support $[\underline{b}_B, \overline{b}_B]$, which has a monotonic increasing hazard rate.[27] Cardholders are assumed to only observe the realization of b_B once in the store.[28] Because the net benefit of paying by card is equal to the difference $b_B - f_B$, a card payment is optimal for the consumer whenever $b_B \geq f_B$. The proportion of card payments at a store that accepts cards is denoted $D(f_B)$:

$$D(f_B) = \Pr(b_B \geq f_B) = 1 - H(f_B). \tag{1}$$

Let $v(f_B)$ denote the average net cardholder benefit per card payment:

$$
\begin{aligned}
v(f_B) &= E[b_B - f_B | b_B \geq f_B] \\[2mm]
&= \frac{\int_{f_B}^{\overline{b}_B} (b_B - f_B) dH(b_B)}{1 - H(f_B)} > 0.
\end{aligned} \tag{2}
$$

The monotonic hazard rate of H implies that $v(f_B)$ decreases in f_B.

---

[26] Price coherence is the key feature that defines a two-sided market. Rochet and Tirole (2006) show that the two-sided market pricing structure (e.g., interchange fees) would become irrelevant without the price coherence condition. In reality, price coherence may result either from network rules or state regulation, or from high transaction costs for merchants to price discriminate based on payment means. In the United States, while merchants are allowed to offer their customers discounts for paying with cash or checks, few merchants choose to do so. On the other hand, card network rules and some state laws explicitly prohibit surcharging on payment cards.

[27] The hazard rate is assumed increasing to guarantee concavity of the optimization problem.

[28] This is a standard assumption introduced by Wright (2004) and used in the subsequent literature, which simplifies the analysis of retailers' acceptance of cards without changing the equilibrium outcome. Alternatively, Rochet and Tirole (2002) assume cardholders differ systematically in the benefit that they derive from card payments. However, as shown in Rochet and Tirole (2011), these two alternative assumptions deliver broadly convergent results.

Merchants

Merchants derive the convenience benefit b_S of accepting payment cards (relative to handling cash). By accepting cards under the price coherence assumption, a merchant is able to offer each of its card-holding customers an additional expected surplus of $D(f_B)v(f_B)$, but faces an additional expected net cost of $D(f_B)(f_S - b_S)$ per cardholder. Denote c as the cost of the good. Competitive merchants then set a retail price equal to marginal cost, namely

$$p = c + D(f_B)(f_S - b_S) \qquad (3)$$

if they accept cards, or $p = c$ if they reject cards. Consumers choose the stores that accept cards if and only if their increased surplus $D(f_B)v(f_B)$ exceeds the price increase $D(f_B)(f_S - b_S)$. Therefore, all merchants accept cards if and only if

$$f_S \leq b_S + v(f_B). \qquad (4)$$

Rochet and Tirole (2011) show that (4) also holds for a variety of other merchant competition setups, including monopoly and Hotelling-Lerner-Salop differentiated products competition with any number of retailers. Wright (2010) shows the same condition holds for Cournot competition.

Acquirers

We assume acquirers incur per-transaction cost c_S and are perfectly competitive. Thus, given an interchange fee a, they charge a merchant discount f_S such that

$$f_S = a + c_S. \qquad (5)$$

Because acquirers are competitive, they play no role in our analysis except passing through the interchange charge to merchants.

Issuers

Issuers are assumed to have market power.[29] We consider a symmetric oligopolistic equilibrium at which all issuers charge the same

---

[29] This is a standard assumption in the literature. As pointed out in Rochet and Tirole (2002), the issuer market power may be due to marketing strategies, search costs, reputation, or the nature of the card. Note that were the issuing side perfectly competitive, issuers and card networks would have no preference over the interchange fee, and so the latter would be indeterminate.

consumer fee f_B, which can be negative if the cardholder receives a reward. Issuers incur a per-transaction cost c_B and receive an interchange payment of a for a card transaction. At equilibrium, the net per-transaction cost for issuers is $c_B - a$. For simplicity, we consider that issuers set a constant markup φ.[30] Hence, the consumer fee f_B is determined as

$$f_B = \varphi + c_B - a. \tag{6}$$

Network

We consider a monopoly network, which sets the interchange fee a to maximize the total profit of issuers from card transactions, namely,

$$\Pi = \varphi D(f_B) = \varphi[1 - H(f_B)] \ .$$

Alternatively, we could consider a regulator who instead sets the interchange fee to maximize social welfare or user surplus.

Timing

The timing of events is as follows.

1. The card network (or the regulator) sets the interchange fee a.

2. Issuers and acquirers set fees f_B and f_S. Merchants then decide whether to accept cards and set retail prices.

3. Consumers observe the retail prices and whether cards are accepted, and choose a store. Once in the store, the consumer receives her draw of b_B and decides which payment method to use.

Model Characterization

We first consider the market equilibrium under a monopoly network. Given the model setup, the network solves the following problem:

$$\max_{a} \varphi[1 - H(f_B)] \tag{7}$$

$$s.t. \quad f_B = \varphi + c_B - a, \tag{8}$$

---

[30] This is a simplifying assumption, and the findings of the model hold if we instead consider an endogenous issuer markup. See Wang (forthcoming).

$$a \le b_S + v(f_B) - c_S, \tag{9}$$

where the condition (9) is derived from equations (4) and (5).

Since the issuers' profit (7) is maximized by setting the highest possible merchant fee at which merchants still accept cards, the condition (9) holds with an equality. Therefore, the profit-maximizing interchange fee is determined as

$$a^m = b_S + v(f_B^m) - c_S, \tag{10}$$

where f_B^m solves

$$b_S + v(f_B^m) - c_S = \varphi + c_B - f_B^m.$$

Here the superscript m denotes market-determined rates.

This simple model helps illustrate the impact of an interchange cap regulation as we found (or expect to find) in empirical evidence. According to the model, if a regulation pushes down the interchange fee to a level a^r, where $a^r < a^m$, we would have the results as follows.

Result 1 *If a regulation pushes down the interchange fee below the market-determined rate such that $a^r < a^m$, the model implies that (i) consumer card fee increases; (ii) merchant retail price decreases; (iii) card usage falls; and (iv) issuers' profit declines.*[31]

Proof.   (i) Conditions (8) and (9) imply that consumer card fee f_B increases as the interchange fee a decreases; (ii) according to (3), merchant retail price p depends on $D(f_B)$ and f_S, both increasing in a; (iii) card usage $1 - H(f_B)$ decreases in f_B; and (iv) issuer profits $\varphi[1 - H(f_B)]$ decrease in f_B. ∎

We now turn to the welfare discussion. We first consider that the card network is run by a regulator who maximizes social welfare. Social welfare is generated if consumers use cards for payment at retailers whenever consumer and merchant joint transaction benefits exceed the joint cost of doing so, namely $b_S + b_B > c_B + c_S$. It can be shown that social welfare is the sum of issuers' profit, consumer surplus, and merchants' profit. Accordingly, the regulator solves the problem

$$\max_{f_B} \int_{f_B}^{\bar{b}_B} (b_S + b_B - c_B - c_S) dH(b_B). \tag{11}$$

The first-order condition with regard to f_B requires that

$$f_B^w = c_B + c_S - b_S,$$

---

[31] In theory, an interchange fee cap can be set too low so that the card market shuts down. For example, for a distribution H with a finite support, consumer fee f_B can become so high that $1 - H(f_B) = 0$.

which implies that the welfare-maximizing interchange fee is

$$a^w = b_S - c_S + \varphi. \tag{12}$$

Here the superscript w denotes welfare-maximizing rates.

Comparing (10) and (12), we have the following findings.

Result 2 *(i) When $\varphi < v(f_B^m)$, the market-determined interchange fee a^m is higher than the welfare-maximizing interchange fee a^w; (ii) when $\varphi \geq v(f_B^m)$, the market-determined interchange fee a^m coincides with the welfare-maximizing interchange fee a^w.*

Proof. (i) Equations (10) and (12) suggest that $a^w = a^m - v(f_B^m) + \varphi$. Therefore, $a^m > a^w$ when $\varphi < v(f_B^m)$. (ii) When $\varphi \geq v(f_B^m)$, we have $a^w \geq a^m$. Because a^m is the highest interchange fee that merchants can accept, a^m then coincides with the welfare-maximizing interchange fee a^w. ∎

Similarly, we can consider the card network run by a regulator who maximizes user surplus. Note that user surplus is the sum of consumer surplus and merchants' profit (but not issuers' profit). In the case of competitive merchants, merchants earn zero profit so user surplus equals consumer surplus. Accordingly, the regulator solves the following problem:

$$\max_{f_B} \int_{f_B}^{\bar{b}_B} (b_S + b_B - f_B - f_S) dH(b_B). \tag{13}$$

Recall (5) and (6), which imply that $f_B + f_S = c_B + c_S + \varphi$. Maximizing the user surplus (13) then requires

$$f_B^u = c_B + c_S + \varphi - b_S, \tag{14}$$

which implies that the user-surplus-maximizing interchange fee is

$$a^u = b_S - c_S. \tag{15}$$

Here, the superscript u denotes user-surplus-maximizing rates.

Comparing (10), (12), and (15), we have the following findings.

Result 3 *(i) The interchange fee a^u maximizing the user surplus is lower than the welfare-maximizing interchange fee a^w; (ii) a^u is also lower than the market-determined interchange fee a^m.*

Proof. (i) Equations (12) and (15) suggest that $a^u = a^w - \varphi$, so $a^u < a^w$. (ii) Equations (10) and (15) suggest that $a^u = a^m - v(f_B^m)$, so $a^u < a^m$. ∎

Results 2 and 3 show that the market-determined interchange fee tends to be too high, based on the criterion of either social welfare maximization or user surplus maximization. The reason is that under price coherence, merchants internalize consumers' expected card usage benefits when they decide whether to accept cards and set retail prices. This allows the card network to charge too high an interchange fee and too low a consumer fee. As a result, cards get used even when consumer and merchant joint card usage costs exceed their joint transaction benefits. Therefore, regulating down the interchange fee may potentially improve payments efficiency.

However, (12) and (15) also clarify that the socially optimal interchange fee is not determined by the issuer cost, c_B, but rather by the merchant transaction benefit of accepting cards, b_S. Particularly, (15) suggests that a regulator may consider setting the merchant discount $f_S = b_S$, at which the resulting interchange fee maximizes the user surplus. This is the criterion proposed by Rochet and Tirole (2011), which they call the "merchant avoided-cost test."[32]

Small-Ticket Effect

Our analysis so far does not explain the regulation's unintended consequence on small-ticket merchants. This is largely because we have only assumed homogenous merchants in the model. However, even if in a model with multiple (heterogenous) merchant sectors, as long as those merchant sectors are independent from one another in terms of card acceptance and usage, it is still a puzzle to think why card networks would abandon the interchange differentiation in response to a cap regulation. In other words, if it was profitable for a card network to charge a lower fee to small-ticket merchants in the absence of regulation, why would the card network want to change the practice because of a non-binding cap? To address this issue, Wang (forthcoming) extends the model of Rochet and Tirole (2011) by considering card demand externalities across merchant sectors.

In the setup of Wang (forthcoming), there are multiple merchant sectors (e.g., large-ticket merchants and small-ticket merchants). Different merchant sectors are charged different interchange fees due to their (observable) heterogenous benefits of card acceptance and usage. In addition, consumers' benefits of using cards in a merchant sector are

---

[32] Focusing on user surplus is legitimate if card issuer profits are not considered or weighed much less by competition authorities. The criterion proposed by Rochet and Tirole (2011) is adopted by the European Commission and renamed the "merchant indifference test," while some other countries, including the United States and Australia, adopt the issuer cost-based cap regulation.

positively affected by their card usage in other sectors, which is called the "ubiquity externalities."[33] Based on this setup, Wang (forthcoming) again finds that market-determined interchange fees tend to exceed the socially optimal level. The reason is similar to before: Under price coherence, consumers are provided with excessive incentives to use cards. In addition, Wang (forthcoming) offers the following new findings.

Result 4 *(i) Card demand externalities across merchant sectors explain why card networks eliminate the interchange fee discount to small-ticket merchants in response to the interchange cap regulation; (ii) the social planner who maximizes social welfare would set a discounted interchange fee for small-ticket merchants; (iii) capping the weighted average interchange fee, instead of the maximum interchange fee, may restore the social optimum and avoid the unintended consequence on small-ticket merchants.*

Wang (forthcoming) offers a formal derivation of the above results. Here we provide an intuitive discussion. First, the "ubiquity" externalities may explain card networks' pricing response to the cap regulation: Before the regulation, card networks offer a discounted interchange fee (i.e., a subsidy) to small-ticket merchants because their card acceptance boosts consumers' card usage for large-ticket purchases from which card issuers can collect higher interchange fees. After the regulation, however, the interchange fees on large-ticket purchases are capped. As a result, card issuers profit less from this kind of externality so card networks discontinued the discount.

Second, despite privately determined interchange fees tending to exceed the socially optimal level, the social planner who maximizes social welfare would behave similar to the private network by setting differentiated interchange fees, i.e., charging a high interchange fee to large-ticket merchants but a low interchange fee to small-ticket merchants. Essentially, both the social planner and the private network treat the small-ticket transactions as a loss leader. By subsidizing small-ticket transactions, they internalize the positive externalities of card usage between the small-ticket and large-ticket sectors.

Third, it is possible to design a cap regulation that may restore the social optimum and avoid the unintended consequence on small-ticket merchants. Conceptually, this can be done by imposing a cap on the weighted average interchange fee instead of the maximum interchange

---

[33] Ubiquity has always been a top selling point for brand cards. This is clearly shown in card networks' advertising campaigns, such as Visa's "It is *everywhere* you want to be," and MasterCard's "There are some things money can't buy. For *everything* else, there's MasterCard."

fee. This alternative regulation provides card networks with incentives to continue charging differentiated interchange fees to different merchant sectors. Note that, under such a cap, a card network can either set an uniform interchange fee limited by the cap, or they could set an above-cap (respectively, below-cap) interchange fee to large-ticket (respectively, small-ticket) transactions as long as the weighted average fee does not exceed the cap. When the cap and weights are appropriately chosen, profit-driven card networks are induced to set differentiated interchange fees at the socially optimal level.

4. CONCLUSION

The recent debit card regulation introduced by the Durbin Amendment to the Dodd-Frank Act has generated significant impact on the U.S. payments industry. In this article, we provide a review of the first-year experience of the regulation.

We first investigate the regulation's empirical impact on different players in the debit card market. We find that the regulation has substantially reduced interchange revenues of large issuers who are covered by the regulation, while small issuers who are exempt have been shielded well so far. We also find that merchants are affected unevenly by the regulation. While merchants as a whole have benefited from the reduced interchange rates, merchants specializing in small-ticket transactions have been adversely affected.

We then provide a theoretical framework to assess the regulation's implications on payments efficiency. We show that market-determined interchange fees tend to be too high compared with the social optimum, so regulating down interchange fees could be welfare enhancing. However, the regulation based on issuer cost is less consistent with theoretical foundation. Rather, policymakers may consider capping interchange fees based on the merchant transaction benefit of accepting cards. Moreover, we discuss that capping the weighted average interchange fee, instead of the maximum interchange fee, may avoid the unintended consequence on small-ticket merchants.

Many issues remain to be addressed for improving the efficiency of the U.S. card payments system. First of all, in order to assess the pricing and performance of payment card markets, policymakers need a good measurement of the costs and benefits of different payment means. These include both private costs and benefits as well as social costs and benefits. Second, policymakers may want to consider policy options other than interchange fee regulation. For instance, in theory, if merchants can set different retail prices conditioning on payment means (e.g., surcharging card usage), the interchange fee becomes less of an

issue. However, those policy options may also have their own limitations, so some cautions need to be taken.[34] Finally and more broadly, we need a better understanding of the functioning of payment card markets, especially the complicated issues regarding the two-sided market nature, the network externalities, and the cooperation and competition between payment platforms.

REFERENCES

Armstrong, Mark. 2006. "Competition in Two-Sided Markets." *RAND Journal of Economics* 37 (Autumn): 668–91.

Baxter, William. 1983. "Bank Interchange of Transactional Paper: Legal Perspectives." *Journal of Law and Economics* 26: 541–88.

Bradford, Terri, and Fumiko Hayashi. 2008. "Developments in Interchange Fees in the United States and Abroad." Federal Reserve Bank of Kansas City *Payments System Research Briefings* (May).

Hayashi, Fumiko. 2012. "Discounts and Surcharges: Implications for Consumer Payment Choice." Federal Reserve Bank of Kansas City *Payments System Research Briefings* (June).

Hung, Christian. 2009. "An Update of Interchange Legislation in the United States." Federal Reserve Bank of Kansas City *Payments System Research Briefings*.

Levitin, Adam. 2010. "Interchange Regulation: Implications for Credit Unions." Filene Research Institute Research Brief 224 (November).

McAndrews, James, and Zhu Wang. 2012. "The Economics of Two-Sided Payment Card Markets: Pricing, Adoption and Usage." Federal Reserve Bank of Richmond Working Paper 12-06 (September).

---

[34] For example, in countries where card surcharging is allowed, few merchants choose to do so. Moreover, for some merchants who are indeed surcharging, they are found surcharging more than card acceptance costs or imposing surcharging in nontransparent ways. See Hayashi (2012).

Prager, Robin A., Mark D. Manuszak, Elizabeth K. Kizer, and Ron Borzekowski. 2009. "Interchange Fees and Payment Card Networks: Economics, Industry Developments, and Policy Issues." *Finance and Economics Discussion Series* 2009-23. Washington: Board of Governors of the Federal Reserve System (May).

Rochet, Jean-Charles, and Jean Tirole. 2002. "Cooperation among Competitors: Some Economics of Payment Card Associations." *RAND Journal of Economics* 33 (Winter): 549–70.

Rochet, Jean-Charles, and Jean Tirole. 2006. "Two-Sided Markets: A Progress Report." *RAND Journal of Economics* 35 (Autumn): 645–67.

Rochet, Jean-Charles, and Jean Tirole. 2011. "Must-Take Cards: Merchant Discounts and Avoided Costs." *Journal of the European Economic Association* 9 (3): 462–95.

Rysman, Marc. 2007. "An Empirical Analysis of Payment Card Usage." *Journal of Industrial Economics* 55 (1): 1–36.

Rysman, Marc. 2009. "The Economics of Two-Sided Markets." *Journal of Economic Perspectives* 23 (Summer): 125–43.

Schmalensee, Richard. 2002. "Payment Systems and Interchange Fees." *Journal of Industrial Economics* 50 (June): 103–22.

Shy, Oz. 2012. "Who Gains and Who Loses from the 2011 Debit Card Interchange Fee Reform?" Federal Reserve Bank of Boston Public Policy Discussion Paper 12-6 (June).

Shy, Oz, and Zhu Wang. 2011. "Why Do Payment Card Networks Charge Proportional Fees?" *American Economic Review* 101 (June): 1,575–90.

Wang, Zhu. 2010. "Market Structure and Payment Card Pricing: What Drives the Interchange?" *International Journal of Industrial Organization* 28 (January): 86–98.

Wang, Zhu. Forthcoming. "Demand Externalities and Price Cap Regulation: Learning from A Two-Sided Market." Federal Reserve Bank of Richmond Working Paper.

Weyl, Glen. 2010. "A Price Theory of Multi-Sided Platforms." *American Economic Review* 100 (September): 1,642–72.

Wright, Julian. 2003. "Optimal Card Payment Systems." *European Economic Review* 47 (August): 587–612.

Wright, Julian. 2004. "Determinants of Optimal Interchange Fees in Payment Systems." *Journal of Industrial Economics* 52 (March): 1–26.