Payment Network Competition

Lulu Wang*

September 16, 2022

Abstract

I estimate how payment network entry affects the cost of payments and welfare. I use bank payment volumes and consumer surveys to estimate an industry model of payment network pricing, consumer payment choice, and merchant acceptance. Incumbent credit card networks respond to entry by increasing merchant fees by 8 bps to fund 13 bps more consumer rewards. Retail prices rise by 16 bps and dissipate the usual welfare gains of entry. Consumer and total surplus fall by $7 and $10 billion, respectively.

*Wang: Stanford GSB, 655 Knight Way, Stanford CA 94305. Email: luluyw@stanford.edu. Job market paper. I am extremely grateful to my advisors Amit Seru, Darrell Duffie, Ali Yurukoglu, and Claudia Robles-Garcia. I thank Lanier Benkard, Greg Buchak, Jacob Conway, Liran Einav, Joseph Hall, and Ben Hebert for helpful comments. I acknowledge support from the National Science Foundation Graduate Research Fellowship under Grant Number 1656518.

1 Introduction

Many have argued that high prices for accepting credit and debit cards in the US reflects a lack of competition. In 2020, the Department of Justice (DOJ) challenged Visa's acquisition of a nascent payment network, Plaid, on the grounds that "both merchants and consumers would be deprived of competition that would drastically lower costs for online debit transactions." The intuitive view stands in contrast to the ambiguous answer from the theoretical literature (Rochet and Tirole, 2003; Guthrie and Wright, 2007; Edelman and Wright, 2015; Liu et al., 2021). Payment networks compete not only over merchants but also over consumers. Under certain conditions, competition can cause networks to raise fees for merchants in order to fund larger consumer rewards.

In this paper I quantify the effects of payment network competition on the cost of payments and welfare. The absence of clean reduced form shocks leads me to build and estimate an industry model featuring three competing multiproduct payment networks: Visa, Mastercard, and Amex. I use data on bank payment volumes, consumer card holdings, and merchant card acceptance to estimate consumer and merchant preferences over payment instruments and networks' costs of processing payments. I model rich consumer and merchant substitution patterns between payment methods. I assume the observed data reflects the equilibrium of a game between competing networks. I simulate the entry of a new payment platform by solving for the new equilibrium with one more network.

I find that entry of a new fintech payment platform that competes with credit cards raises the cost of payments by 16 bps, exacerbates regressive transfers, and reduces average consumer surplus by -7 bps. Although many consumers adopt the new product, the overall increase in retail prices outweighs the traditional welfare gains from a new entrant. Scaling the overall loss in surplus to the roughly \$10 trillion in consumer to business payments translates to \$7 billion a year in lost consumption. Network profits also fall, leading to a \$10 billion reduction in total surplus.

The market failure in payments reflects externalities, not market power. The central friction is that merchants charge consumers the same price no matter how they pay (Stavins, 2018). Uniform pricing means consumers do not internalize the effect of their payment adoption on the prices other consumers pay. Consumers are locked in a prisoner's dilemma. While they collectively prefer a world with low card use, they individually adopt cards to take advantage of higher rewards. Networks exploit this coordination failure to charge merchants high fees that fund large consumer rewards. Unlike in traditional one sided markets where the concern is that output is too low, in payments the concern is that card adoption is too high.

My reduced form analysis guides how I model demand. A regulatory shock that ended debit rewards at large banks but not small banks caused a relative decline in debit card use at large banks. Thus consumer payment choice is sensitive to rewards. Primary credit card users' secondary card is likely to be a credit card from a different network, even though debit cards are popular among the general population. I interpret primary and secondary cards as first and second choices and conclude that consumers have unobserved heterogeneity in their preferences over payment instruments. Accepting cards increases sales by a large amount for the average merchant, but not all merchants accept cards. Thus merchants accept cards to increase sales, but merchants vary in their benefits.

I build and estimate a structural model in which consumers adopt cards that pay high rewards and are widely accepted, and merchants accept cards when the the increase in sales exceed the loss from fees. The model has three kinds of players: consumers, merchants, and payment networks. Consumers choose which cards to put in their wallet and consumption over merchants. Consumers buy more from merchants that accept the cards in their wallet and set low prices. Merchants choose the optimal subset of payments to accept and set prices. Networks maximize profits by adjusting consumer rewards and merchant fees, accounting for the effect of prices on adoption decisions. Consumers vary unobservably in their preferences over payment instruments, and merchants vary unobservably in their benefits of card acceptance.

I estimate the model with indirect inference. I assume that the observed data reflects an equilibrium of the game. Consumer preference parameters are pinned down by aggregate shares and the reduced form facts on debit rewards and consumer card holdings. The distribution of merchant benefits comes from the share of card consumers' spending at merchants that accept card and an estimate of the causal effect of card acceptance on sales. I estimate network costs by inverting the first order conditions at the observed equilibrium.

The model fits well. I match the fact that primary credit card consumers are less likely to use a debit card for their secondary card. I am also able to match out of sample predictions for how credit card volumes change after a shock to debit rewards. The estimated network marginal cost parameters are consistent with accounting data.

I estimate that consumers are price sensitive while merchants are not. A one basis point increase in Visa credit rewards increases Visa credit's market share by 2.8% while decreasing the market share of Mastercard credit, Mastercard debit, and cash by -2.0%, -0.7%, and -0.6%, respectively. In contrast, a 1 bps increase in merchant fees for Visa credit would only cause a -0.16% change in the share of firms who accept Visa credit.

Introducing a fintech payment network causes incumbent credit card networks to raise merchant fees by 8 bps to fund 13 bps larger subsidies. I assume the entrant competes for credit card consumers but is not a substitute for credit cards at the point of sale[1].

Entry hurts debit and cash consumers more than credit consumers. The combination of higher credit card fees and more consumers using credit cards causes retailers to raise prices by 16 bps. This directly hurts cash users who do not switch. For consumers who do not switch, cash and debit users' welfare falls by -16 bps and -11 bps, respectively. Even incumbent credit card consumers' welfare falls by -6 bps.

I estimate that consumer surplus *declines* by -7 bps as debit consumers switch to credit to take advantage of rewards. In traditional markets, new product entry is associated with higher consumer surplus due to gains from variety and reductions in markups (Petrin, 2002). In payments, this result does not hold. Competition improves the value of cards relative to cash by 10 bps, but the 16 bps increase in retail prices eliminates these gains.

Another interpretation is that consumer effort spent to chase credit card rewards is socially wasteful. Credit cards pay rewards while debit cards do not. Therefore consumers must value the non-pecuniary characteristics of debit[2]. Agents who switch sacrifice non-pecuniary utility to take advantage of rewards. However, the private gains from rewards are not social gains because the merchant fees that fund rewards inflate retail prices for other consumers (Edelman and Wright, 2015). On net, higher rewards cause consumers to use payment methods with lower non-pecuniary value, reducing consumer surplus. Competition causes networks to raise subsidies by more than they increase fees, which causes a 4 bps increase in utility, but consumers lose -10 bps in non-pecuniary utility. The loss in non-pecuniary utility can be interpreted as effort spent chasing rewards and is a source of social losses.

I also explore two more counterfactuals that suggest an important source of welfare loss is the fact that debit cards cannot pay rewards while credit cards do. This rationalizes why entry is harmful. Because entry increases the rewards gap between credit and debit cards, it hurts welfare.

In the first counterfactual I let debit cards charge merchants 1% instead of the current

[1] This substitution reflects facts about new fintech platforms that I discuss in section 8

[2] Another possibility is that debit card users do not have access to credit cards, but in section 4 I show that the marginal consumers between debit and credit tend to have access to both debit cards and rewards credit cards. The source of this non-pecuniary utility could be either consumers who value the simplicity of not having to worry about interest charges or late fees, or debt aversion.

0.72% effective price ceiling. Debit rewards rise by 22 bps. Although retail prices rise by 3 bps, consumers as a whole get 5 bps higher utility as fewer debit card consumers choose credit cards for the rewards.

In the second counterfactual, I merge Amex and Mastercard and assume marginal costs do not change. More market power causes credit card rewards to fall by -11 bps and credit card fees to rise by 3 bps. Surprisingly, consumer surplus rises by 1 bps as consumers no longer face strong incentives to adopt credit cards. The merger counterfactual highlights the importance of retail price externalities, as mergers without efficiencies in typical one sided markets never raise consumer surplus.

Although I focus on payment networks, my empirical approach is also relevant to other two sided markets where network pricing can affect retail prices. Advertising platforms like Facebook or Google connect merchants with consumers. Just as in payments, merchants charge consumers the same price regardless of whether consumers saw the ad. Competing platforms make larger investments in consumer benefits funded by higher prices of advertising. Higher advertising prices get passed through to final goods prices, leaving an ambiguous effect on consumers. The net effects depend on how much consumers multihome across platforms and how effective investment is at attracting users. The empirical approach in my paper can then be used to study the price and welfare effects of competition between other platforms that connect buyers and sellers.

1.1 Related Literature

The primary contribution of the paper is to demonstrate that payment competition can raise merchant fees and reduce consumer surplus in a realistic model of the US payments market. The closest related work is Huynh et al. (2022), who also estimate a structural model of consumer and merchant adoption of cards[3]. However, my result depends on the interaction between merchant pricing and network competition, neither of which they model. In my model higher interchange can reduce *consumer* surplus, whereas the model in Huynh et al. (2022) would predict higher interchange always benefits consumers.

Several papers have similar goals of evaluating the welfare effects of payment pricing. Mukharlyamov and Sarin (2022) study Durbin and evaluate the passthrough of interchange to consumers and merchants. Their identification approach and analysis of the Durbin Amendment motivate my reduced form work. Felt et al. (2020) is also

[3]Li et al. (2020) also build an equilibrium model but assume that merchants accept cards to reduce costs. As a result, in their model card users cross subsidize cash users.

concerned about the distributional effects of credit card rewards holding fixed consumer and merchant adoption. I extend their welfare analysis to include aggregate consumer surplus and total surplus effects.

My paper contributes to the literature on two sided markets and payments by simultaneously capturing the key factors that shape the effects of competition on prices and welfare. Past work has shown that merchant heterogeneity (Rochet and Tirole, 2003; Guthrie and Wright, 2007), consumer multihoming (Armstrong, 2006; Anderson et al., 2018; Liu et al., 2021; Bakos and Halaburda, 2020; Gentzkow et al., 2022), and merchant price setting behavior (Edelman and Wright, 2015) all have important implications for the effects of competition. In contrast to many existing empirical and theoretical papers in the two sided markets literature such as Rysman 2004; Rosaia 2020; Tan and Zhou 2021, I allow platform competition to affect the value of the outside option.

Lastly, my paper contributes to the finance literature on payments by providing a model of how important drivers of payment adoption are determined in equilibrium. Many papers have documented the importance of adoption externalities (Gowrisankaran and Stavins, 2004; Rysman, 2007; Higgins, 2020; Crouzet et al., 2020), unobserved preference heterogeneity (Koulayev et al., 2016; Huynh et al., 2020), and rewards (Arango et al., 2015; Ru and Schoar, 2020) in determining consumer payment choice. My model explains how networks take these preferences into account in setting prices. Berg et al. (2022) document that consumers of many new fintech products seem unwilling to substitute to traditional cards even when available. I use this fact in defining my counterfactual. Ghosh et al. (2021); Parlour et al. (2022) study how payment data complement traditional banking activities. My model provides a benchmark for how payment networks that do not bundle with other services should compete.

2 Institutional Details

Payment card networks in the United States connect consumers with merchants. Networks, unlike stores, compete over two "sides" of the market: consumers and merchants. Because merchants charge consumers the same retail price no matter how they pay, networks gain market share by setting high merchant fees to fund consumer rewards.

2.1 Network Structure, Pricing, and the Interchange Fee

Card networks shape both the prices that merchants pay to accept cards as well as the subsidies consumers receive from using cards. In doing so they also can shape consumers' incentives to adopt cards.

Incumbent payment networks like Visa and MC connect four players: the merchant, the merchant's bank, the consumer's bank, and the consumer[4]. Figure 1 illustrates the typical flow of money between these players. When a consumer uses her Chase Freedom Unlimited credit card to buy $100 of product at a large retailer, the merchant might pay a $2.25 merchant discount fee to her acquiring bank to process the transaction. The Acquiror can be a bank like Wells Fargo or a fintech player like Square who works with a bank to connect the merchant to the Visa network. The acquiror will use some of that fee to cover its costs, but then must also send $1.75 to the issuing bank, Chase, in the form of interchange. The issuer and the acquiror collectively then pay around 14 cents in assessment fees to Visa. While some of the $1.75 of interchange fees goes towards covering the issuer's costs, a large part of it is also rebated back to the consumer. In the case of the Chase Freedom Unlimited card, this is literally $1.50 cents.

Visa and MC influence merchant costs and consumer benefits by adjusting interchange. Gans (2007) shows that after Australia regulated credit interchange in 2003, the merchant discount fee fell almost one for one. Valverde et al. (2016) document that merchant discounts moved one for one with interchange fees over a 10 year window in Spain when interchange fees fell due to regulation. On the consumer side of the market, banks ended rewards debit programs after the Durbin Amendment capped debit card interchange fees (Hayashi, 2012; Schneider and Borra, 2015). Agarwal et al. (2018) show that around 70% of issuers' credit card interchange income goes towards consumer rewards and covering the cost of fraud. When Australia capped credit interchange fees, banks reduced credit rewards (Chang et al., 2005).

2.2 The Role of Issuers

My model abstracts away from the strategic decisions of issuers. In my model, networks directly set the merchant discount fee and consumer rewards. This is correct for proprietary networks like Amex or fintechs like PayPal, for whom there are no issuers or acquirors. In the case of Visa and MC, this assumption can be justified under the assumption that Visa internalizes the profits of issuers and acquirors and so the three types of entities can operate in a vertically integrated manner. This is reasonable given Visa's need for both issuers and acquirors to make long term investments and the extensive

[4]Proprietary payment networks like American Express have similar pricing structure, but can be thought of as a vertically integrated entity that combines Visa, the acquiror, and the issuer. In their case, they directly set the prices that merchants pay and the benefits consumers receive.

incentive payments Visa pays to both sides of the market[5].

Treating both merchant fees and rewards as variables set at the network level is essential for understanding how entry simultaneously affects merchant fees and consumer rewards. Competition between Chase and Bank of America shapes the rewards they offer. But to understand how merchant fees might change in response to entry and how increased merchant fee income may fund higher rewards, it's essential to have a player that can simultaneously influence merchant fees and consumer subsidies. Historical shocks to interchange highlight how the primary determinant of consumer rewards in many cases was not issuer competition, but rather limits on how much networks could charge merchants.

I abstract away from issuers' choices of interest rates and credit lines. My model will assume that the subsidies paid to the consumer side of the market are passed through to rewards, not to rates. This is a reasonable assumption to the extent that there are two segments of credit card consumers – revolvers who are sensitive to interest rates and transactors who are sensitive to rewards (Adams and Bord, 2020). Under this assumption, issuer fee income from transactions should only have a small effect on interest rates. To the extent that consumers value paying with credit lines, that will show up in product fixed effects.

2.3 Uniform Prices for Different Payment Methods

Merchants charge uniform prices to consumers who use different payment methods, and this shapes network competition. Historically, uniform prices were the result of no surcharge laws imposed by the networks and state governments. Despite the repeal of these laws, merchants are reluctant to surcharge for higher cost payments (Stavins, 2018).

When merchants charge uniform prices, networks gain consumer market share by charging merchants higher fees to fund consumer rewards. If Visa raises both merchant discounts and cash back by one cent, Visa customers benefit from the one cent increase in cash back, but only bear part of the cost of higher merchant fees through retail prices. Consumers now have stronger incentives to use Visa cards. In a frictionless world where merchants pass on the cost of payments to the consumer on each transaction, networks have no incentive to raise merchant fees to fund customer rewards since consumers would still pay the same net prices (Gans and King, 2003).

Theoretically, an individual merchants' gain from passing on the cost of payments

[5] In their 2019 10k, Visa reported \$6.2 billion in client incentives to issuers and acquirors on a total of \$29.2 billion in gross revenues.

(i.e. surcharging) is small if consumer payment choice is inelastic. Starting from the optimal surcharge, the reduction in profit from charging uniform prices is second order in the transaction fees. Potential first order costs to surcharging such as menu costs or reputational costs could then overwhelm the benefits of surcharging[6]. Even though the network level consequences of surcharges are large, no individual merchant can influence consumers' adoption decisions. Therefore no one merchant can realize large gains from surcharging.

3 Intuition for the Results

This section explains why entry of new payment platforms can lead to higher merchant fees and lower consumer surplus. In a typical one sided market, entry reduces prices as firms cut margins in order to steal business from competitors. Consumers benefit from lower prices. Payment markets differ because they are two sided. Platforms set prices to balance consumer and merchant interests. If entry increases competition over consumers more than it increases competition over merchants, platforms respond by *raising* merchant fees to fund larger consumer rewards. Consumers' adoption of payment methods with high merchant fees inflates retail prices, potentially reducing consumer surplus in equilibrium.

The solid lines in panel A of figure 2 shows how a monopoly nonprofit output maximizing network sets fees by balancing consumer and merchant interests. While my model will feature profit maximizing networks, output maximizing networks' incentives are easier to visualize. The network has one control variable – the merchant fee – and rebates all profits to consumers to subsidize adoption. The left graph shows that consumer adoption increases in the merchant fee due to higher subsidies. The middle graph shows that higher merchant fees reduce merchant acceptance. The right graph shows payment volume as a function of merchant fees. Starting from a low level of fees, volume is increasing in fees as more consumers join the platform. At high fees, volume decreases as merchants leave. The network therefore chooses a fee at the peak of the solid green curve.

The dotted lines in panels A and B show how entry can have an ambiguous effect on merchant fees. Panel A shows a case where entry leads to higher merchant fees. The solid lines represent the incumbent's curves without entry, and the dotted lines repre-

[6]Consider a CES merchant facing a demand curve with an elasticity of $\sigma = 6$. Suppose $1/5$ of its customers use cash with a cost of 0.3% and the remainder use a card with an average cost of 1.5%. The loss from not charging payment method specific prices is 7 bps of profit.

sent the incumbent's curves after entry. In this example, entry causes the incumbent's consumer adoption curve to become lower and steeper. Fewer consumers join the incumbent at every level of fee, and the introduction of a substitute good means consumer adoption is more sensitive to rewards. In the middle panel, the new merchant acceptance line for the incumbent is now also lower and steeper, but crucially the change in the slope is not as dramatic as the change in the consumers' slope. The change might be small because merchant demand is price insensitive or because the new network is not a good substitute for merchants. The changes to consumer and merchant demand result in a total transactions curve that is both lower but with a peak that is to the right. The network charges a higher merchant fee to maximize output. Panel B illustrates how merchant fees fall in the opposite case when merchant demand changes more than consumer demand.

Edelman and Wright (2015) show that the increase in fees and rewards can reduce total consumer surplus. Figure 3 illustrates the intuition. While my model features debit, credit, and cash consumers, here I show a simplified world with three types of consumers: those who always use cash, those who always use card, and those who would be willing to switch from cash to cards if rewards are high enough. Higher fees that fund higher rewards have three effects. First, merchants pass on the higher fees to prices, hurting all consumers. Second, higher rewards benefit both the consumers who always use cards and those who switch. Third, switching consumers lose non-pecuniary utility. Because cards pay rewards but cash does not, then the consumers who switched must have valued the non-price characteristics of using cash higher than the non-pecuniary characteristics of cards. If the increase in rewards and merchant fees exactly offset each other, the total change in consumer surplus reflects only the decline in non-pecuniary utility. In practice, rewards and merchant fees will not exactly offset and my model will allow me to compute how each terms changes in equilibrium.

The market failure from payment network competition is not market power, but rather externalities. Consumers do not internalize the effect of their own payment choice on the aggregate price level, leading to a coordination failure (Edelman and Wright, 2015). While consumers would collectively prefer a world with low card use and low retail prices, each consumer has private incentives to adopt the high fee, high reward card. Networks exploit this coordination failure to encourage excessive adoption. In contrast to traditional antitrust cases where the problem is of high prices and low output, the problem in payments is that output and merchant prices are both too high. Whereas price elastic consumers mitigate market power, they exacerbate externalities by adopting payment methods with high merchant fees.

My estimated model recovers the empirical analogues of figures 2 and 3. Whether or not merchant fees increase with entry depends on the level and slopes of the consumer adoption and merchant acceptance curves, as well as how the curves change with entry. The welfare effects will depend on the distribution of consumers' non-pecuniary preferences over payment instruments and the model's prediction of how profit maximizing networks adjust fees and rewards. Identifying the shape of consumer and merchant demand curves for payments as well as the distribution of preferences will be the focus of my estimation in the sections to follow.

4 Data

I combine bank level data from a payments trade journal, the Nilson Report, with consumer level data from the Nielsen Homescan panel and the Federal Reserve's Diaries and Surveys of Consumer Payment Choice. These data provide key moments for estimating consumer and merchant demand for payments.

4.1 Issuer Payment Volumes

I construct an imbalanced panel of issuer level payment volumes from the Nilson Report. This panel tracks the effects of a regulatory shock to interchange fees on subsequent usage. The Nilson Report publishes the dollar payment volumes of the top credit and debit card issuers every year. These issuers include both banks and large credit unions.

I merge the issuer panel from the Nilson Report with call report data to identify which issuers were affected by the Durbin Amendment. In particular, for credit unions I use the NCUA credit union call reports and for the banks I use bank call reports from the FFIEC. I also drop any institution that experiences a large merger, which I define as a merger for which the new equity acquired is more than 50% of the pre-merger book equity.

My main difference in difference analysis focuses on a subset of 39 issuers, 19 of them above the Durbin cutoff and 20 below. Table 1 reports the main summary statistics for this sample. These issuers have assets in 2011 between $2.5 billion and $200 billion. The smallest issuers are small regional credit unions like the Pennsylvania State Employees' Credit Union, while the largest of these issuers are regional banks like Suntrust Bank or Fifth Third Bank. Signature debit transactions cover around two thirds of the total signature debit plus credit transactions.

4.2 Consumer Payment Surveys

I combine the Atlanta Federal Reserve's Diary of Consumer Payment Choice and Survey of Consumer Payment Choice to build a transaction level dataset on consumers' payment choices over three day windows. I use the data from the $2015 - 2020$ waves of both surveys for my main sample, although to study credit versus debit acceptance I also use data from the $2008 - 2014$ waves of the SCPC. This data is both useful in establishing basic facts about how consumers use different payment methods as well as establishing the relationship between consumer demand for payments and merchants' acceptance policies.

Table 2 reports the key summary statistics for the transactions in the dataset, whereas table 3 shows the demographic statistics for consumers with different payment preferences. I focus on non-bill, in person purchase payments in consumer retail and service sectors with ticket sizes of less than 100 dollars for most of my analysis. Around 42% of the payments are made in cash, with the remainder split relatively evenly over credit and debit. Most transactions (95%) are at a merchant who accepts card.

Consumers display a wide range of different payment behaviors. Households are spread out over four categories: cash, debit, credit, and mixed households who use both credit and debit cards in large amounts. Credit households' annual incomes are roughly twice that of cash and debit households. Mixed households have incomes between the debit and credit households.

Credit cards are widely available and frequently pay rewards. Around 73% of the cash and credit consumers own a credit card, and 46% of cash and debit consumers earn rewards on their credit cards. Around 80% of consumers who use a mix of debit and credit have a rewards credit card.

4.3 Nielsen Homescan Panel

The Nielsen Homescan tracks the method of payment of around 90,000 households at large consumer packaged goods stores. I use this to build measures of primary and secondary cards at the consumer level. Table 4 reports the main summary statistics. I focus on a sample of households that has a low share of missing payment data. The average household is in the sample for 3 years and records 500 transactions. Around a quarter of the transactions are on cash. Of the card transactions, slightly more than half are on debit and the remainder are on credit cards. I drop Discover transactions since it is a small share of aggregate payment volumes documented in Nilson.

The main shortcoming of the Homescan panel is that it cannot cover certain spending categories, such as travel or restaurants, that tend to have a high prevalence of card use. Appendix table B.3 shows the aggregate payment shares I derive from the Homescan panel as compared to tables from the Nilson Report on aggregate payment shares. Nielsen overrepresents cash and debit transactions while underrepresenting American Express. This is reasonable given Nielsen's sector composition.

5 Reduced Form Analysis

The reduced form analysis motivates a model of network competition where rewards influence consumer payment adoption and merchants adopt payment methods to increase sales. The magnitudes will also discipline the model estimates of consumer and merchant demand over payments. Throughout I will relate the reduced form facts to the shapes of consumer and merchant demand curves in figure 2, and how the curves change with competition. The facts point towards a world where new entry likely increases competition over consumers more than it increases competition over merchants, leading to higher merchant fees and consumer subsidies.

5.1 Consumer Substitution Between Credit and Debit

A regulatory shock that reduced debit interchange rates, thereby ending debit rewards, led to large reallocation of spending from debit to credit. This motivates why a model of competition in two sided markets is important, since networks will need to balance consumers' sensitivity to rewards against merchants' sensitivity to fees. A high consumer price sensitivity also corresponds to a steep consumer demand curve in the intuition diagram in figure 2.

Ideal variation would shock rewards at some banks and then to study how consumers change their payments. As an approximation to the ideal variation, I exploit the caps on debit card interchange fee introduced as part of the 2010 Dodd Frank Financial Reform Act[7]. Illinois Senator Richard Durbin inserted an amendment into the bill to cap debit interchange at large banks and credit unions[8] with more than $10 billion in assets to 22 cents and 0.05% of the purchase value. Credit interchange was untouched. By changing

[7]While there have been a few empirical papers on the effects of interchange fee regulation (Chang et al., 2005; Valverde et al., 2016), these papers cannot identify consumer preferences because of potential merchant responses. It is important that the shock to interchange is small because the goal is to isolate consumer preferences holding fixed merchant adoption. In models of interchange, the level of the interchange fee should affect both consumer utilization and merchant adoption (Rochet and Tirole, 2002).

[8]While the regulation covered both banks and credit unions, for the rest of the discussion I will refer to these financial institutions as simply "banks".

banks' income from debit card issuance, this led to a change in rewards. All large banks ended their debit reward programs, while small banks largely kept their rewards programs intact (Schneider and Borra, 2015; Orem, 2016).

To study the effect of the Durbin Amendment on payment volumes, I employ a difference in difference approach that compares payment volumes at large banks versus small banks around the time the Durbin Amendment was implemented. I estimate the regression

$$y_{it} = \sum_{k=-3}^{3} \beta_k I\{t = k\} \times \text{Treated}_i + \delta_i + \delta_t + \epsilon_{it} \tag{1}$$

where y_{it} is the log signature debit or credit card payment volume at a bank. Treated_i refers to whether the bank had more than \$10 billion in assets in 2010, and δ_i and δ_t represent bank and year fixed effects, respectively. By comparing large versus small banks I am able to difference out the effects of the CARD act or the effect of the Durbin Amendment on debit routing. I define time relative to 2011 and use 2010 as my base year. Figure 4 shows that signature debit volumes fell by -27 percent whereas credit card volume rose by 35 percent. Volume largely substituted between payment forms, as overall card spending fell by a small but statistically insignificant -5 percent. Appendix table B.4 reports the exact coefficients and standard errors.

The Durbin Amendment evidence suggests that consumers are sensitive to rewards. Hayashi (2012) estimates that the average debit rewards program paid consumers around 25 bps of transaction value, yet even that small change led to a -27% decline in signature debit volumes. Consumers' high sensitivity to rewards may be a surprise given consumers' general price insensitivity in other household finance settings. Part of this may reflect that banks tend to make rewards salient in advertising (Ru and Schoar, 2020). My estimated results are also consistent with Mukharlyamov and Sarin (2022) who find that geographic areas that were more affected by the Durbin interchange caps also saw larger increases in credit card volumes.

In the appendix I include additional robustness checks. Figure C.1 shows that the two groups of banks did not suffer large relative shocks in their asset values at the same time. Figure C.2 shows that the overall debit cards, which included PIN debit cards that were not affected by the regulation, declined less[9]. The differential pattern across debit cards suggests the effect I am identifying is about relative prices for credit and debit, and not just other shocks to big and small banks during this time period.

[9] Besides signature debit, many banks offered PIN debit. PIN debit was not affected by Durbin since the interchange rates were already low (Hayashi, 2012).

5.2 Consumer Substitution Between Networks

Data on consumers' primary and secondary cards shows that credit cards from different networks are good substitutes for each other. Therefore entry of a new credit network should have a large effect on consumer demand. The incumbent's consumer demand curve in figure 2 should shift down by a large amount in response to entry.

I identify substitution patterns by studying co-holding data. Amex users often carry Visa credit cards as a backup card for when Amex is not accepted. I infer from this fact that primary Amex users are therefore likely to use Visa credit cards in an alternative world without Amex. Formally, I interpret data on consumers' primary and secondary cards as data on hypothetical first and second choices. By identifying first and second choices, I can then use standard micro BLP techniques to identify the strength of segmentation between different types of cards (Berry et al., 2004).

To interpret primary and secondary cards as first and second choices, I need to assume that consumers ignore complementarities or substitution effects in deciding which pair of cards to put in their wallet. For example, two credit cards in different rewards categories (e.g. grocery + travel) could be complements. I ignore this because complementarities are small at the network level. Visa, Mastercard, and Amex all offer a robust suite of rewards credit cards across various categories. Rewards thus no longer create strong complementarities across networks. Cards could also be substitutes, as a credit card user does not bother to get a second card that offers a similar service. I ignore this possibility and therefore underestimate the extent to which consumers view credit cards from different networks as substitutes. Appendix F derives a dynamic microfoundation that features no complementarities or substitution effects. Under the microfoundation, the stationary distribution of primary and secondary cards (viewed as a Markov Chain) corresponds exactly to the joint distribution of first and second choices among primary payment methods.

From the Homescan shopping data, I define each consumers' primary payment card as the card network that is used for the highest share of trips[10]. Table 5 shows both the aggregate shares of each primary payment method and the conditional probability of each payment option occurring as the second choice. The bottom row of the table shows that debit cards are the most popular primary payment method, followed by cash, then Visa credit cards, MC credit cards, and lastly Amex. If credit cards, debit cards, and cash were all equally good substitutes for each other, then we should expect Amex users' second choices to mostly be debit cards and cash with a small share of Visa and MC. In

[10] In Appendix table B.6, I show that the total number of trips is highly correlated with the card with the highest share of spending.

reality, Amex users are more likely to have a Visa or a MC than they are to have a debit card, even though debit cards are dramatically more popular in the general population. Similar patterns are present for Visa and MC users.

5.3 Merchant Benefits from Accepting All Cards

The average merchant's sales increase around 30% from card acceptance yet not all merchants choose to accept cards. The large average sales benefit combined with less than universal card acceptance motivates a model where merchants accept cards to increase sales, but merchants differ in how much their sales increase. The large benefit relative to the level of fees also suggests that merchant demand is not price sensitive, so that the merchant demand curve in figure 3 should be flat.

I exploit variation in consumer payment preferences to identify how much merchants' sales increase from card acceptance. Ideal variation shocks merchant adoption of payment methods and measures the effect on sales[11]. As an alternative, I hold fixed merchant adoption but assume that variation in payment preferences among consumers is random. If card acceptance increases sales, then, relative to cash consumers, card consumers should transact more at merchants who accept card. This approach requires card consumers have similar baseline preferences over merchants as cash consumers. If card consumers travel more, and payment convenience is more important when one is traveling, then my approach will overestimate merchants' benefit from card acceptance. I try to adjust for these differences by saturating my regressions with fixed effects, but inevitably there are confounds. The main benefit is that by using the survey data I can provide an integrated view of how the US payment market works.

I use a logistic regression to measure how much sales increase from card acceptance. Index consumers by i and transactions by t. Let y_{it} be an indicator for whether the transaction t occurred at a store that accepts cards. Let X_{it} be an indicator of whether the consumer prefers cards. Let δ_{it} be a vector of fixed effects such as the survey respondent's income and education, the ticket size of the transaction, and the merchant category (e.g. restaurant versus retail). I run the logistic regression

$$P\left(y_{it}=1\right) \sim \phi X_i + \delta_{it} + \epsilon_{it} \tag{2}$$

The coefficient ϕ can be interpreted as the average increase in sales experienced by the merchants who accept card, netting both the positive effect from increased convenience

[11] Higgins (2020) and Berg et al. (2022) use this ideal variation to show that payment adoption can generate $20-50\%$ increases in sales.

and the negative effect of higher fees that get passed through to higher prices.

Table 6 shows the result of this regression with different options for fixed effects. My preferred estimate includes both the consumer and merchant controls, and suggests that the average consumer who prefers cards is around 30% more likely to shop at a store that accepts cards than a consumer who does not prefer to use cards. I interpret this as saying merchants are more likely to attract card consumers if they accept cards. The relative stability of the results even as I adjust the consumer and transaction fixed effects suggests there is little unobserved variation driving the result. Appendix table B.5 shows that this effect does not vary much across debit versus credit card users, those who hold one or multiple cards, or high or low income respondents.

Most but not all consumer spending is done at stores that accept cards. Therefore even though most merchants gain a large amount from accepting cards, some merchants do not. From the diary data I see that consumers who prefer to use cards spend around 97% of their money at stores that accept cards.

5.4 Merchant Substitution Between Networks

Merchants do not view credit card and debit card acceptance as substitutes, and consumer holding data suggests that different credit card networks are imperfect substitutes for each other. Therefore a new credit entrant should cause the merchant demand curve in figure 2 to only fall a small amount.

I use a large change in the cost of debit versus credit acceptance to test merchant substitution patterns between debit and credit. Intuitively, two goods are good substitutes if changes in relative prices induce large changes in relative quantities. Yet when the Durbin Amendment cut debit card fees, there was no significant change in the number of merchants that accepted credit cards. Figure 5 plots of credit and debit card fees around the implementation of the Durbin Amendment and survey measures of credit and debit acceptance around the same time. The left panel shows that the cost of debit acceptance fell by half and the cost of credit card acceptance continued to rise. Yet consumer ratings of debit and credit card acceptance were unchanged.

One reason debit acceptance may not substitute for credit acceptance is because the consumers who use both a debit and credit card use them for different purposes[12]. For example, a consumer might use a debit card when she has money in her checking ac-

[12]This is the logic behind defining credit cards as a separate market from debit cards in past antitrust cases. In footnote 8 of the decision in the US v. VISA USA, Judge Jones argues for a separate credit card market based on Visa and Mastercard analysis showing that possession of debit cards did not reduce credit card spending. Jones (2001)

count, but then switch to a credit card when she does not. Therefore when the consumer wants to use the credit card, using a debit card is not an acceptable substitute. Table 3 shows that when compared to consumers who only pay with credit cards, consumers who use a mix are more likely to carry balances. Another reason for the lack of substitution is if merchants have to accept Visa credit and debit together. However, there is no longer any legal requirement to do so[13]. Such a theory would also counterfactually predict that Amex acceptance should drop in response to Durbin.

Accepting one network's credit cards is an imperfect substitute for accepting a different network's credit cards because around 40% of consumers carry a card from only one of the three major networks (Visa, MC, Amex). The literature refers to consumers who carry multiple cards as multihomers and those who carry only one as singlehomers. Multihomers empower merchants to reject high fee cards. If every Visa consumer owns a MC, then accepting MC also enables merchants to serve customers with a Visa card. Visa and MC would compete down merchant fees since merchants would only accept the cheapest option. In contrast, if consumers singlehome, then Visa can charge merchants high prices for exclusive access to Visa customers. Table 7 compares my estimates of the probability that a credit card consumer singlehomes. Across both Homescan and the DCPC I find that 40% of credit card consumers use only credit cards from one network. This is somewhat lower than the data in Rysman (2007), and is consistent with the general growth of the card industry in the past 15 years.

6 Model

The model maps reduced form facts into predictions for how networks compete. Once I estimate the parameters, solving the game under different conditions will enable me to predict network behavior under different market structures and decompose the welfare effects of changes in competition and regulation. A model is necessary because I am interested in studying market structures that have not yet occurred and because many of the essential cross price elasticities on the merchant side are hard to measure with only reduced form evidence.

[13]Before a 2003 court settlement, Visa and Mastercard did have "honor all cards" rules that tied the acceptance of debit and credit cards. However, the 2003 settlement dropped these rules and in the following years Visa and Mastercard cut interchange fees on debit cards (Constantine, 2012). Therefore there was no formal requirement that prevented merchants from substituting from credit to debit acceptance.

6.1 Structure of the Game

I model competition between card networks as a static game with three stages with three kinds of players: networks, consumers, and merchants. I solve for a subgame perfect equilibrium of this game.

In the first stage, profit maximizing networks set per transaction fees for merchants and promised utility levels for consumers.

In the second stage, consumers and merchants make adoption and pricing decisions. Consumers choose up to two cards to put in their wallet, and are more likely to choose those with high promised utility. Merchants choose which cards to accept and set prices. Merchants are more willing to accept payment methods many consumers use. Merchants pass on transaction fees to higher prices.

In the third stage, consumers decide how much to consume from each merchant and pay with the cards in their wallet. Consumers are more likely to consume at merchants that accept their preferred payment methods and charge low prices. Below, I walk through the stages of the game in reverse order.

6.2 Payment Choice at the Point of Sale

At the point of sale, consumers will first try to use their primary card, then their secondary card if it is the same type as their primary card, and otherwise cash.

Define the set of all inside payment methods (i.e. cards) as $\mathcal{J}_1 = \{1, \ldots, J\}$, and the set of all payment methods as $\mathcal{J} = \{0\} \cup \mathcal{J}_1$, where 0 refers to cash. Although I have in mind a fintech platform entering in the counterfactual, for simplicity I will refer to all inside payment methods as cards. Each payment method has a type, $\chi^j \in \{C, D, R, A\}$ for cash, debit, credit, or an app.

Each consumer has a wallet w with zero, one or two cards[14] that they chose at an earlier stage of the game. Let $\mathcal{W} = \{(j,k) : j,k \in \mathcal{J}, j \neq k\}$ denote the set of all possible wallets. For a wallet $w = (w_1, w_2)$, the term w_1 is the primary payment method and w_2 is the secondary payment method.

If the merchant accepts M, then an indicator for whether a wallet w consumer pays

[14] I choose two cards because two networks covers 95% of an average consumer's card spending (see table B.7).

with payment method j can be defined as

$$I_{j,M}^{w} = \begin{cases} 1 & w_1 = j, j \in M \\ 1 & w_1 \neq j, w_2 = j, \chi^{w_1} = \chi^{w_2}, j \in M \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

where $I_{j,M}^{w} = 1 \iff$ wallet w pays with j at a merchant that accepts M. Note consumers only pay with their secondary card if the primary card is not accepted, the secondary card is, and the secondary card is the same type as their primary card.

By modeling wallets, I am able to unify cash consumers, consumers who use only one card, and consumers who use multiple cards under one framework. Diagrams of how different types of consumers pay are shown in figure 6. A cash only consumer's primary payment method is cash, $w_1 = 0$. A consumer who only carries a Visa has $w_1 = \text{Visa}$ but $w_2 = \text{Cash}$. A consumer who carries an Amex as their primary card and a Visa as a backup has $w_1 = \text{Amex}$, $w_2 = \text{Visa}$. Note that the Amex + Debit consumer either pays with Amex or cash, skipping over the debit card. This occurs because Amex and debit cards are different types of payment. Imposing that consumers only substitute within types of payment explains why merchants do not treat debit and credit acceptance as substitutes.

Rewards do not affect consumer behavior at the point of sale. Arango et al. (2015) show that conditional on the cards a consumer has in their wallet, rewards do not have a large effect on payment choice. In the model, high reward cards are more likely to become primary cards, but conditional on being the primary card rewards have no effect on the intensive margin.

6.3 Consumer Consumption Decisions Over Merchants

Consumers value both card acceptance and low prices. Fix a consumer who has chosen a wallet w at an earlier stage. Each merchant ω has already decided to accept cards $M^*(\omega) \subset \mathcal{J}_1$ and has set prices $p^*(\omega)$. Define a card payment indicator as

$$v_M^w = \begin{cases} 1 & \text{Consumer } w \text{ pays with a card if merchant accepts } M \\ 0 & \text{Otherwise} \end{cases}$$

This term is therefore 1 for a consumer with two cards of the same type provided that one of their cards is accepted.

The consumer has symmetric CES preferences over merchants, where payment acceptance enters into quality. Each merchant is characterized by a type $\gamma(\omega) \geq 0$ that determines the importance of payment availability for consumer shopping behavior at the merchant. Let the elasticity of substitution be σ. The consumer has income y^w. The consumer chooses a consumption vector $q^w(\omega)$ to solve

$$U^w = \max_{q^w} \left(\int_0^1 \left(1 + \gamma(\omega) v^w_{M^*(\omega)}\right)^{\frac{1}{\sigma}} q^w(\omega)^{\frac{\sigma-1}{\sigma}} \, d\omega \right)^{\frac{\sigma}{\sigma-1}}$$

$$\text{s.t.} \int_0^1 q^w(\omega) p^*(\omega) \, d\omega \leq y^w$$

Standard CES results mean that the quantity consumed at a merchant ω depends on the type γ, the price p, the payments accepted M, income y^w, and an aggregate price index P^w that summarizes the impact of the pricing and adoption decisions of all other merchants. The demand from a consumer with wallet w for a merchant of type γ is:

$$q^w(\gamma, p, M, y^w, P^w) = (1 + \gamma v^w_M) p^{-\sigma} \frac{y^w}{(P^w)^{1-\sigma}} \tag{4}$$

$$(P^w)^{1-\sigma} = \int \left(1 + \gamma(\omega) v^w_{M^*(\omega)}\right) p^*(\omega)^{1-\sigma} \, d\omega$$

In this demand curve, only γ, v^w_M, and p vary across merchants. The price index P^w and the income y^w are not affected by any one merchant's actions.

The merchant's γ parameter determines the percentage increase in sales from a card consumer who originally had to pay in cash, but who can now pay with a card. A low γ firm might be a small business with a loyal customer base, for whom the method of payment is not important. A high γ firm could be an e-commerce firm, who can benefit from significantly higher sales if the online checkout process is convenient (Berg et al., 2022).

I assume consumers only care about *whether* they use a card from their wallet and not about which card is used. This allows me to use the share of singlehoming credit card consumers in the data to discipline how much demand merchants lose if they were to stop accepting one credit card network while continuing to accept others.

The CES assumption underpins my welfare analysis. I infer from card consumers' higher consumption at merchants who accept card as indicating consumer utility goes up from card acceptance. CES provides a disciplined framework for adding up the utility benefits across merchants to arrive at an aggregate change in consumer welfare.

Two merchants with the same γ will choose the same price and acceptance policy.

Therefore the merchant variety ω can be dropped from the analysis. I can describe the equilibrium in terms of a equilibrium price schedule $p^*(\gamma)$ and a set valued adoption schedule $M^*(\gamma)$. This reparameterization means that the price index can now be expressed as

$$(P^w)^{1-\sigma} = \int (1 + \gamma v^w (M^*(\gamma))) \, p^*(\gamma)^{1-\sigma} \, \mathrm{d}G(\gamma) \tag{5}$$

where $G(\gamma)$ is the distribution of the γ parameter across merchants.

In equilibrium, consumers will consume according to a consumption schedule $q^{w*}(\gamma)$ for each merchant type γ that satisfies

$$q^{w*}(\gamma) = q^w(\gamma, p^*(\gamma), M^*(\gamma), y^w, P^w) \tag{6}$$

6.4 Merchant Pricing

Merchants are single product firms that maximize profits by setting prices and choosing the optimal subset of payments to accept.

I first solve for optimal pricing conditional on a given acceptance decision M. Collapse the wallet specific price indices from the consumer problem to $P = (P^w)_{w \in \mathcal{W}}$. Let the merchant fee for payment method j equal τ_j of sales. Let the share of consumers with wallet w be $\tilde{\mu}^w$. This should be thought of as the share of dollars in the economy in a wallet of type w. Normalize the firm's marginal costs to 1.

The profit function as a function of that one merchant's price is

$$\Pi(p, \gamma, M, P, \tau) = \sum_{w \in \mathcal{W}} \tilde{\mu}^w \left[\underbrace{q^w p (1 - \tau^w_M)}_{\text{Net Revenue}} - \underbrace{q^w}_{\text{Costs}} \right] \tag{7}$$

Where the fee τ^w_M for wallet $w = (w_1, w_2)$ is the fee of the payment method that is finally used. Formally it is $\tau^w_M = \sum_{j \in \mathcal{J}} I^w_{j,M} \tau_j$, where the indicators $I^w_{j,M}$ were defined in equation 3 and pick up whether payment method j was used.

The expression for profit in equation 7 is a wallet weighted average of revenues less costs. The first term is revenue net of card fees. The second captures the costs of production, which have been normalized to 1. The merchant's optimal pricing problem is

$$\hat{p} = \underset{p}{\operatorname{argmax}} \, \Pi(p, \gamma, M, P, \tau) \tag{8}$$

The optimal price passes on the average transaction fee to the consumer.

$$\implies \hat{p} = \frac{\sigma}{\sigma - 1} \times \frac{1}{1 - \hat{\tau}} \tag{9}$$

$$\hat{\tau}\left(\gamma, M, P^*, \tau\right) = \frac{\sum_{w \in \mathcal{W}} \tilde{\mu}^w \frac{y^w}{\left(P^w\right)^{1-\sigma}} \left(1 + \gamma v^w\right) \tau^w}{\sum_{w \in \mathcal{W}} \tilde{\mu}^w \frac{y^w}{\left(P^w\right)^{1-\sigma}} \left(1 + \gamma v^w\right)} \tag{10}$$

When payment methods have no fees, the above formula becomes the standard CES optimal pricing equation where prices are equal to $\frac{\sigma}{\sigma-1}$ times marginal costs. When transaction fees τ_j are positive, prices are inflated relative to the standard CES benchmark by $\left(1 - \hat{\tau}\right)^{-1}$. The realized transaction fee $\hat{\tau}$ captures the average transaction fee the merchant pays, averaging across all the consumers it serves[15].

In equilibrium, merchants set optimal prices given the optimal pricing and adoption strategies of other merchants.

$$\hat{p}\left(\gamma, M^*\left(\gamma\right), P, \tau\right) = p^*\left(\gamma\right) \tag{11}$$

6.5 Merchant Acceptance

Merchants also choose the profit maximizing bundle of payments to accept. Define the profit function for a particular bundle of payments $M \in 2^{\mathcal{J}_1}$, taking into account the merchant's optimal pricing policy, as

$$\widehat{\Pi}\left(\gamma, M, P, \tau\right) = \Pi\left(\hat{p}, \gamma, M, P, \tau\right)$$

Merchants then solve the profit maximization problem

$$\widehat{M}\left(\gamma, P, \tau\right) = \underset{M \in 2^{\mathcal{J}_1}}{\operatorname{argmax}} \hat{\Pi}\left(\gamma, M, P, \tau\right) \tag{12}$$

Where the price index P already depends on the equilibrium pricing p^* and adoption decisions M^*. In equilibrium, merchants adopt optimal bundles holding fixed the optimal

[15]The pricing formula 9 states that fees are passed through to prices more than one for one. The high rate of pass through reflects a combination of market power and log-convex demand (Weyl and Fabinger, 2013; Pless and van Benthem, 2019).. Empirical work on the passthrough of card transaction fees suggests that merchants tend not to change prices in response to transaction fees in the short run (Higgins, 2020). A model with high passthrough is reasonable for long run analysis since menu costs might prevent short run adjustment while allowing for long run passthrough. This issue is also unlikely to be resolved empirically, as tests of the long run response of prices to transaction fees lack power.

adoption and pricing behavior of other merchants.

$$\widehat{M}(\gamma, P, \tau) = M^*(\gamma) \tag{13}$$

Merchants are able to accept any subset of cards, including subsets that are not played on the equilibrium path. This is essential to discipline networks' incentives to raise merchant fees. In the current US market, almost all merchants who accept Visa credit cards also accept MC credit cards. That is rational in the current equilibrium where both networks charge similar fees. However, it's essential that merchants are able to drop Visa credit while still accepting MC credit in an alternative world where Visa were more expensive than MC. Otherwise, Visa would face strong incentives to raise merchant fees since doing so would not have a large effect on merchant acceptance.

6.6 Consumer Adoption

Consumers put the two payment instruments that offer the highest payment utility to put in their wallet. If a consumers top two options are cards, the consumer will also carry cash.

I define the log payment utility V_i^j from a single payment method $j \in \mathcal{J}$ as

$$\log V_i^j = \underbrace{\log U^j}_{\text{CES}} + \underbrace{\Xi^j}_{\text{Intercepts}} + \frac{1}{\alpha}\left(\underbrace{\eta_{ij}}_{\text{Unobs Char}} + \underbrace{\beta_i X^j}_{\text{R.C.}}\right) \tag{14}$$

Each component of payment utility plays an important role. The first term is the CES utility, U^j. This combines the consumer's utility from rewards with the gains from card acceptance. If the network pays consumers who singlehome on card j a subsidy f^j, then the consumer's log optimal utility from solving her consumption problem is U^j where

$$U^j = \frac{1 + f^j}{P^j} \implies \log U^j \approx f^j - \log P^j \tag{15}$$

where P^j is the CES price index associated with a customer who only uses j. The utility from the CES system increases for a payment method that earns a large subsidy (which increases the numerator) and also increases for a payment method that is widely accepted (which decreases the denominator[16]). The variables Ξ^j represent card specific intercepts that rationalize market shares. I normalize $\Xi^0 = 0$. The parameter α is a

[16] A more widely accepted payment instrument will have a lower price index by equation 5

measure of how elastic consumers are to increases in the subsidy. A large value of α will mean that a small increase in f^j leads to a large increase in j's market share. The shocks η_{ij} represent unobserved characteristics of payment methods. The characteristics X^j are indicators for whether a payment method is a card or cash and whether it has a credit function. The random coefficients are distributed $\beta_i \sim N(0, \Sigma)$ for some covariance matrix Σ. This unobserved heterogeneity across consumers generates rich substitution patterns between credit and debit cards and between cards and cash.

The share of demand at each merchant from consumers of each kind of wallet is pinned down by the joint distribution of the largest and second largest values of V_i^j. The payment method with the highest utility becomes the primary payment and the second highest utility becomes the secondary payment in the wallet. I define *insulated* market shares for the wallet $w = (l, k)$ as

$$\mu^{(l,k)} = P\left(\left(V_i^l = \max_{j \in \mathcal{J}} V_i^j\right) \cap \left(V_i^k = \max_{j \in \mathcal{J} \setminus \{l\}} V_i^j\right)\right) \tag{16}$$

These shares add up to 1 and represent the share of demand for a cash only merchant from consumers of each wallet type.

Market shares $\tilde{\mu}$ are constructed so that each merchant's decision on which cards to accept will only depend on the insulated shares μ, and not on the underlying price index P^w or the subsidies f^w. Actual market shares among consumers for different wallets are derived from the insulated shares as

$$\tilde{\mu}^{(l,k)} = \frac{1}{C} \frac{\mu^{(l,k)} \left(P^{(l,k)}\right)^{1-\sigma}}{1 + f^{(l,k)}} \tag{17}$$

$$C \equiv \sum_{w \in \mathcal{W}} \frac{\mu^w (P^w)^{1-\sigma}}{1 + f^w} \tag{18}$$

where f^w is the total subsidy paid[17] to a consumer with wallet w. The constant C has been defined in a way to make the market shares add up to 1. If one alternatively defined the market shares $\tilde{\mu}$ in terms of the joint distributions of the payment methods delivering the top two highest V_i^j, that would create a strategic substitutability where merchants are less likely to adopt payment methods if other merchants have already adopted. A

[17] The subsidies f^w for consumers who only hold one card will be set by the networks, and the subsidies for the consumers who multihome will be based off of the singlehoming subsidies under the assumption that the subsidy from a card is proportional to the amount of spending done on that card. I discuss the calculation of these subsidies in the next subsection.

pure strategy equiliubrium for consumer and merchant adoption may no longer exist.

6.7 Network Profits

Network profits come from transaction fees charged to merchants, less the subsidies paid to consumers. A useful quantity for computing profits is the total dollar amount $\tilde{d}_j^w$ consumers with wallet w spend on card j. This is

$$\tilde{d}_j^w = \tilde{\mu}^w \int I_{M(\gamma),j}^w q^w(\gamma)\, p(\gamma)\ \mathrm{d}G(\gamma)$$

where the indicator $I_{M,j}^w$ picks up whether payment method j was used if the merchant accepts M and the consumer has a wallet w defined in equation 3.

Total profits from the merchant side of the market for card j is

$$T_j = \sum_{w \in \mathcal{W}} \tilde{\mu}^w \tilde{d}_j^w \left(\tau_j - c_j\right)$$

where c_j is the cost of processing one dollar on method j. The total cost of subsidies is

$$S_j = \sum_{w \in \mathcal{W}} \tilde{\mu} f_j^w$$

where f_j^w is the amount of subsidies that need to be paid to consumer w for her use of j.

In the model, networks own multiple cards. For a network n that owns cards $\mathcal{O}_n \subset \mathcal{J}_1$, the profit is

$$\Psi_n = \sum_{j \in \mathcal{O}_n} \left(T_j - S_j\right) \tag{19}$$

I describe how to calculate each of the terms below. First, note that the total dollars can also be expressed in terms of insulated shares μ^w and a new expression for insulated dollars, d_j^w, that does not depend on the normalizing constant C.

$$\tilde{d}_j^w = \frac{\mu^w}{C} \underbrace{\int I_{M(\gamma),j}^w \left(1 + \gamma v_M^w\right) p(\gamma)^{1-\sigma}\ \mathrm{d}G(\gamma)}_{d_j^w}$$

The profits networks earn from merchants can then be re-expressed as

$$T_j = \frac{1}{C} \times \left(\tau_j - c_j\right) \sum_{w \in \mathcal{W}} d_j^w \tag{20}$$

To calculate the cost of subsidies to consumers, I assume that consumers receive subsidies according to a fixed percent of their equilibrium spending. This assumption is only relevant in the off equilibrium path where different networks that are substitutes for merchants charge different fees. If a singlehoming American Express user spends 50 cents on American Express and earns 2 cents in subsidies, and a singlehoming Visa user spends \$1 on Visa and earns 2 cents in subsidies, a multihoming consumer who spends 50 cents on Visa and 50 cents on American Express should earn 3 cents of subsidies.This assumption is equivalent to networks who pay lump sum subsidies according to knowledge of how much the consumer will spend in equilibrium.

To implement the above assumption, I calculate the total subsidies the card j pays out as

$$S_j = f^j \tilde{\mu}^j \left(\frac{\sum_{k \neq j} d_j^{(j,k)} + d_j^{(k,j)}}{d_j^{(j,0)}} \right) \tag{21}$$

where $\tilde{\mu}^j \equiv \mu^{(j,0)}$. Intuitively, the network must pay out $f^j \tilde{\mu}^j$ to the singlehoming agents. To compute how much needs to be paid out to the multihoming agents, I compute how many dollars the multihoming agents spend on j relative to the amount of dollars the singlehoming agents spend on j. It is possible for multihoming agents to spend on j only when j is in the wallet, and therefore the sum iterates over both kinds of wallets (j,k) and (k,j). The amount of the subsidy is then inflated by the ratio of the total dollars spent on j by all agents compared to the total dollars spent on j only by the singlehoming agents.

There is one last fixed point between the normalizing constant C, the actual market shares $\tilde{\mu}$, and the subsidies paid to each type of multihoming agent. To get around this issue, I make a simplifying assumption that, for the purpose of calculating network profits, the the multihoming agents can be assumed to receive the subsidy of their primary card. This is a small adjustment since it reflects a second order effect of differences in subsidies causing consumers to have differences in income, changing spending, and thus affecting transaction fee income. Thus I approximate the profits from merchants $\tilde{T}_j$ and the subsidy bill $\tilde{S}_j$ as

$$\begin{aligned}
\tilde{T}_j &= \frac{1}{\tilde{C}} \times \left(\tau_j - c_j\right) \sum_{w \in \mathcal{W}} d_j^w \\
\tilde{S}_j &= \frac{1}{\tilde{C}} f^j \frac{\mu^{(j,0)} \left(P^j\right)^{1-\sigma}}{1+f^j} \left(\frac{\sum_{k \neq j} d_j^{(j,k)} + d_j^{(k,j)}}{d_j^{(j,0)}} \right) \\
\tilde{C} &= \sum_{w=(w_1,w_2) \in \mathcal{W}} \frac{\mu^w \left(P^w\right)^{1-\sigma}}{1+f^{w_1}}
\end{aligned} \tag{22}$$

6.8 Solving the Merchant Adoption Subgame

While the profit maximization problem in equation 12 is conceptually clean, it will be both computationally easier and yield more economic intuition to have merchants adopt bundles M that maximize a linear approximation of the profit function $\widehat{\Pi}$, which I will call quasiprofits $\overline{\Pi}$. As long as fees are small, true profits $\widehat{\Pi}$ will be approximately equal to a linear function of γ with weights and slopes that have an intuitive meaning. I formalize this in the theorem below.

Theorem 1. *True profits are approximately linear in γ. For any γ, M, P, τ,*

$$\widehat{\Pi} - \overline{\Pi} = O\left((1+\gamma) \left(\tau^{\max}\right)^2 \right)$$

where

$$\overline{\Pi}(\gamma, M, P, \tau) \equiv \frac{1}{C} \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \left\{ -a_M + b_M \gamma + \frac{1}{\sigma} \right\} \tag{23}$$

$$a_M = \sum_{w \in \mathcal{W}} \mu^w \tau_M^w$$

$$b_M = \frac{1}{\sigma} \sum_{w \in \mathcal{W}} \mu^w v_M^w \left(1 - \sigma \tau_M^w\right)$$

$$\tau^{\max} = \max_j \tau_j \tag{24}$$

Proof. See Appendix A ☐

The linear form of quasiprofits means the adoption equilibrium among merchants can be computed by solving for the upper envelope of a collection of linear functions. Payment bundles with a high fee have a low intercept and a flatter slope. Payment bundles that serve a large share of consumers have a steeper slope. The linear form of quasiprofits can be used to illustrate the similarities between my model of merchant adoption and that of Rochet and Tirole (2003). I elaborate on these issues in Appendix E.

6.9 Network Conduct and Equilibrium Determinacy

Networks maximize profits by adjusting promised CES utility levels for consumers U^j and transaction fees for merchants τ_j, holding fixed the utility levels and transaction fees from other networks. This conduct assumption is in line with the insulating tariffs framework of Weyl (2010) and guarantees that for every vector of network choice variables, the merchant and consumer subgame is unique.

One central challenge in modeling network competition is dealing with a potential multiplicity of equilibria due to a "chicken or egg" problem (Caillaud and Jullien, 2003; Chan, 2021). Consumers only adopt cards if they are widely accepted, and merchants only accept cards if there are enough consumers who want to use them. In monopoly settings, a common approach is to select the Pareto undominated equilibrium. This is a bad approach to modeling competition because Pareto dominance would not be able to select between competing networks.

To deal with the "chicken or egg" problem, I instead assume that networks promise CES utility levels for consumers. This is in line with the insulating tariffs outlined in Weyl (2010); White and Weyl (2016). Networks promise consumers if merchants do not adopt, the networks will compensate the consumers with higher subsidies. Consumers then each have a dominant strategy, and merchant actions are determined as soon as consumer actions are determined. Weyl (2010) argues that this is a reduced form way of capturing penetration pricing where networks subsidize consumer adoption when the network is small.

I am able to solve for the merchant and consumer subgame and network profits given the promised utility levels U^j and merchant fees τ set by the networks. The utility levels give insulated shares μ^w by equation 16. Merchant adoption then follows from solving for the upper envelope of quasiprofit functions from equation 23. The merchant adoption strategy yields the CES price indices P^w according to equation 5. The CES price indices combined with the CES utility levels U^j give implied subsidy levels f^j from equation 15. Equations 20, 21, and 19 then yield the network profits.

I employ a refinement to deal with the non-differentiability of network profits with respect to the merchant fees. The source of this non-differentiability comes from the Bertrand like assumption that merchants always accept the bundle of payments that delivers the highest profit. Rochet and Tirole (2003) do not encounter this issue in their symmetric, two network model, but subsequent work has shown that transaction volumes are generally non-differentiable in transaction fees when consumers can multi-home[18] (Liu et al., 2021). I assume that when each network chooses utility levels and transaction fees, it maximizes expected profits while assuming small trembles in the choice variables. Appendix D explains why this makes the profit function differentiable and how to efficiently calculate the derivative of the expectation.

I now have the tools to formally state the conduct assumption. For each network $n = 1, \ldots, N$, networks set promised utility levels U^{j*} and transaction fees τ_j^* for the cards that they own $\mathcal{O}_n$ such that

$$\left(U^{j*}, \tau_j^*\right)_{j \in \mathcal{O}_n} = \underset{\left(U^j, \tau_j\right)_{j \in \mathcal{O}_n}}{\operatorname{argmax}} \mathbb{E}\left[\Psi_n\left(\tilde{U}^j, \tilde{\tau}_j, \tilde{U}^{-j}, \tilde{\tau}_{-j}\right)\right] \tag{25}$$

$$\tilde{U}^j \sim N\left(U^j, \sigma^2\right) \text{iid}$$

$$\tilde{\tau}_j \sim N\left(\tau_j, \sigma^2\right) \text{iid} \tag{26}$$

where σ^2 is a small variance that I set to 10^{-10}, and U^{-j}, τ_{-j} capture all the singlehoming utilities and fees set by the other networks. I model cash as a network who sets fees to the cost of cash $\tau_j = c_0$ and sets a utility level U^0 equal to $1/P^0$ so as to not pay any subsidies.

6.10 Equilibrium

A full equilibrium is characterized by fees τ^*, CES utility levels U^*, insulated shares μ, a merchant pricing schedule $p^*(\gamma)$, a merchant adoption schedule $M^*(\gamma)$, and wallet specific consumer demand schedules $q^{w*}(\gamma)$ that satisfy five conditions.

1. The demand schedule $q^{w*}(\gamma)$ is optimal given each consumer's wallet choice, the network's choice of subsidy, and merchants' acceptance and pricing policies (Eqn 6).

2. For each merchant of type γ, she maximizes quasiprofits by accepting $M^*(\gamma)$ and sets the price $p^*(\gamma)$ (Eqn 11 + 13), holding fixed the adoption and pricing decisions of all other merchants, consumers' choices of wallets, and networks' choices.

[18]Starting from the symmetric equilibrium, a network that raises its merchant fee is now competing with the option to accept all other card networks. A network that cuts its merchant fee is now competing with cash. In these two regions, the marginal revenue from raising fees is very different, and therefore profits are not differentiable in the neighborhood of the original symmetric fee.

3. The insulated shares μ reflect consumers' optimal wallet choices, holding fixed the networks' promised utility levels (Eqn 16).

4. Networks are maximizing profits at the fees τ^* and promised utility levels U^*, holding fixed the promised utility levels and fees of other networks (Eqn 25)

5. Cash pays no subsidy and charges a fee τ_0 equal to the cost of cash c_0.

7 Estimation

I estimate the parameters of my model with indirect inference by matching the reduced form facts. By estimating the model I am able to make quantitative statements about how competition affects market outcomes. The key primitives to recover are (1) consumers' demand parameters over the different payment options, (2) the distribution of merchant types describing how much merchants benefit from payment acceptance, and (3) the networks' marginal cost parameters. I assume the aggregate shares and prices are the equilibrium of the model with three multicard payment networks – Visa, Mastercard, and Amex. Both Visa and Mastercard each own two cards (debit and credit) while Amex only owns their credit card network.

7.1 Consumer Substitution Patterns

I first estimate how consumers substitute between payment methods with different characteristics and how consumers respond to changes in rewards. Substitution and price sensitivity are governed by the distribution of random coefficients summarized in Σ and the price sensitivity parameter α. These parameters are identified by the reduced form facts on consumers' primary and secondary cards (section 5.2) and the effects of Durbin on payment volumes (section 5.1).

The covariance matrix Σ is important for matching the Homescan data on primary and secondary cards. Let cash be the outside option, and order the cards in Homescan as debit, Visa credit, MC credit, and Amex. For each possible wallet (j,k) where j is not cash, let $\hat{s}_{jk}$ be the estimated probability that a Homescan consumer is a primary j user and a secondary k user. Stack these shares as $\hat{s}$.

I match the empirical shares with a simple model where the utility from each option j for each consumer i is

$$u_{ij} = \delta_j + \eta_{ij} + \beta_i X^j \tag{27}$$

where $\delta_0 = 0$, $\beta_i \sim N(0, \Sigma)$ and $\eta_{ij} \sim$ T1EV. The characteristics X^j are indicators for whether the payment option is card or cash and whether it has a credit function. The mean utilities capture the equilibrium values of CES utility and intercepts from equation 14. Crucially the distribution of random coefficients Σ is the same as in the full model. The probability of a given combination of primary and secondary card is

$$s_{jk}(\Sigma, \delta) = \int \frac{\exp\left(\delta_j + \beta_i X^j\right)}{\sum_l \exp\left(\delta_l + \beta_i X^j\right)} \times \frac{\exp\left(\delta_k + \beta_i X^j\right)}{\sum_{l \neq j} \exp\left(\delta_l + \beta_i X^j\right)} \, \mathrm{d}H(\beta_i) \tag{28}$$

where H is the joint distribution of β_i. This formula follows Berry et al. (2004). I compute this with Monte Carlo integration. Stack the mean utilities as δ and the model implied shares as s.

The price sensitivity coefficient α is important for matching the effects of Durbin. From the Nilson panel, I estimate two micro-moments: the effect of the Durbin Amendment on signature debit volumes (Figure 4), and the share of signature debit card volumes of total signature debit + credit volumes (Table 1). I impose a third aggregate moment that 20% of overall transactions by value are done by cash (Table B.1). Combine these moments as $\hat{m}$.

Next I simulate Durbin in my model. The log payment utility V_i^j can be approximated as

$$\log V_i^j \approx f^j - \log P^j + \Xi^j + \frac{1}{\alpha}\left(\eta_{ij} + \beta_i X^j\right)$$

The payment utilities generate identical market shares as a simple discrete choice model where the utility for consumer i from option j at time $t \in \{\text{pre}, \text{post}\}$ is

$$v_{ijt} = \tilde{\delta}_j + \alpha f_t^j + \eta_{ij} + \beta_i X^j \tag{29}$$

where $\tilde{\delta}_j$ differ from δ_j, but the distribution of β_i is the same as above in equation 27. I use this simple model to generate the moments. I model the two cards seen in Nilson as signature debit and credit cards and cash is again the outside option. Let $\Delta f = 25$ bps, which is the change in debit rewards as a result of Durbin. The model implied moments are

$$m(\Sigma, \alpha, \phi) = \begin{pmatrix} \log \int \frac{\exp\left(\tilde{\delta}_1 - \alpha \Delta f + \beta_i X^1\right)}{\sum_k \exp\left(\tilde{\delta}_k - \alpha \Delta f I\{k=1\} + \beta_i X^k\right)} - \log \int \frac{\exp\left(\tilde{\delta}_1 + \beta_i X^1\right)}{\sum_k \exp\left(\tilde{\delta}_k + \beta_i X^k\right)} \\ \int \frac{\exp\left(\tilde{\delta}_1 + \beta_i X^1\right)}{\sum_k \exp\left(\tilde{\delta}_k + \beta_i X^k\right)} \times \left(\int \frac{\exp\left(\tilde{\delta}_1 + \beta_i X^1\right)}{\sum_k \exp\left(\tilde{\delta}_k + \beta_i X^k\right)} + \int \frac{\exp\left(\tilde{\delta}_2 + \beta_i X^2\right)}{\sum_k \exp\left(\tilde{\delta}_k + \beta_i X^k\right)}\right)^{-1} \\ \int \frac{1}{\sum_k \exp\left(\tilde{\delta}_k + \beta_i X^k\right)} \end{pmatrix}$$

where all integrals are against the distribution H of random coefficients β_i.

I estimate the parameters with indirect inference with the optimal weight matrix. I estimate the covariance matrices of the micro-moments in $\hat{s}$, $\hat{m}$ with the Bayesian bootstrap. I assume that the aggregate cash moment is independent of the other moments and is observed with only a small 1 bps standard error. Denote these estimates as $\hat{S}_1, \hat{S}_2$ respectively. Since the empirical moments are from different datasets, we have that the optimal weight matrix W is block diagonal with $\hat{S}_1^{-1}$ and $\hat{S}_2^{-1}$. Stack the moments as $g(\Sigma,\alpha,\delta,\phi) = \begin{pmatrix} s(\Sigma,\delta) & m(\Sigma,\alpha,\tilde{\delta}) \end{pmatrix}^T$ and $\hat{g} = \begin{pmatrix} \hat{s} & \hat{m} \end{pmatrix}^T$. Stack the parameters as $\theta_1 = \begin{pmatrix} \Sigma & \alpha & \delta & \tilde{\delta} \end{pmatrix}^T$. I estimate θ_1 by solving $\hat{\theta}_1 = \operatorname{argmin}_{\theta_1} (g(\theta_1) - \hat{g})^T W (g(\theta_1) - \hat{g})$.

I extract both $\hat{\Sigma}, \hat{\alpha}$ and the associated covariance matrix of the estimates. The mean utility parameters $\delta, \tilde{\delta}$ are nuisance parameters. Ultimately my model will match aggregate shares, and so I ignore the mean utility levels implied from the surveys.

7.2 Merchant Benefits, Network Costs, and Consumer Intercepts

I estimate the remaining parameters by matching the estimated effect of card acceptance on sales, matching the share of card consumers's spending at card merchants (both from section 5.3), and inverting the networks' first order conditions at the observed aggregate prices and shares.

I first make two assumptions on fees and the cost of cash. First, I assume that the aggregate fees are observed with error because my model cannot rationalize three credit card networks of different sizes charging identical fees. Instead of matching the surveyed fees in table B.2, I instead assume that MC credit charges a fee $\tau_{\text{Visa Credit}} + \Delta\tau_{\text{MC}}$ and that Amex charges a fee $\tau_{\text{Visa Credit}} + \Delta\tau_{\text{MC}} + \Delta\tau_{\text{Amex}}$, where $\Delta\tau_{\text{MC}}$ and $\Delta\tau_{\text{Amex}}$ are fee adjustment parameters to be estimated. I set the cost of cash $c_0 = \tau_0 = 0.003$ to match past studies[19] (European Commission, 2015; Felt et al., 2020).

I parameterize the distribution of merchant benefits G as a Gamma distribution with a mean $\overline{\gamma}$ and a standard deviation of σ_γ and adjust the mean and standard deviations to match the facts from the payment surveys. Let the first data moment $\hat{\phi}_1$ be the share of card consumers' spending at card stores (97%). Let the second data moment $\hat{\phi}_2$ be the logistic regression coefficient of whether a transaction is at a card merchant on whether the consumer prefers card (30). Stack these data moments as $\hat{\phi}$.

To calculate the analogous model moments, define an expenditure $e^w\left(\gamma'\right)$ for a con-

[19] In the US, many small business bank accounts charge around 30 bps to deposit cash.

sumer with wallet w as

$$e^{w}\left(\gamma'\right)=\int_{\gamma>\gamma'} q^{w*}\left(\gamma\right)p^{*}\left(\gamma\right)\ \mathrm{d}G\left(\gamma\right)$$

Let $\mathcal{M}=\{w\in\mathcal{W}:w_1\in\{\text{Visa Credit},\text{MC Credit},\text{Amex}\}\}$ be the set of wallets that are primary credit card consumers. Let $\mathcal{C}=\{w\in\mathcal{W}:w_1=\text{Cash}\}$ be the set of wallets of primary cash users. Let γ^* be the lowest merchant type that accepts all the cards. Then the analogous model moments are

$$\begin{aligned}\phi_1&=\frac{\sum_{w\in\mathcal{M}}e^{w}\left(\gamma^*\right)}{\sum_{w\in\mathcal{M}}e^{w}\left(0\right)}\\ \phi_2&=\ell\left(\phi_1\right)-\ell\left(\frac{\sum_{w\in\mathcal{C}}e^{w}\left(\gamma^*\right)}{\sum_{w\in\mathcal{C}}e^{w}\left(0\right)}\right)\\ \ell\left(p\right)&=\log\frac{p}{1-p}\end{aligned}$$

The first moment is the expenditure share of credit card consumers at card stores. The second moment is the difference in the logits of two expenditure shares: the share of credit consumers' spending at card stores and the share of cash consumers' spending at card stores. Stack these two model moments as ϕ.

Intuitively, a higher average benefit $\overline{\gamma}$ means that card consumers spend more at card stores relative to cash consumers. A higher dispersion σ_γ means card consumers spend less at card stores because a greater share of merchants decide to not accept card.

I estimate the parameters jointly by finding the 15 parameters to match 2 moment conditions $\hat{\phi}=\phi$, 8 first order conditions, and 5 share constraints. The fifteen parameters are the average $\overline{\gamma}$ and standard deviation σ_γ of merchant benefits, the 5 marginal cost parameters c for each card, the 5 utility intercepts Ξ for each card, the two fee adjustments $\Delta\tau_{\text{MC}},\Delta\tau_{\text{Amex}}$, and the CES substitution parameter σ. The 8 first order conditions are the 3 first order conditions of each credit card network with respect to its merchant fee and the 5 first order conditions of each card with respect to the promised utility U^j to the consumer. Debit card fees are not at a first order condition due to Durbin. The 5 share constraints require that at the profit maximizing promised utility for each card, the resulting aggregate shares $\tilde{\mu}$ match the data. I solve the moment conditions and the first order conditions jointly because the distribution of merchant types affects the networks' first order conditions.

I calculate the standard errors through the Delta Method. Denote all the parameters to be estimated in this step as θ_2. Stack all the first order conditions and moment conditions into a function F. The estimate $\hat{\theta}_2$ solves the equation

$$F\left(\hat{\theta}_2, \hat{\theta}_1, \hat{\phi}\right) = 0$$

The implicit function theorem gives a representation of $\hat{\theta}_2$ as $\hat{\theta}_2\left(\hat{\theta}_1, \hat{\phi}\right)$ with a known Jacobian. I calculate the covariance matrix of $\left(\hat{\theta}_1, \hat{\phi}\right)$ by assuming that the two are independent and by using the Bayesian bootstrap for the distribution of $\hat{\phi}$. The delta method converts the covariance matrix and the Jacobian into a full covariance matrix for $\hat{\theta}_2$.

7.3 Results

I estimate precise consumer elasticities, merchant elasticities, and network marginal costs. Table 8 contains all of the parameter estimates, and below I walk through the interpretation of the parameters.

The consumer parameters indicate that consumers are highly willing to substitute between payment methods, especially between payment methods with similar characteristics (e.g. credit vs debit). I transform the consumer parameters into their implications for semi-elasticities in table 9. Each column shocks the reward for a different payment method by one basis point. Each row then records how the market share of that payment method changes in response to the shock.

The first column of table 9 shows that a one basis point shock to Visa debit rewards, holding all else equal, increases the share of Visa debit primary card users by 2.4%, with a standard error of 0.4%. The new consumers mostly come from MC debit, which declines by -2.5%. In contrast, MC credit only declines by -0.6%. The difference reflects the large amount of estimated heterogeneity in consumers' valuation of credit versus debit products. Cash use only declines by -0.3%. The small change reflects the large amount of heterogeneity in consumers' valuation of cash versus cards. The third column shows a similar pattern for Visa credit. A shock to rewards steals consumers from MC credit and Amex, while having a relatively muted effect on cash and debit users.

The merchant parameters show that merchants are price insensitive. I estimate that, starting from an equilibrium where three symmetric credit card networks charge the same price, a one basis point increase in the fees to one credit card network leads to only a -0.16% decrease in the number of merchants who accept that card, with a standard error of only 0.01%. This is roughly one tenth of the sensitivities I estimate for consumers.

To calculate the merchant elasticity while holding consumer demand fixed, I use a fact proven in Appendix E.3. If all other credit card networks charge a fee of τ^* and one network deviates to a fee of τ, the lowest merchant type that accepts the deviating network is $\gamma^{'}$ where

$$\gamma^{'}(\tau) = \begin{cases} \frac{\sigma\tau}{1-\sigma\tau} & \tau < \tau^{*} \\ \frac{\sigma\rho\tau+\sigma(1-\rho)(\tau-\tau^{*})}{\rho(1-\sigma\tau)-\sigma(1-\rho)(\tau-\tau^{*})} & \tau \geq \tau^{*} \end{cases}$$

and ρ is the share of credit card holders that only carry one card. This expression is continuous but not differentiable at τ^{*}. To be consistent with the equilibrium refinement I use to solve the model I calculate the percentage change in acceptance by averaging the effects of deviations to higher and lower fees. Formally, I calculate the percentage change as

$$\frac{1}{2} \times \frac{G\left(\gamma^{'}\left(\tau^{*}-10^{-4}\right)\right)-G\left(\gamma^{'}\left(\tau^{*}+10^{-4}\right)\right)}{1-G\left(\gamma^{'}\left(\tau^{*}\right)\right)}$$

where G is the CDF of merchant types. I calculate the standard error of this change with the delta method.

The network supply parameters are also precisely estimated. I estimate marginal cost parameters that average around 43 bps with standard errors around 0. This is reasonable given accounting estimates of issuer costs around $20-40$ bps, acquiror costs of around 5 bps, and network costs of around 5 basis points[20].

The merchant elasticity is estimated more precisely than the underlying parameters governing merchant types because different combinations of primitives deliver the same merchant elasticity. Merchants have high willingness to pay both when card acceptance has a small effect on sales but markups are high (low $\overline{\gamma}$, low σ), or the sales effect is large but markups are low (high $\overline{\gamma}$, high σ). Both cases deliver similar implications for merchants price elasticity with respect to fees. So even though the CES substitution parameter of 6.4 and the average sales benefit of 34% have standard errors roughly one third of the main estimate, the standard error of the percentage change is much smaller relative to the size of the estimate.

[20]For issuer costs, Mukharlyamov and Sarin (2022) note that Durbin was crafted to target an interchange fee "reasonable and proportional" to the costs of debit cards. Initial rules considered a 30 bps interchange fee (12 cents / average ticket size of $40), which was ultimately raised to 60 bps. In Australia, credit and debit card interchange fees were also regulated by a cost based benchmark, which led to credit interchange of around 50 bps and debit interchange of 20 bps. Analyses from NACHA suggest acquirors take around 5% of the fees of credit card acceptance, such that their costs are likely between $5-10$ bps (NACHA, 2017). Visa's operating profits are around two-thirds of revenue, and so at most has a marginal cost of around 5 bps.

7.4 Goodness of Fit

The model matches the co-holding data and the effects of Durbin on credit and overall card volumes.

I show the fit of the Homescan co-holding data in figure 7. Each facet shows how well the model predicts the secondary cards of the primary card holders plotted in the facet. Each bar shows the magnitude of a log prediction error, or the log ratio of the probability of seeing the (primary, secondary) card combination in the model versus the probability in the data. The red bars show the prediction errors from my estimated model. Most red bars are close to zero, reflecting a good fit. With 7 parameters, I am able to fit the shares for all 16 shares.

The same figure shows the random coefficients play an important role in matching the secondary card data. The blue bars show the prediction errors without random coefficients. Turning off random coefficients removes the rich substitution patterns between different payment methods. The model without random coefficients fails on three dimensions for primary credit card consumers: it overpredicts the share with a secondary debit card, it overpredicts the share whose secondary payment is Cash (i.e. someone who only carries one card), and underpredicts the share who carry a secondary credit card. This is evident in the blue bars in the Visa panel. The Debit and cash bars are above zero, whereas Amex and MC are below zero.

Table 10 shows how the model performs on an out of sample test: matching the facts on Durbin. The model exactly matches the percentage change in debit volumes and the share of signature debit transactions because those were target moments in the estimation. However, the percentage changes in credit and overall card volumes serve as out of sample tests. This is an out of sample test because the random coefficients that govern substitution patterns were estimated on co-holding data from Homescan, whereas the substitution patterns revealed by the changes in credit and total card volume use exogenous price variation. The model moments are within half a standard error of the truth for both tests. This also provides evidence that co-holding data provides realistic estimates of substitution patterns in payments.

7.5 Relating Identification to the Model Intuition

The estimated parameters are closely related to the shapes of consumer and merchant demand curves and the effects of entry. As discussed in section 3, these values determine the effects of competition on prices for consumers and merchants.

The slope of the consumer demand curve is determined by the price sensitivity α, which is identified by the change in debit card volumes in response to Durbin. Because the change in debit volumes was large, sensitivity is high and the demand curve is steep. The level is determined by the card fixed effects, which are identified by the aggregate shares. The change of the consumer demand curve in response to entry is pinned down by how willing consumers are to substitute between products with similar characteristics. This is determined by the parameter Σ governing the distribution of random coefficients β_i. The parameter is identified by the Homescan data on co-holding patterns. The fact that credit card consumers' carry secondary credit cards reveals that different networks' credit cards are good substitutes for each other. The consumer demand curve should therefore fall by a large amount in response to entry.

The slope of the merchant demand curve is determined by the average benefit merchants receive from accepting cards. Merchant demand is flat because the average benefit of a 34% increase in sales is large relative to the level of fees. The intercept is determined by the share of merchants who accept cards, which is identified from the survey data. The amount the merchant curve changes in response to entry is determined by the share of consumers who hold multiple payment methods. This is again determined by the random coefficients β_i and identified by the co-holding patterns. The random coefficients are able to determine changes in both consumer and merchant demand curves because of my assumption that a consumer who carries two payment methods of the same type is *indifferent* between the two payment methods.

8 Counterfactuals

In my main counterfactual I study the effects of network entry on consumers, merchants, and the networks. I show that more competition increases merchants' cost of payments as payment methods with high merchant fees and high rewards take a larger share of the market. Consumers are worse off in aggregate. I also find that relaxing the Durbin Amendment's restrictions on debit card interchange fees and merging Mastercard and American Express would both be good for consumers. I use these counterfactuals to propose a principle for payment market regulation centered around reducing differences in rewards rates across payment options.

8.1 Credit Fintech Payment Entry

I first model the entry of a new fintech payment app that competes for credit card consumers. Introducing a new product requires specifying the characteristics X^j, Ξ^j that enter consumer utility, the type χ^j of the payment method, and network costs. I give the app consumer characteristics X^j that are the same as a credit card and the same utility intercept Ξ^j as Amex. Consumers who like cards and who like credit cards in particular will prefer the new product. I assume the new app is a new payment type $\chi^j = A$, so that at the point of sale the new app does not substitute with credit and debit cards. Given these characteristics and costs, I can solve for the new equilibrium after the fintech platform enters.

The assumption that the merchant does not treat card acceptance and app acceptance as substitutes is consistent with studies of e-commerce that consumers who prefer alternative payment methods are unwilling to substitute to cards when their preferred method is not available (Berg et al., 2022). The assumption can also be justified by the way new platforms are combining commerce and other financial services with payments, so that not accepting the app would reduce demand from consumers who use the app even if those consumers own credit and debit cards[21].

I first show that my simulated equilibrium without entry looks like the real world. Table 11 shows the baseline prices and quantities in my model. The shares are slightly different than in table B.2 because I have scaled Visa, Mastercard, and Amex up to the entire card sector. The merchant prices are similar, although I slightly under predict American Express' merchant fee. To implement Durbin, I cap debit card merchant fees at the observed equilibrium fees.

In response to entry, incumbent credit card networks raise their fees by 8 bps and drive credit card adoption by paying out 13 bps more subsidies. Incumbent debit cards raise rewards by only around 6 bps. Incumbent debit cards are unable to raise their merchant fees due to the Durbin Amendment. Merchants' cost of payments increases both because fees have risen and because more consumers are using the high cost payment options. On average, merchant prices rise by 16 bps.

Entry exacerbates regressive transfers from cash and debit consumers to credit consumers and hurts all consumers who do not switch to the new platform. For consumers who do not switch, the change in welfare is simply the change in subsidies less the change in the price index. Panel A of figure 8 illustrates the welfare effects for these consumers. The welfare of cash users who do not switch falls the most. Their consumption falls by -16 bps due to higher retail prices. Debit card users lose -11 bps of consumption, and incumbent credit card users lose -6 bps.

[21] For example, in their 2021 financial results buy now pay later platform Klarna argues "the Klarna app is now the single largest driver of GMV across the Klarna ecosystem, fuelling growth for Klarna and its retail partners through consumer acquisition and referrals...our app is becoming a central place in our consumers' financial lives."

To study the effects of entry on all consumers and the sources of welfare changes, I decompose consumer surplus into three terms – retail prices, the average subsidy paid, and non-pecuniary utility. Let E_i^k be an indicator that consumer i chooses payment method k. I compute and decompose consumer surplus[22] into three terms:

$$
\begin{aligned}
W &= \mathbb{E}_i \left[\max_k \log V_i^k \right] \\
&\approx -\log P^0 + \mathbb{E}_i \left[\max_k f^k - \log \frac{P^k}{P^0} + \Xi^k + \frac{1}{\alpha}\left(\eta_{ij} + \beta_i X^k \right) \right] \\
&= \underbrace{-\log P^0}_{\text{Retail Prices}} + \underbrace{\sum_k \mu^k f^k}_{\text{Subsidies}} + \underbrace{\mathbb{E}_i \left[\sum_k E_i^k \left(-\log \frac{P^k}{P^0} + \Xi^k + \frac{1}{\alpha}\left(\eta_{ij} + \beta_i X^k \right) \right) \right]}_{\text{Non-Pecuniary Utility}}
\end{aligned} \tag{30}
$$

where μ^k is the (insulated) market share of instrument k among all primary payment methods. Equation 30 yields three terms. The first term captures the loss to all consumers from higher retail prices. In contrast to a standard model that normalizes the value of the outside option to zero, I set the value of the outside option to the welfare of a cash consumer. The second term captures the average level of subsidies paid to consumers, weighted by the market share of each payment instrument. The third term is then the residual, and captures the extent to which consumers choose payment methods that offer high non-pecuniary utility. Any gains from variety are in this term. In practice, since α is large this term mostly reflects whether consumers choose payment methods with high unobserved characteristics Ξ^k.

Aggregate consumer surplus falls by -7 bps. Scaled up to the $10 trillion in consumer to business payments, this represents $7 billion in lost consumption. The decline in consumer surplus is surprising because entry typically raises consumer surplus by reducing markups and increasing variety (Petrin, 2002). However, because consumer adoption of credit cards and the entrants' app raises retail prices, consumers are worse off in equilibrium. Panel B in figure 8 shows how the three terms contribute to consumer surplus. Higher retail prices reduce surplus by -16 bps, higher subsidies increase surplus by 20 bps, but the shift to payment instuments with lower non-pecuniary utility hurts consumers by -10 bps. The ultimate loss is -7 bps of consumption.

The retail price externality creates a wedge between private and social incentives to

[22] Aggregating consumer surplus requires a strong assumption that the planner puts equal welfare weights on credit and debit users, which is unlikely given that credit card consumers are much higher income. Given that we already saw entry exacerbates the regressive transfers, my calculation should be considered a lower bound on the harms to consumers.

switch. Consumers who switch are privately better off because the higher rewards offsets the decline in non-pecuniary utility. However, higher rewards create no social gains. The associated higher merchant fees cause retail prices to be higher for other consumers. The above analysis highlights how the market failure in payments is not market power, but rather externalities. After entry, margins fall. Nonetheless, consumer surplus declines because consumers do not internalize the effect of using credit cards on the aggregate price level.

The retail price externality changes the sign of welfare calculations. If one ignored the equilibrium effect of retail prices, a standard discrete choice analysis based on observed market shares would lead to a 10 bps increase in consumer surplus from entry. But after including the retail price externalities we arrive at a loss of -7 bps in consumer surplus.

Total surplus declines because network profits fall as well. To measure total surplus, I assume all of the profits from either merchants or the networks are rebated to consumers equally. Even though different consumers face different CES price indices, the change in log CES utility can still be calculated by just adding the dollar value of the profits. Panel B in figure 8 decomposes the total welfare effects. Merchant profits rise by a negligible amount because consumers have higher incomes from higher network subsidies that offset higher transaction fees. Total network profits, including the entrant, fall by -4 bps or -12% of industry profits. Profits fall because networks are now competing harder to attract consumers and merchants. The net result is that total surplus falls by -10 bps.

8.2 Relaxing Durbin

Relaxing the Durbin Amendment would create a progressive transfer from credit to debit consumers and aggregate consumer benefits. I relax the Durbin Amendment in the model by raising the cap on debit card fees to 1% from their current level around 0.72%. As a result, merchant fees for debit cards rise by 28 bps and debit subsidies rise by 22 bps. Consumers switch to debit cards. The market share of debit cards rises by 5 percentage points (pp) and the market share of credit cards falls by -3% pp.

I illustrate the welfare effects in figure 9. Relaxing Durbin increases consumption of debit card users by 19 bps but reduces consumption of credit card and cash users by -1 and -3 bps, respectively. Since cash is a relatively small share of the population, on balance this is a progressive transfer from credit to debit users. Consumers as a whole gain 5 bps of consumption. Although higher retail prices cost consumers -3 bps of consumption, slightly higher subsidies and 8 bps of higher non-pecuniary utility more than compensate. Total surplus thus rises by 6 bps as networks enjoy higher profits from stealing market share from cash.

8.3 Merger Counterfactual

Merging Amex and Mastercard without any efficiencies would generate a small increase in consumer and total surplus. This result illustrates how the effects of competition in two sided markets differs starkly from one sided markets. Whereas mergers without efficiencies in one sided markets always create deadweight loss, merging two payment platforms can increase consumer and total surplus. A classic result in environmental economics is that, in the presence of negative externalities, markups can function as a tax that brings output closer to the socially efficient level (Barnett, 1980). Credit card networks create a retail price externality when they raise subsidy rates to induce consumers to use more credit cards. Adoption is excessive. A merger can reduce output, reduce the externality, and thereby raise welfare.

When Amex and MC merge, merchant fees for credit cards rise by 3 bps but more importantly subsidies for credit cards fall by -11 bps. Consumers switch to cash and debit cards. The market share of cash rises by 2 pp and the market share of debit cards rises by 1 pp.

I illustrate the welfare effects in figure 10. The merger creates progressive transfers. Cash and debit consumers gain 7 and 4 bps of consumption, whereas credit users lose -5 bps. All consumers benefit from lower retail prices, but only the credit card users are hurt by a large decline in network subsidies. Consumers as a whole gain 1 bps of consumption. Although lower subsidies cost consumers -11 bps of consumption, higher non-pecuniary utility and lower retail prices more than compensate. Total surplus rises by 6 bps as networks enjoy dramatically higher profits from the reduction in competition.

8.4 Principles for Regulation

A key principle that emerges is that regulatory policy should seek to reduce differences in rewards across payment methods. Entry exacerbates the gap in credit and debit rewards, causes consumers to sacrifice non-pecuniary utility to chase high credit card rewards, and lowers consumer surplus. Relaxing Durbin reduces the gap and thus raises consumer surplus. Mergers without efficiencies in one sided markets always reduce consumer surplus. Yet because a MC and Amex merger would lower credit card rewards in equilibrium, it raises consumer surplus from payments.

9 Conclusion

In this paper, I study how a new payment platform would change prices and welfare in the United States payment market. I find that a new credit fintech platform increases the total fees merchants pay to handle payments and can lower consumer surplus. Entry can reduce consumer surplus as consumers who do not value the non-pecuniary features of credit cards switch in order to take advantage of rewards. Such switching behavior inflates the aggregate price level and generates social losses.

I find that the market failure in payments is not market power, but rather excess adoption and retail price externalities. Unlike in standard settings in industrial organization where market power is identified with high prices and low output, the market failure in payment markets is associated with high prices and *high* output. My counterfactual results on relaxing existing price regulations on debit card fees and on merging Amex and Visa point to a new principle that regulatory and competitive changes that equalize rewards rates across payment options should be encouraged.

References

Adams, Robert and Vitaly Bord, "The Effects of the COVID-19 Shutdown on the Consumer Credit Card Market: Revolvers versus Transactors," Technical Report, Federal Reserve Board of Governors October 2020.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel, "Do Banks Pass through Credit Expansions to Consumers Who Want to Borrow?*," *The Quarterly Journal of Economics*, February 2018, *133* (1), 129–190.

Anderson, Simon P., Øystein Foros, and Hans Jarle Kind, "Competition for Advertisers and for Viewers in Media Markets," *The Economic Journal*, February 2018, *128* (608), 34–54.

Arango, Carlos, Kim P. Huynh, and Leonard Sabetti, "Consumer Payment Choice: Merchant Card Acceptance versus Pricing Incentives," *Journal of Banking & Finance*, June 2015, *55*, 130–141.

Armstrong, Mark, "Competition in Two-Sided Markets," *The RAND Journal of Economics*, September 2006, *37* (3), 668–691.

Bakos, Yannis and Hanna Halaburda, "Platform Competition with Multihoming on Both Sides: Subsidize or Not?," *Management Science*, December 2020, *66* (12), 5599–5607.

Barnett, A H, "The Pigouvian Tax Rule Under Monopoly," *American Economic Review*, December 1980, *70* (5), 1037–1041.

Berg, Tobias, Valentin Burg, Jan Keil, and Manju Puri, "On the Rise of Payment Firms," *Working Paper*, May 2022, p. 55.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, February 2004, *112* (1), 68–105.

Caillaud, Bernard and Bruno Jullien, "Chicken & Egg: Competition among Intermediation Service Providers," *The RAND Journal of Economics*, 2003, *34* (2), 309.

Chan, Lester T., "Divide and Conquer in Two-sided Markets: A Potential-game Approach," *The RAND Journal of Economics*, December 2021, *52* (4), 839–858.

Chang, Howard, David S. Evans, and Daniel D. Garcia Swartz, "The Effect of Regulatory Intervention in Two-Sided Markets: An Assessment of Interchange-Fee Capping in Australia," *Review of Network Economics*, January 2005, *4* (4).

Commission, DG Comp European, "Survey on Merchants' Costs of Processing Cash and Card Payments: Final Results, March 2015.," Technical Report, Publications Office, LU March 2015.

Constantine, Lloyd, *Priceless: The Case That Brought Down the Visa/Mastercard Bank Cartel*, Skyhorse Publishing, 2012.

Crouzet, Nicolas, Apoorv Gupta, and Filippo Mezzanotti, "Shocks and Technology Adoption: Evidence from Electronic Payment Systems," November 2020, p. 97.

Edelman, Benjamin and Julian Wright, "Price Coherence and Excessive Intermediation*," *The Quarterly Journal of Economics*, August 2015, *130* (3), 1283–1328.

Felt, Marie-Hélène, Fumiko Hayashi, Joanna Stavins, and Angelika Welte, "Distributional Effects of Payment Card Pricing and Merchant Cost Pass-through in the United," Federal Reserve Bank of Boston Research Department Working Papers, Federal Reserve Bank of Boston December 2020.

Gans, Joshua S, "Evaluating the Impact of the Payment System Reforms," Technical Report January 2007.

— **and Stephen P King**, "The Neutrality of Interchange Fees in Payment Systems," *Topics in Economic Analysis & Policy*, 2003, *3* (1), 1–16.

Gentzkow, Matthew, Jesse M Shapiro, Frank Yang, and Ali Yurukoglu, "Pricing Power in Advertising Markets: Theory and Evidence," *Working Paper*, July 2022, p. 73.

Ghosh, Pulak, Boris Vallee, and Yao Zeng, "FinTech Lending and Cashless Payments," *SSRN Electronic Journal*, 2021.

Gowrisankaran, Gautam and Joanna Stavins, "Network Externalities and Technology Adoption: Lessons from Electronic Payments," *The RAND Journal of Economics*, 2004, *35* (2).

Guthrie, Graeme and Julian Wright, "Competing Payment Schemes," *Journal of Industrial Economics*, March 2007, *55* (1), 32.

Hayashi, Fumiko, "The New Debit Card Regulations: Initial Effects on Networks and Banks," *Federal Reserve Bank of Kansas City Economic Review*, 2012, p. 38.

Higgins, Sean, "Financial Technology Adoption," *Working Paper*, December 2020, p. 80.

Huynh, Kim, Gradon Nicholls, and Oleksandr Shcherbakov, "Equilibrium in Two-Sided Markets for Payments: Consumer Awareness and the Welfare Cost of the Interchange Fee," *Bank of Canada Working Paper*, March 2022.

Huynh, Kim P, Jozsef Molnar, Oleksandr Shcherbakov, and Qinghui Yu, "Demand for Payment Services and Consumer Welfare: The Introduction of a Central Bank Digital Currency," *Bank of Canada Currency Department Working Paper*, March 2020, p. 34.

Jones, Barbara, "U.S. v. VISA U.S.A., et Al.," October 2001.

Koulayev, Sergei, Marc Rysman, Scott Schuh, and Joanna Stavins, "Explaining Adoption and Use of Payment Instruments by US Consumers," *RAND Journal of Economics*, 2016.

Li, Bin Grace, James McAndrews, and Zhu Wang, "Two-Sided Market, R&D, and Payments System Evolution," *Journal of Monetary Economics*, November 2020, *115*, 180–199.

Liu, Chunchun, Tat-How Teh, Julian Wright, and Junjie Zhou, "Multihoming and Oligopolistic Platform Competition," *SSRN Working Paper*, 2021, p. 47.

Mukharlyamov, Vladimir and Natasha Sarin, "Price Regulation in Two-Sided Markets: Empirical Evidence from Debit Cards," *Working Paper*, September 2022.

NACHA, "Understanding the Cost of Processing Card Payments," Technical Report, Optimized Payments Consulting 2017.

Orem, Tina, "Credit Unions Revive Debit Rewards," *Credit Union Times*, January 2016.

Parlour, Christine A, Uday Rajan, and Haoxiang Zhu, "When FinTech Competes for Payment Flows," *Review of Financial Studies*, April 2022, p. 64.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 2002, *110* (4).

Pless, Jacquelyn and Arthur A. van Benthem, "Pass-Through as a Test for Market Power: An Application to Solar Subsidies," *American Economic Journal: Applied Economics*, October 2019, *11* (4), 367–401.

Rochet, Jean-Charles and Jean Tirole, "Cooperation among Competitors: Some Economics of Payment Card Associations," *The RAND Journal of Economics*, 2002, *33* (4), 549.

— **and** — , "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, June 2003, *1* (4), 990–1029.

Rosaia, Nicola, "Competing Platforms and Transport Equilibrium: Evidence from New York City," *Working Paper*, November 2020, p. 77.

Ru, Hong and Antoinette Schoar, "Do Credit Card Companies Screen for Behavioural Biases?," *BIS Working Paper*, 2020, p. 53.

Rysman, Marc, "Competition between Networks: A Study of the Market for Yellow Pages," *Review of Economic Studies*, April 2004, *71* (483), 512.

— , "AN EMPIRICAL ANALYSIS OF PAYMENT CARD USAGE," *Journal of Industrial Economics*, March 2007, *55* (1), 1–36.

Schneider, James and S.K. Prasad Borra, "The Future of Finance: Redefining "The Way We Pay" in the Next Decade," Technical Report, Goldman Sachs Equity Research March 2015.

Stavins, Joanna, "Consumer Preferences for Payment Methods: Role of Discounts and Surcharges," *Journal of Banking & Finance*, September 2018, *94*, 35–53.

Tan, Guofu and Junjie Zhou, "The Effects of Competition and Entry in Multi-sided Markets," *The Review of Economic Studies*, March 2021, *88* (2), 1002–1030.

Valverde, Santiago Carbó, Sujit Chakravorti, and Francisco Rodríguez Fernández, "The Role of Interchange Fees in Two-Sided Markets: An Empirical Investigation on Payment Cards," *Review of Economics and Statistics*, May 2016, *98* (2), 367–381.

Weyl, E. Glen, "A Price Theory of Multi-Sided Platforms," *American Economic Review*, September 2010, *100* (4), 1642–1672.

— **and Michal Fabinger**, "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, June 2013, *121* (3), 528–583.

White, Alexander and E. Glen Weyl, "Insulated Platform Competition," *SSRN Electronic Journal*, 2016.

Tables and Figures

Figure 1: Illustration of payment flows in a payment network

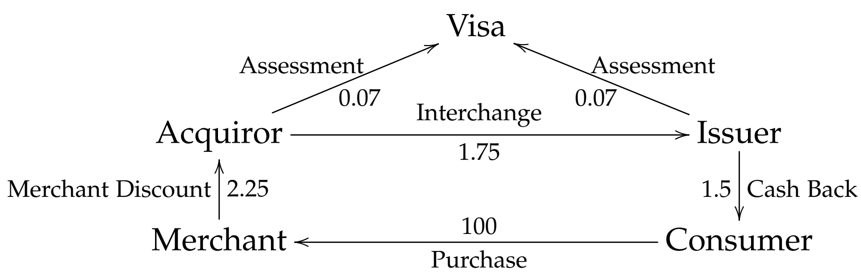


Figure 2: Intuition on the relationship between competition and fees. Dotted lines illustrate counterfactual demand curves after entry of a new competitor.

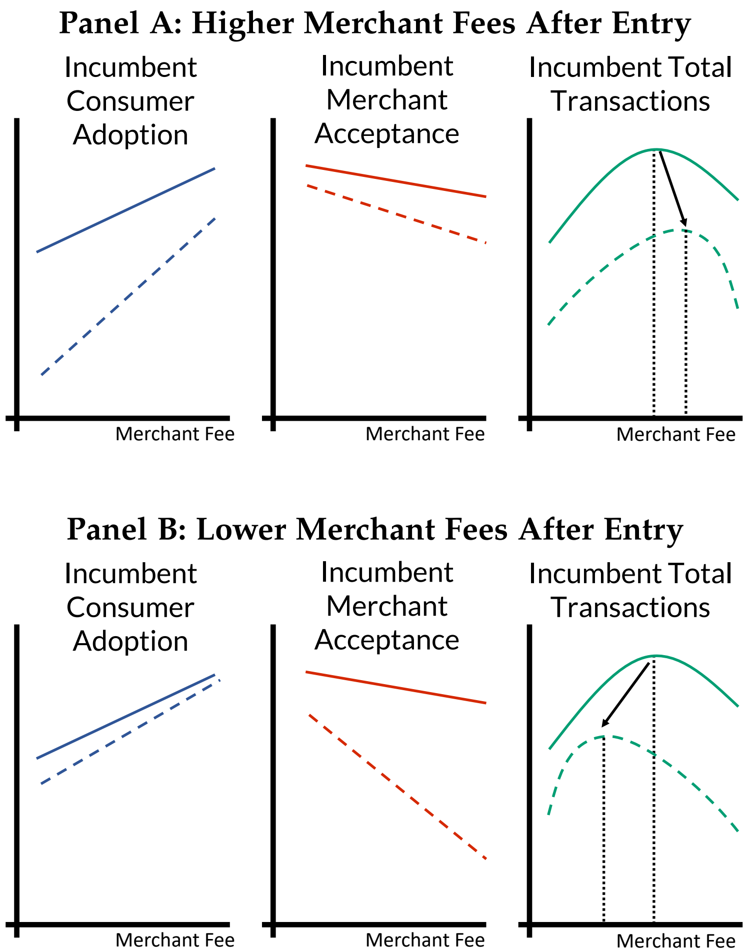


Figure 3: Intuition on the relationship between fees and welfare

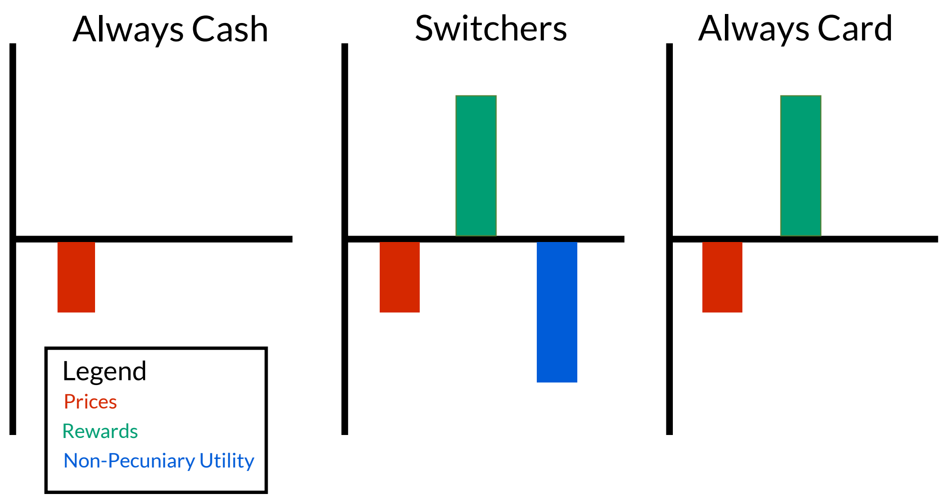


Figure 4: The effect of the Durbin Amendment on debit card and credit card volume. The vertical line marks the year before the policy announcement. The policy started in Q3 2011 and went into full effect in year 2012, which is at $t = 1$.

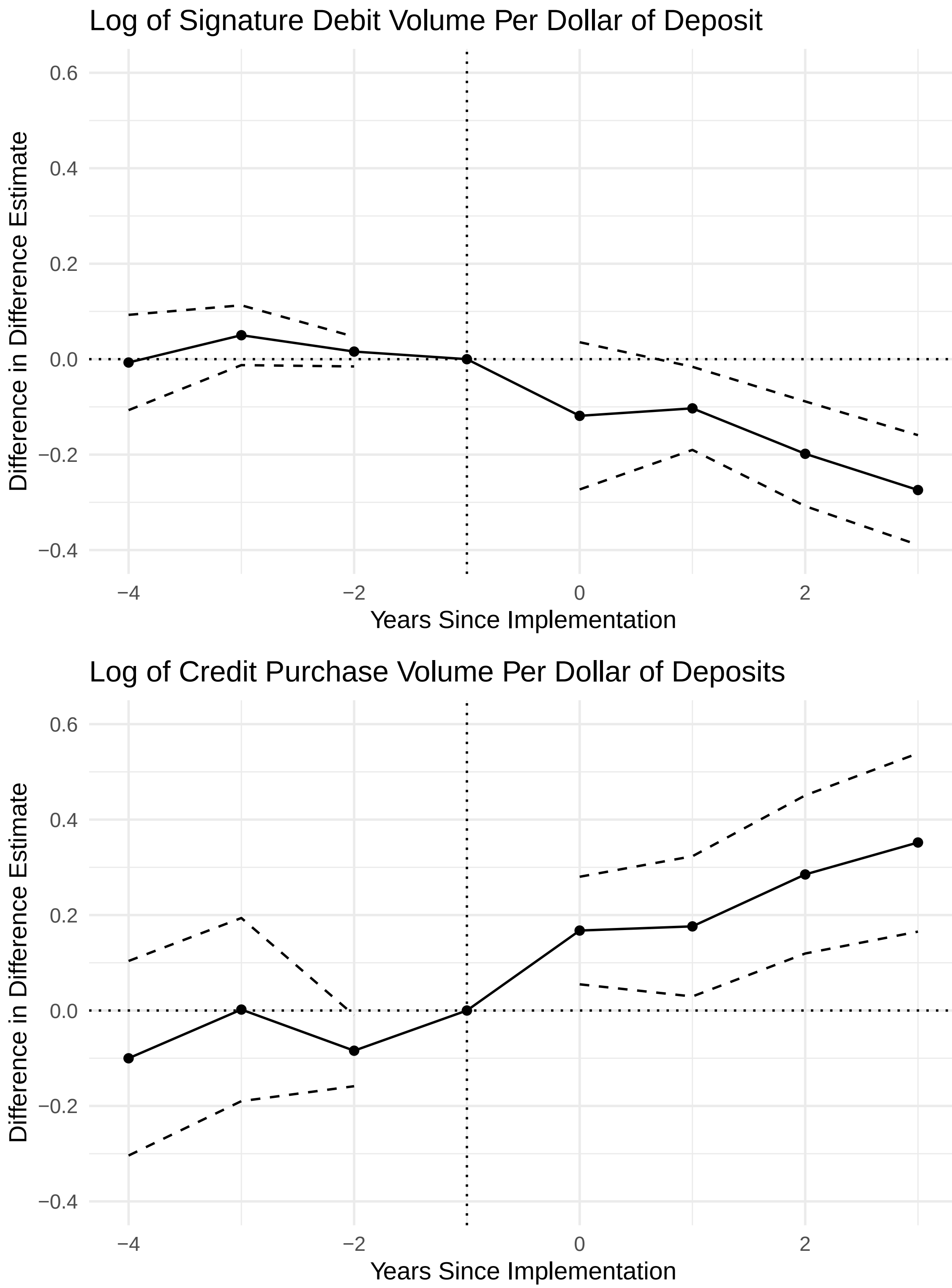


Figure 5: Card fees and acceptance around Durbin

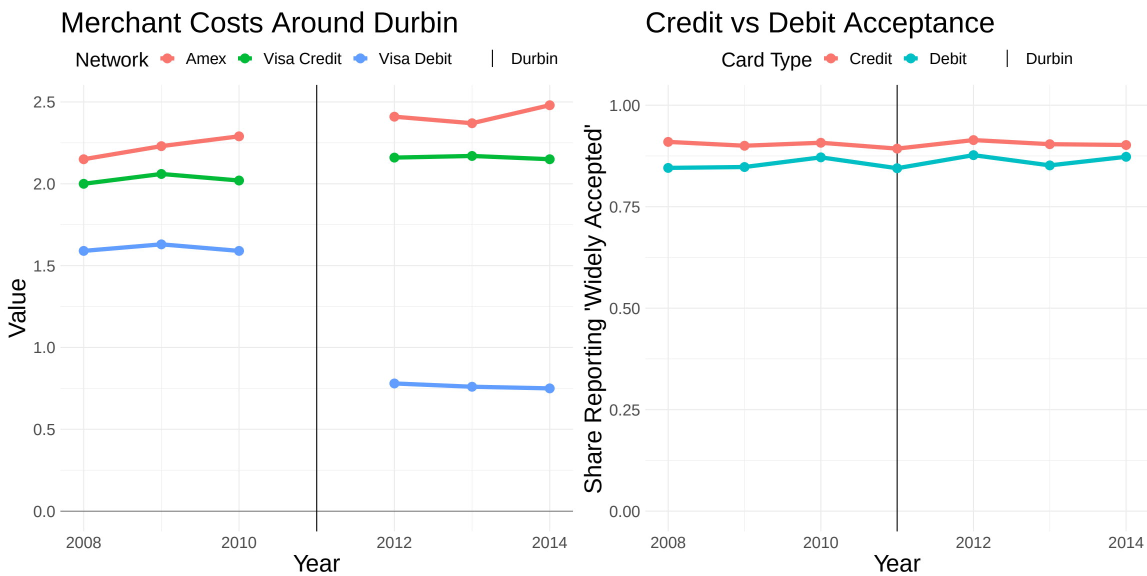


Figure 6: Illustration of how consumers choose payment methods at the point of sale. Note how the Amex/Debit consumer does not spend on her debit card because it is not the same type as her primary card.

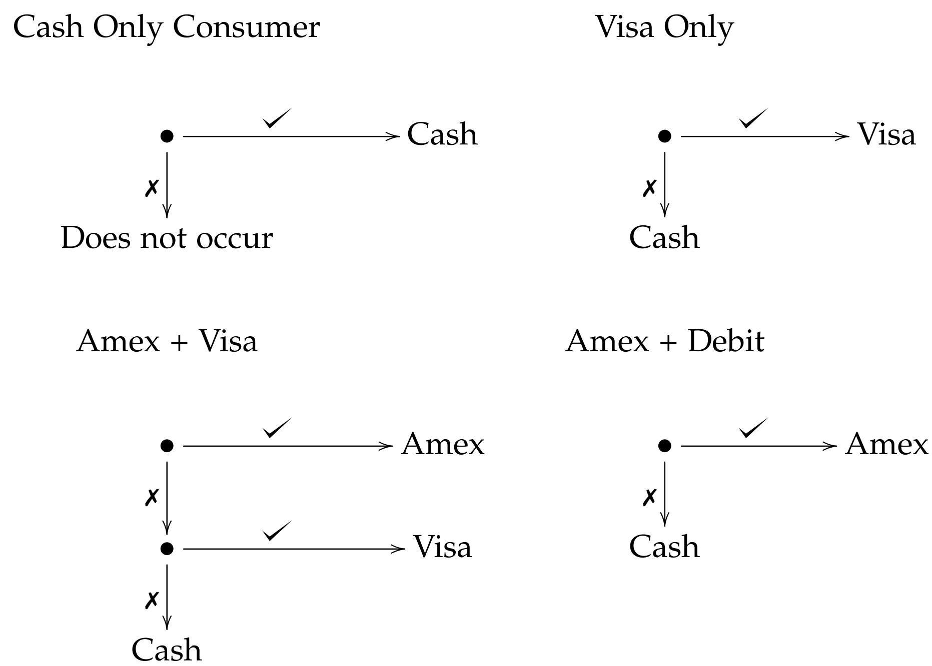


Figure 7: Value of random coefficients in matching co-holding data

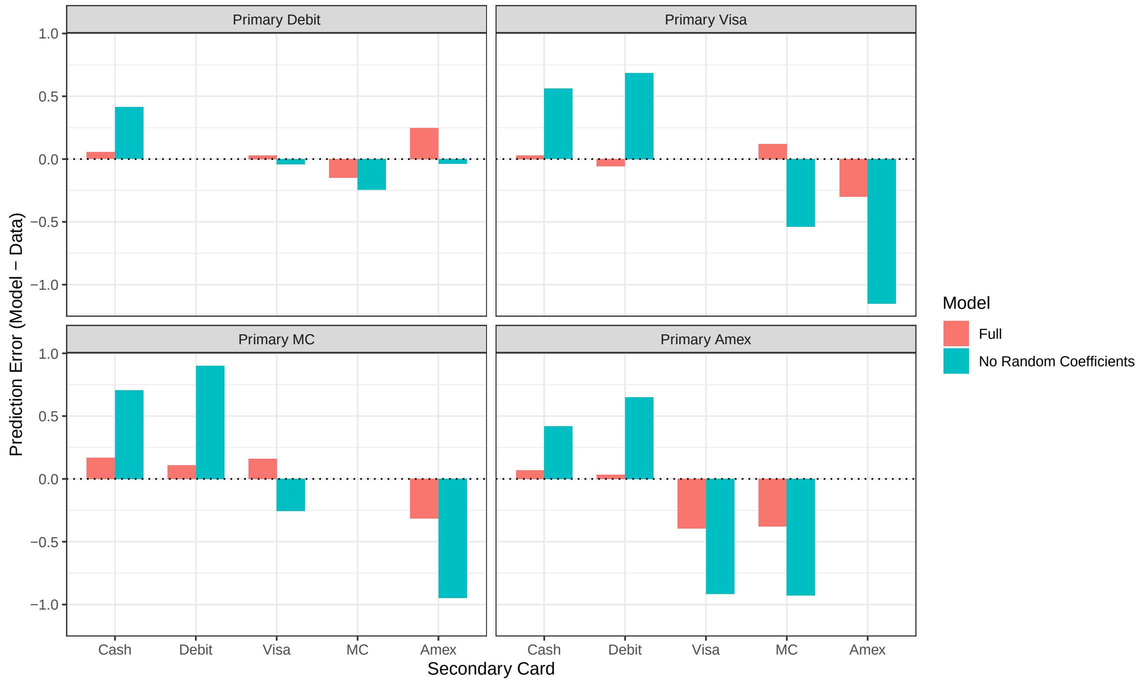
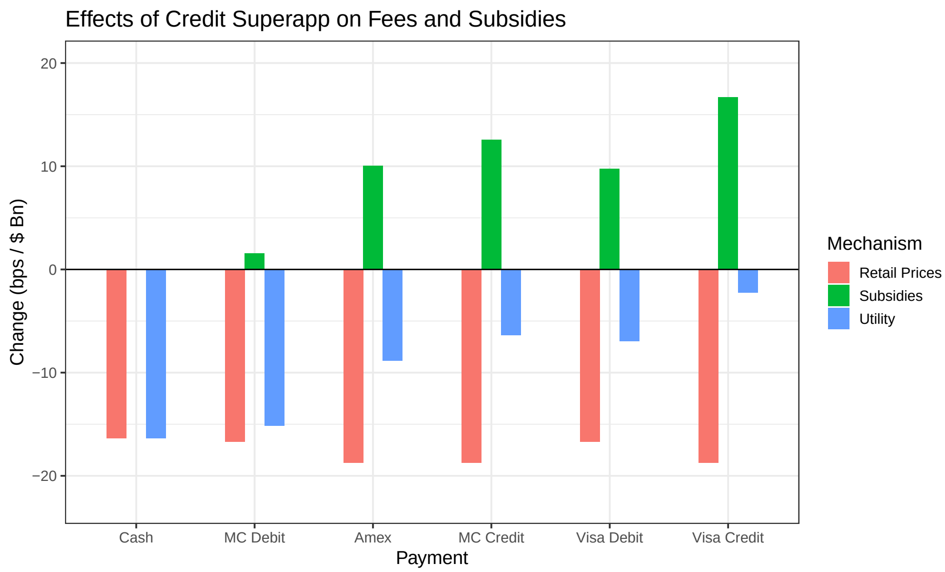


Figure 8: Welfare effects of entry of a credit fintech payment platform

Panel A: Distributional Effects for Non-Switchers



Panel B: Decomposition of Consumer Surplus and Total Surplus Effects

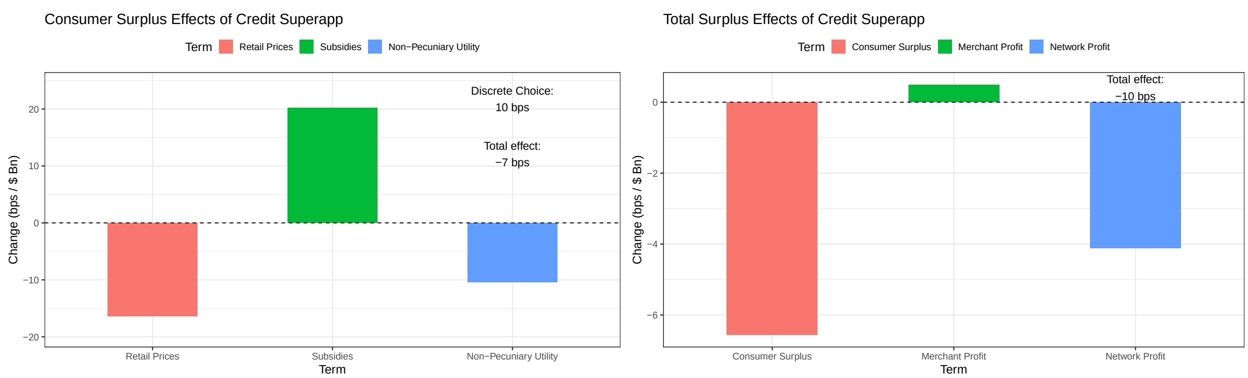
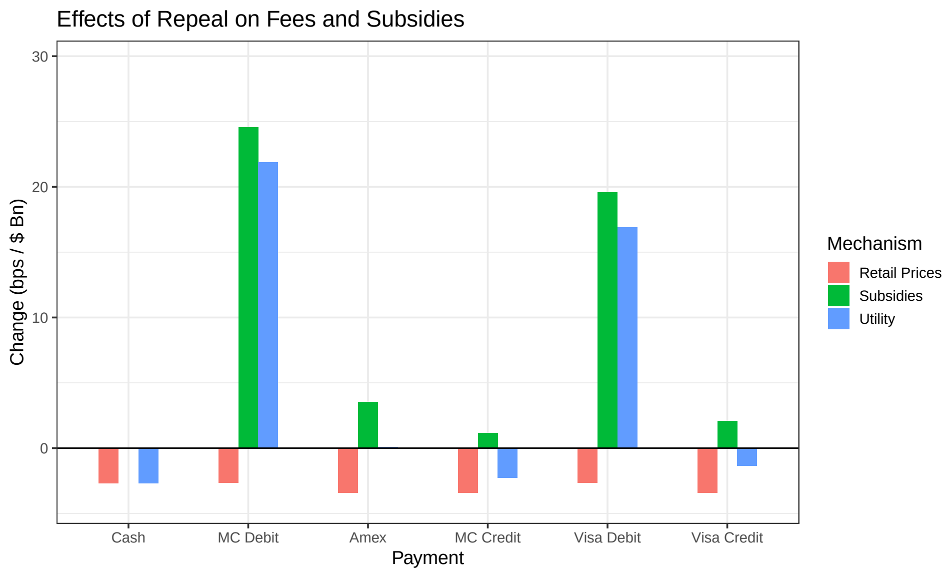


Figure 9: Welfare effects of relaxing the Durbin Amendment cap on debit card fees

Panel A: Distributional Effects for Non-Switchers



Panel B: Decomposition of Consumer Surplus and Total Surplus Effects

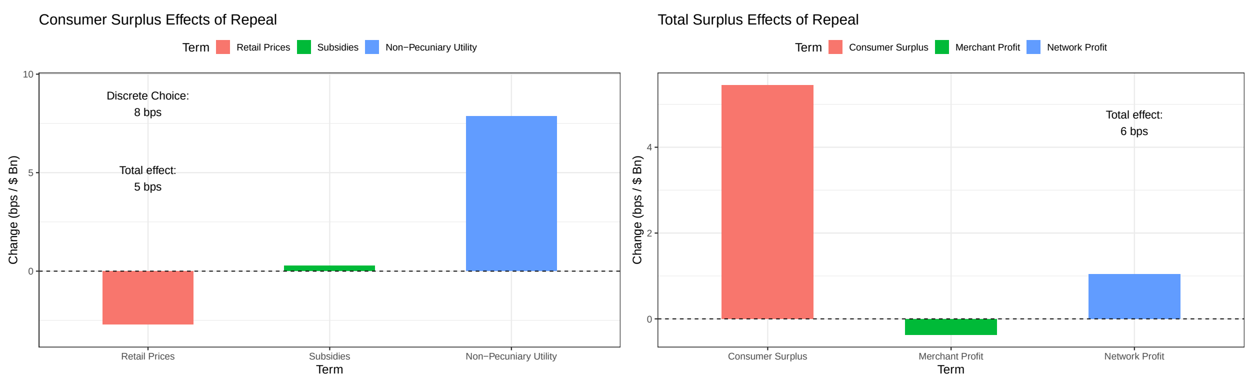
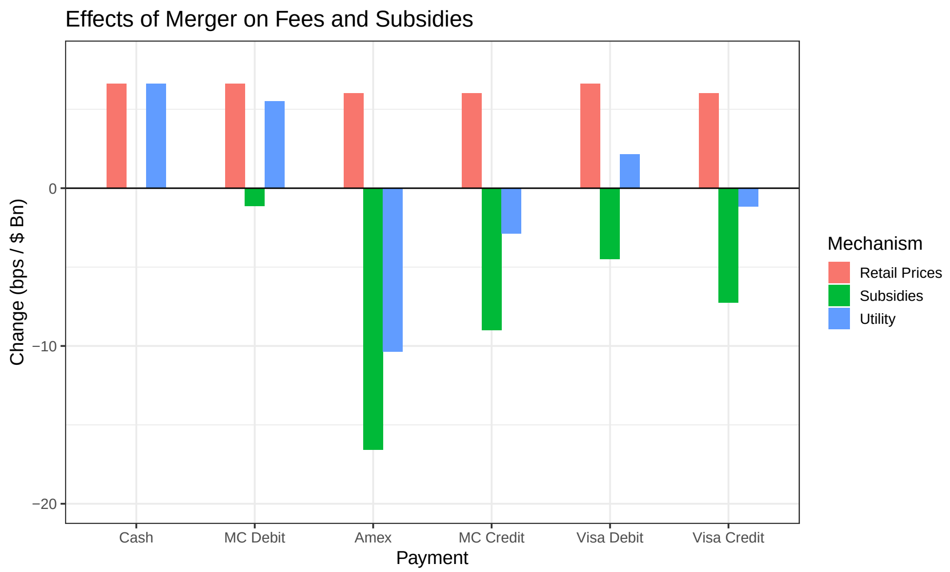


Figure 10: Welfare effects of relaxing the Durbin Amendment cap on debit card fees

Panel A: Distributional Effects for Non-Switchers



Panel B: Decomposition of Consumer Surplus and Total Surplus Effects

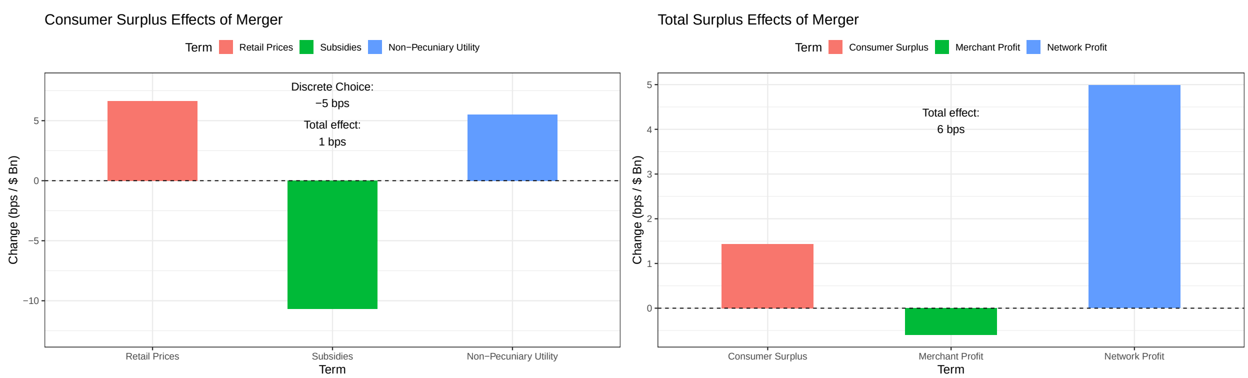


Table 1: Summary statistics of Nilson Report panel

	N	Mean	P25	P50	P75
Assets	309	29126.09	4207.34	9673.76	34162.04
Credit	296	1431.63	365.44	554.50	1455.00
Debit	294	4928.75	1237.25	2526.00	5435.00
Signature Debit	292	3035.46	783.75	1270.50	2715.25
Sig Debit Ratio	279	0.66	0.58	0.68	0.78
Treated	309	0.48	0.00	0.00	1.00

Table 2: Summary statistics of the transactions in the payment diary sample

	N	Mean	P25	Median	P75
Ticket Size	31763	22.08	6.67	15.01	30.00
Prefers Card	31763	0.78	1.00	1.00	1.00
Prefers Credit	31763	0.35	0.00	0.00	1.00
Prefers Debit	31763	0.43	0.00	0.00	1.00
Prefers Cash	31763	0.20	0.00	0.00	0.00
Merchant Accepts Card	29658	0.95	1.00	1.00	1.00

Table 3: Summary statistics of the consumer types in the payment diary sample. The share variable reports the share of the sample in each column. All other variables report averages within the group of consumers of a given payment choice. All averages and shares are calculated with individual level sampling weights.

	Cash	Debit	Credit	Credit + Debit
Share	0.24	0.30	0.26	0.20
Carry Balance	0.41	0.57	0.28	0.65
Has CC	0.71	0.75	1.00	1.00
Rewards CC	0.48	0.44	0.88	0.78
Rewards DC	0.13	0.16	0.12	0.16
HH Income ($k)	60.11	67.13	107.83	92.43
Age	52.19	48.74	55.08	50.04

Table 4: Summary statistics of the Homescan sample

	N	Mean	P25	Median	P75
Years per Household	92107	3.06	1.00	2.00	5.00
Transactions	92107	500.49	134.00	306.00	669.00
Average Tx Size	92107	56.62	35.41	49.56	69.43
Debit	92107	0.36	0.01	0.29	0.67
Visa	92107	0.17	0.00	0.04	0.21
MC	92107	0.08	0.00	0.01	0.06
Amex	92107	0.03	0.00	0.00	0.01
Discover	92107	0.03	0.00	0.00	0.00
Cash	92107	0.27	0.06	0.17	0.40

Table 5: Shares of consumers with each primary and secondary card combination. If a consumer only uses one type of card, the secondary "card" is defined as cash.

	Secondary Card				
Primary Card	Cash	Debit	Visa	MC	Amex
Debit	0.22		0.45	0.26	0.07
Visa	0.16	0.38		0.29	0.17
MC	0.13	0.29	0.45		0.13
Amex	0.09	0.20	0.49	0.22	
Primary Card Share	0.26	0.44	0.18	0.08	0.04

Table 6: Logistic regressions predicting the probability that a given transaction occurs at a merchant who accepts credit cards as a function of consumer preferences.

	No Controls	Tx Controls	Consumer Controls	Both
Prefer Card	0.352***	0.342***	0.358***	0.299***
	(0.073)	(0.085)	(0.078)	(0.088)
N	29658	29658	29658	29658
Year FE	X	X	X	X
Merch Type FE		X		X
Ticket Size FE		X		X
FICO Category FE			X	X
Age Group FE			X	X
Income Category FE			X	X
Education FE			X	X
State FE		X	X	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Number of credit cards and debit cards carried by typical consumer

	Rysman (2007) Visa Payments Panel	DCPC	Homescan
Share of Credit Card Singlehomers	0.51	0.39	0.38

Table 8: Estimated parameters

Panel A: Consumer Parameters

Parameter	Estimate	SE
SD of Credit RC	2.0	0.0
SD of Card RC	4.9	0.1
Correlation of RC	-0.3	0.0
Price Sensitivity α	483.7	87.3
Visa Debit Intercept $\times 100$	-4.6	0.2
Visa Credit Intercept $\times 100$	-5.7	0.2
MC Debit Intercept $\times 100$	-4.8	0.2
MC Credit Intercept $\times 100$	-5.8	0.2
Amex Intercept $\times 100$	-5.9	0.2

Panel B: Merchant Parameters

Parameter	Estimate	SE
CES σ	7.2	1.9
Average γ	0.3	0.1
Log Ratio of $\frac{\sigma_\gamma}{\overline{\gamma}}$	-1.1	0.1

Panel C: Network Parameters (bps)

Parameter	Estimate	SE
Visa Debit Cost	43.3	0.2
Visa Credit Cost	13.7	0.4
MC Debit Cost	52.3	0.1
MC Credit Cost	56.1	0.3
Amex Cost	57.7	0.4
MC Fee Adj	0.1	0.0
Amex Fee Adj	0.0	0.0

Table 9: Estimated consumer own price and cross price semi-elasticities. Each entry shows the effect of a one basis point change in the rewards of the column payment method on the market share of the row payment method. The change is measured as a percentage of the row payment method's market share.

Payment	V Debit	MC Debit	V Credit	MC Credit	Amex
Cash	$-0.3\,(0.1)$	$-0.1\,(0.0)$	$-0.6\,(0.1)$	$-0.2\,(0.0)$	$-0.2\,(0.0)$
V Debit	$+2.4\,(0.4)$	$-1.0\,(0.2)$	$-0.7\,(0.1)$	$-0.3\,(0.0)$	$-0.2\,(0.0)$
MC Debit	$-2.5\,(0.4)$	$+3.9\,(0.7)$	$-0.7\,(0.1)$	$-0.3\,(0.0)$	$-0.2\,(0.0)$
V Credit	$-0.6\,(0.1)$	$-0.2\,(0.0)$	$+2.8\,(0.5)$	$-0.8\,(0.1)$	$-0.7\,(0.1)$
MC Credit	$-0.6\,(0.1)$	$-0.2\,(0.0)$	$-2.0\,(0.4)$	$+4.0\,(0.7)$	$-0.7\,(0.1)$
Amex	$-0.6\,(0.1)$	$-0.2\,(0.0)$	$-2.0\,(0.4)$	$-0.8\,(0.1)$	$+4.1\,(0.7)$

Table 10: Fit of Durbin Facts

Reduced Form Fact	Data	Model	Standard Error
Effect on Debit	-0.27	-0.27	0.06
Ratio of Debit to Credit Volumes	0.65	0.65	0.02
Effect on Credit	0.35	0.30	0.09
Effect on All Card Volume	-0.05	-0.04	0.05

Table 11: Baseline equilibrium prices and quantities

Variable (%)	Cash	Visa Debit	Visa Credit	MC Debit	MC Credit	Amex
Merchant Fee	0.3	0.72	2.25	0.72	2.25	2.25
Consumer Subsidy	0.0	0.00	1.30	0.00	1.30	1.36
Market Share	20.0	23.88	26.27	9.55	10.75	9.55

A Proofs

Proof of Theorem 1. I first prove the theorem for $\gamma = 0$, and when $\tau = 0$. I then usethe envelope theorem to extrapolate to positive values of τ. From the definition in equation 7, profits in general are

$$\begin{aligned}\hat{\Pi}(\tau) &= \left(\sum_{w \in \mathcal{W}} \tilde{\mu}^{w} q^{w}(\gamma, \hat{p}, M, P, y^{w}) \right) \times \left(\hat{p}\left(1 - \tau_{M}^{w}\right) - 1 \right) \\ &= \frac{1}{C} \sum_{w \in \mathcal{W}} \mu^{w} \left(1 + \gamma v_{M}^{w}\right) \hat{p}^{-\sigma} \left(\hat{p}\left(1 - \tau_{M}^{w}\right) - 1 \right)\end{aligned}$$

Suppress the $\mathcal{W}$ and leave out the $\frac{1}{C}$ normalizing factor. When $\gamma = 0$, profit simplifies to

$$\begin{aligned}\hat{\Pi} &= \sum_{w} \mu^{w} \hat{p}^{-\sigma} \left(\hat{p}\left(1 - \tau_{M}^{w}\right) - 1 \right) \\ \hat{p} &= \frac{\sigma}{\sigma - 1} \frac{1}{1 - \hat{\tau}}\end{aligned}$$

At a fee of zero, $\hat{\tau} = 0$. Hence profits are

$$\begin{aligned}\hat{\Pi}(0) &= \frac{1}{\sigma - 1} \times \hat{p}^{-\sigma} \left(\sum_{w} \mu^{w} \right) \\ &= \left(\frac{\sigma}{\sigma - 1} \right)^{1 - \sigma} \frac{1}{\sigma} \left(\sum_{w} \mu^{w} \right)\end{aligned}$$

We next establish the result for small τ. By the envelope theorem, the derivative of the optimized profit for a $\gamma = 0$ firm with respect to the transaction fees τ at zero is

$$\begin{aligned}\left. \frac{\partial \hat{\Pi}}{\partial \tau_{j}} \right|_{\tau_{j} = 0} &= \sum_{w} \mu^{w} \hat{p}^{-\sigma} \left(-\hat{p} I_{j,M}^{w} \right) \\ &= -\left(\frac{\sigma}{\sigma - 1} \right)^{1 - \sigma} \sum_{w} \mu^{w} I_{j,M}^{w}\end{aligned}$$

where the indicator $I_{j,M}^{w}$ is an indicator capturing whether payment method j is used by the wallet w when the merchant accepts M. Crucially, we have that the indicators multiplied by the fees gives the card fee that the consumer will cause the merchant to pay $\sum_{j=1}^{J} I_{j,M}^{w} \tau_{j} = \tau_{M}^{w}$. We can then compute profits at a generic level of fees with a Taylor approximation. Up to second order terms in τ, this should equal

$$\begin{aligned}
\hat{\Pi}(\tau) &= \hat{\Pi}(0) + \sum_{j=1}^{J} \frac{\partial \hat{\Pi}}{\partial \tau_j} \tau_j + O\left((\tau^{\max})^2 \right) \\
&\approx \sum_w \mu^w \left[\left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \frac{1}{\sigma} - \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \tau_M^w \right] \\
&= \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \left[-\sum_w \mu^w \tau_M^w + \frac{1}{\sigma} \right]
\end{aligned}$$

This establishes the theorem for $\gamma = 0$ and small τ.

Next we prove the result for generic γ. Recall that $\hat{\tau}$ is the realized average card fee that the merchant incurs, and enters into optimal pricing. Drop terms that are of order $O\left(\tau^2\right)$. By the envelope theorem we can ignore the effect of changing γ on the optimal price. Hence the derivative of optimized profit with respect to γ is

$$\begin{aligned}
\frac{\partial \hat{\Pi}}{\partial \gamma} &= \sum_w \mu^w v_M^w \hat{p}^{-\sigma} \left(\hat{p} \left(1 - \tau^w \right) - 1 \right) \\
&\approx \sum_w \mu^w v_M^w \left(\frac{\sigma}{\sigma - 1} \right)^{-\sigma} \left(1 - \sigma \hat{\tau} \right) \left(\frac{\sigma}{\sigma - 1} \left(1 + \hat{\tau} \right) \left(1 - \tau^w \right) - 1 \right) \\
&\approx \left(\frac{\sigma}{\sigma - 1} \right)^{-\sigma} \frac{1}{\sigma - 1} \times \sum_w \mu^w v_M^w \left(1 - \sigma \hat{\tau} \right) \left(1 + \sigma \hat{\tau} - \sigma \tau^k \right) \\
&\approx \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \frac{1}{\sigma} \sum_w \mu^w v_M^w \left(1 - \sigma \tau^k \right)
\end{aligned}$$

Integrating the derivative from 0 to γ gives the desired result. $\square$

B Additional Tables

Table B.1: Aggregate shares and cost of card acceptance derived from the Nilson Report

Payment Method	Volume in 2019 (Tr)	Share of Total
Total	9.6	
Cash + Check	1.9	20%
Cards	7.7	80%
Credit	4.0	42%
Visa	2.1	22%
MC	0.9	9%
Amex	0.8	8%
Discover	0.1	1%
Other	0.1	1%
Debit	3.3	34%
Visa	1.9	20%
MC	0.8	8%
Other	0.6	6%
o/w Other Cards	0.4	4%

Table B.2: Aggregate prices for merchants and consumers, as well as acceptance locations. Merchant discount fees are calculated from a survey of acquirers. Rewards come from Amex's 10K statements and Agarwal et al. (2018). Acceptance locations are also estimates from the Nilson Report.

Card	Average Merchant Discount	Rewards	Number of Acceptance Locations (Mln)
Visa + MC Credit	2.25%	1.30%	10.7
Amex	2.27%	1.36%	10.6
Visa + MC Debit	0.72%	0%	

Table B.3: Comparing Nielsen payment shares to aggregate shares

Payment Method	Homescan	Nilson
Amex	0.04	0.10
Cash	0.24	0.20
Debit	0.37	0.33
MC	0.11	0.11
Visa	0.24	0.26

Figure B.1: The effect of the Durbin Amendment on interchange revenue. The vertical line marks the year before the policy announcement. The policy went into full effect in year 2012, which is at $t = 1$.

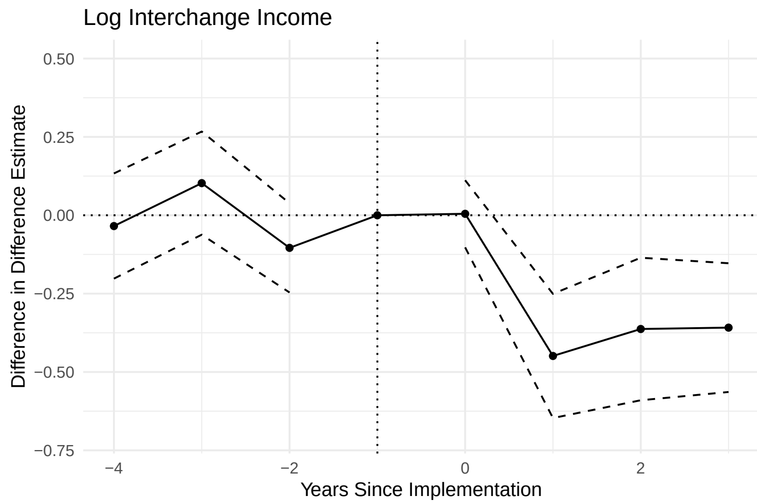


Table B.4: Event study estimates for the effect of the Durbin Amendment on signature credit, debit card, and total volume

	Interchange	Signature Debit	Credit	All Cards
Treat, t=-4	-0.034	-0.007	-0.100	-0.111+
	(0.086)	(0.051)	(0.104)	(0.060)
Treat, t=-3	0.103	0.050	0.002	-0.006
	(0.084)	(0.032)	(0.098)	(0.050)
Treat, t=-2	-0.104	0.016	-0.084*	-0.016
	(0.073)	(0.016)	(0.038)	(0.027)
Treat, t=0	0.005	-0.119	0.168**	-0.006
	(0.055)	(0.079)	(0.057)	(0.056)
Treat, t=1	-0.449***	-0.103*	0.176*	0.020
	(0.101)	(0.044)	(0.075)	(0.048)
Treat, t=2	-0.363**	-0.198**	0.285**	0.002
	(0.116)	(0.056)	(0.085)	(0.057)
Treat, t=3	-0.358**	-0.274***	0.352***	-0.048
	(0.105)	(0.059)	(0.095)	(0.057)
N	292	292	296	281
Bank FE	X	X	X	X
Year FE	X	X	X	X
Cluster N	39	39	39	39

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.5: Subgroup analysis for the effect of card preference on the likelihood the consumer shops at a store that accepts card

	Credit vs Debit	Singlehome	Singlehome CC	Income Group
Prefer Credit	0.253*			
	(0.110)			
Prefer Debit	0.329***			
	(0.095)			
Singlehome X Prefer Card		0.109	0.064	
		(0.129)	(0.091)	
Prefer Card		0.269**	0.277**	0.452***
		(0.093)	(0.094)	(0.127)
High Income X Prefer Card				-0.267
				(0.168)
N	29658	29089	29241	29658
Year FE	X	X	X	X
Merch Type FE	X	X	X	X
Ticket Size FE	X	X	X	X
FICO Category FE	X	X	X	X
Age Group FE	X	X	X	X
Income Category FE	X	X	X	X
Education FE	X	X	X	X
State FE	X	X	X	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.6: Correlation between being the card with the top number of trips and the card with the top share of spending.

		Top Card by Spend			
Top Card by Trips		Amex	Debit	MC	Visa
Amex	N	11568	111	142	527
	% row	93.7	0.9	1.1	4.3
Debit	N	639	132097	1422	2670
	% row	0.5	96.5	1.0	2.0
MC	N	444	426	26806	1057
	% row	1.5	1.5	93.3	3.7
Visa	N	871	910	1079	61791
	% row	1.3	1.4	1.7	95.6

Table B.7: The average share of total card spending on consumers' top two cards split by the primary card of each consumer

Primary Card	Primary Share	Secondary Share	Top Two Total
Amex	0.76	0.18	0.94
Visa	0.81	0.15	0.97
MC	0.77	0.18	0.95
Debit	0.86	0.11	0.97

C Additional Figures

Figure C.1: The effect of the Durbin Amendment on deposits and assets

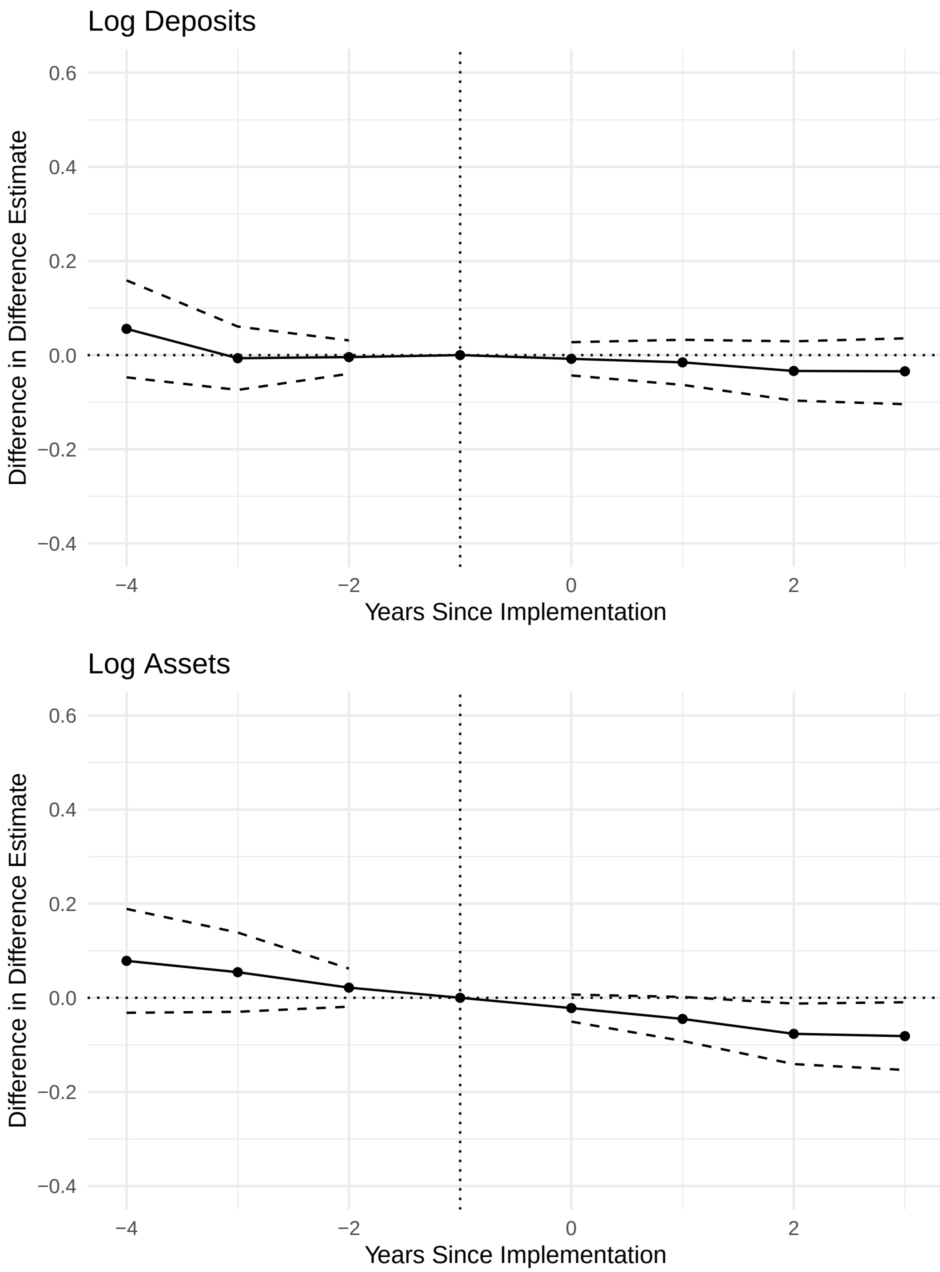


Figure C.2: The effect of the Durbin Amendment on overall debit volumes

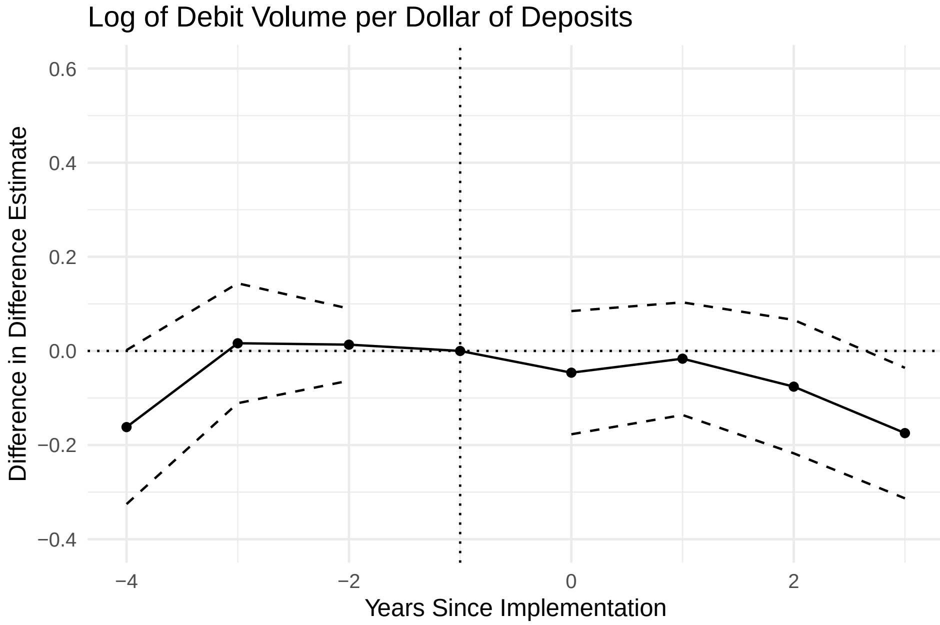
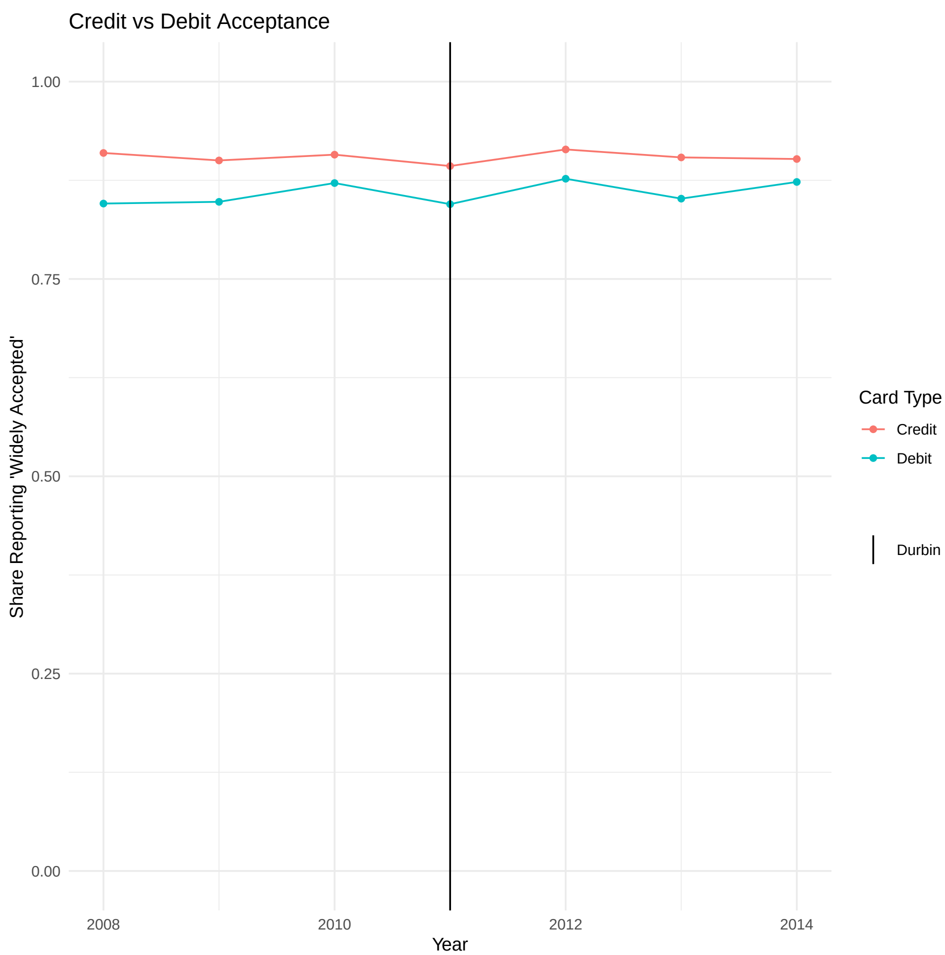


Figure C.3: Consumer ratings of acceptance of credit and debit cards around Durbin



D A Method for Calculating Derivatives of Expectations of Nondifferentiable Functions

Suppose $f : \mathbb{R}^N \to \mathbb{R}$ is continuous but non-differentiable. Then by a standard convolution theorem

$$\begin{aligned} h : \mathbb{R}^N &\to \mathbb{R} \\ \mu &\mapsto \mathbb{E}\left[f\left(X\right)\right], X \sim N\left(\mu,\sigma^2 I\right) \end{aligned}$$

is differentiable. This note explains how to efficiently compute an approximation to the partial derivatives of h. This is non-trivial because the standard monte carlo technique to approximate h as $\hat{h} = N^{-1} \sum_{i=1}^{N} f\left(X_i\right)$ where $X_i \sim N\left(\mu,\sigma^2 I\right)$ does not generate a differentiable function in μ.

The key trick is to use the fact that convolution and differentiation commute. Let $g\left(x\right) = \mathbb{E}\left[f\left(X_1,\ldots,X_N\right)|X_1 = x\right]$. Then by the law of iterated expectations,

$$\mathbb{E}\left[f\left(X\right)\right] = \mathbb{E}\left[g\left(X_1\right)\right]$$

By the law of iterated expectations, we have that

$$\begin{aligned} \mathbb{E}\left[f\left(X\right)\right] &= \mathbb{E}\left[g\left(X_1\right)\right] \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} g\left(z\right) \exp\left(-\frac{1}{2\sigma^2}\left(z-\mu_1\right)^2\right) \, \mathrm{d}z \end{aligned} \tag{31}$$

where μ_1 is the first term in μ. Interchanging differentiation and integration yields

$$\frac{\partial}{\partial\mu_1}\mathbb{E}\left[f\left(X\right)\right] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} g\left(z\right) \frac{z-\mu_1}{\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left(z-\mu_1\right)^2\right) \tag{32}$$

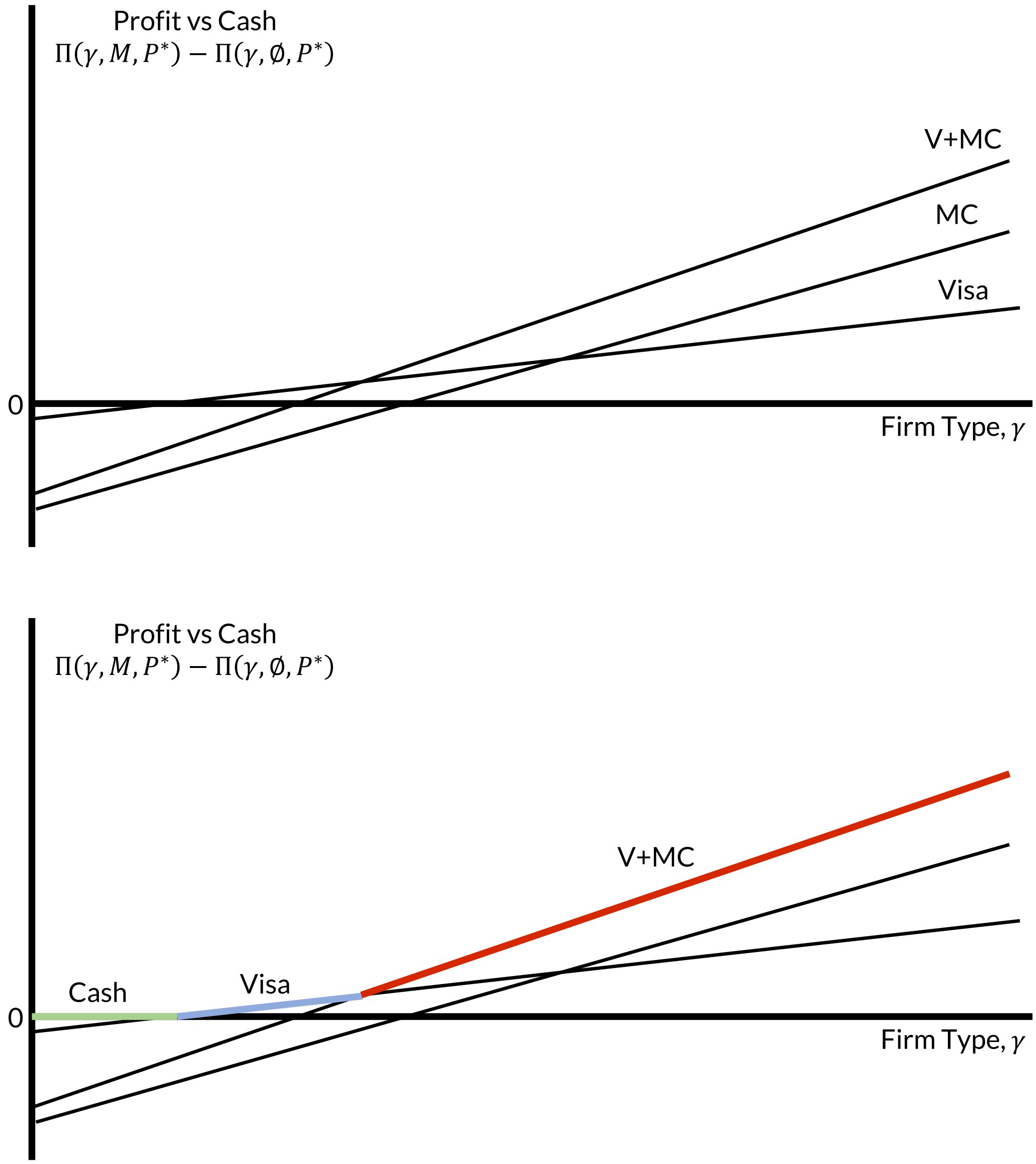
Equations 31 and 32 provide integral expressions for the expectation and the derivative of the expectation. To approximate these expectations, one can simulate g with standard monte carlo techniques as $\hat{g}$. While $\hat{g}$ will not be differentiable, by the convolution theorem expressions 31 and 32 will both be differentiable even if g is replaced by $\hat{g}$. The remaining integral can then be calculated efficiently by Gauss-Hermite quadrature.

E Quasiprofits

E.1 Example of Calculating the Equilibrium

Figure E.1 shows an example of computing an equilibrium when Visa charges merchants low fees but has a low market share among consumers, MC charges high fees and has a high market share, and cash is free. At $\gamma = 0$, because cards cost more than cash, all of the quasiprofit functions for bundles M that include cards are less than the quasiprofit for cash. Therefore merchants with low benefit parameters γ choose to only accept cash. However, because Visa's fee is lower, its y-intercept is closer to zero and its quasiprofit function crosses zero first. The crossing point marks the start of a region of merchants who only accept Visa. When the quasiprofit function for the combination of Visa plus MC exceeds the quasiprofit function for Visa, all merchants of that type or higher will then accept both.

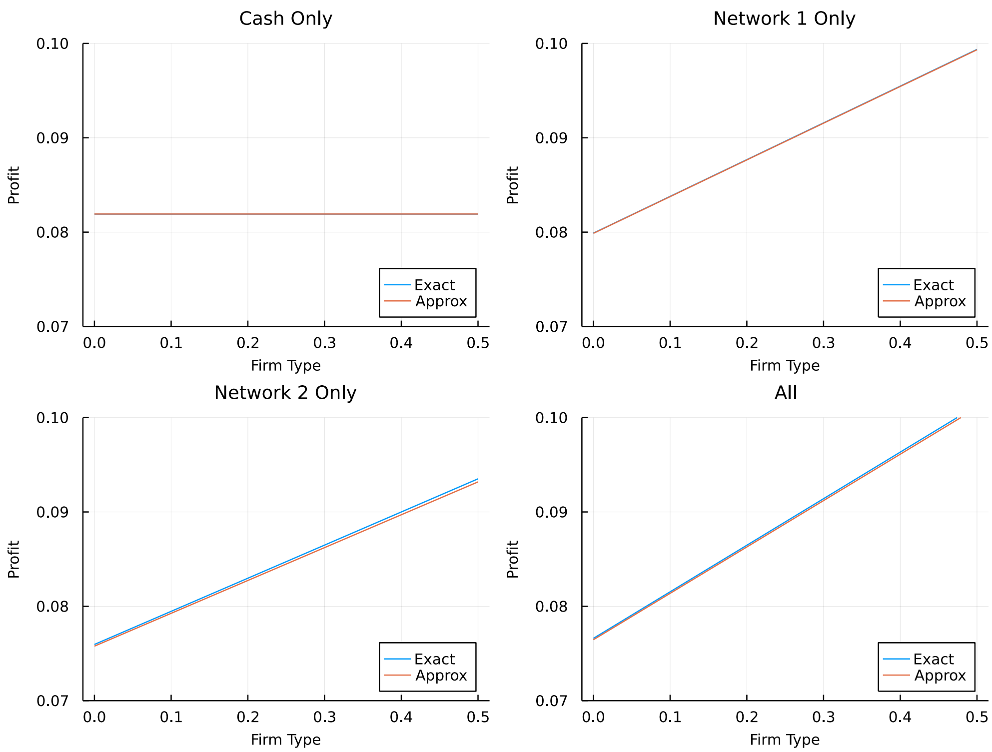
Figure E.1: Illustration of how to compute the merchant adoption subgame.



E.2 Quality of Approximation

A natural question is whether the quasiprofit functions are a good approximation of true profits. Figure E.2 compares exact and approximate profits in a case with two networks with symmetric market shares, differentiated only by the two networks charge different fees. The fit is very close for all values of the merchant type γ.

Figure E.2: Numerical example of how quasiprofit functions approximate true profit functions for a case of two networks with symmetric consumer parameters but who set merchant fees of $\tau_1 = 0.02$ and $\tau_2 = 0.04$



E.3 Comparison with Rochet and Tirole (2003)

The linearity of quasiprofits also reveals how the extent to which consumers hold one card or two will shape merchants willingness to substitute between accepting different cards, as in (Rochet and Tirole, 2003).

Consider a simplified economy in which consumers pay with cash and two cards, Visa (v) and American Express (a). Visa and American Express charge merchant fees of $0 < \tau_v < \tau_a$. Let the insulated shares be μ. Then the merchant adoption equilibrium will feature three regions:

1. Merchants of types $\gamma \in \left[0, \frac{\sigma\tau_v}{1-\sigma\tau_v}\right]$ accept only cash
2. Merchants of type $\gamma \in \left[\frac{\sigma\tau_v}{1-\sigma\tau_v}, \frac{\mu^{a,v}(\tau_a-\tau_v)+\mu^{a,0}\tau_a}{-\sigma\mu^{a,v}(\tau_a-\tau_v)+\mu^{a,0}(1-\sigma\tau_a)}\right]$ accept Visa only, where $\mu^{a,v}$ is the insulated share of consumers who primarily use American Express but who also have a Visa, and $\mu^{a,0}$ is the insulated share of consumers who only have an American Express and do not have a Visa.
3. Merchants of type $\gamma > \frac{\mu^{a,v}(\tau_a-\tau_v)+\mu^{a,0}\tau_a}{-\sigma\mu^{a,v}(\tau_a-\tau_v)+\mu^{a,0}(1-\sigma\tau_a)}$ accept both

When many American Express holders carry Visa, then $\mu^{a,v}$ is large and fewer merchants will accept American Express if Visa charges a low fee. Merchants become unwilling to accept American Express because doing so would force the merchant to raise higher prices, lowering demand, while getting few additional sales. When fewer merchants accept American Express, Visa is better off and so Visa has strong incentives to compete for merchants if most American Express consumers hold Visa cards. In contrast, if no American Express users carry a Visa, then $\mu^{a,v}$ is zero and the lowest type merchant who accepts American Express is $\frac{\sigma\tau_a}{1-\sigma\tau_a}$. In this case, the set of merchants that accepts American Express no longer depends on the fees that Visa charges. This would dramatically weaken Visa's incentives to compete for merchants.

F A Microfoundation for Interpreting Co-holding Data as Hypothetical First and Second Choices

This note outlines a microfoundation by which consumers' secondary cards can be used to identify hypothetical second choices for primary card. I assume consumers have wallets with two cards: a primary card and a secondary card. The consumer usually uses the primary card and with some small probability uses the secondary card. Periodically, consumers re-assess their primary card and choose primary cards of different brands with some probabilities. If the brand of the primary card changes, the consumer then downgrades the existing primary card to secondary status, and the new card becomes the primary card. The conditional distribution of the secondary card conditional on the brand of the primary card will then have the same distribution as second choices for primary cards conditional on the primary card. In other words, the fact that Visa cards are often found in wallets of primary Amex users will mean that Visa is a close substitute for Amex.

F.1 Environment and Proof

Let time be discrete $t = 1,2,\ldots$. For consumer i at time t, suppose that the utility from choosing a card $j \in \{1,\ldots,n\} \equiv J$ is

$$u_{ijt} = \delta_i + \epsilon_{ijt}$$

Suppose her wallet at time t contains two cards, $w_t = (p_t, s_t)$, where $p_t \in J$ is the primary card and s_t is the secondary card. Then at time $t+1$, the consumer draws new utilities and chooses a new primary card $p_{t+1} \in J$ that yields the highest utility. If $p_{t+1} = p_t$, then the wallet does not change and $w_{t+1} = w_t$. Otherwise, the new primary card changes, and then the new secondary card $s_{t+1} = p_t$, and so $w_{t+1} = (p_{t+1}, s_{t+1})$.

Theorem 2. *The joint stationary distribution of w_t is the same as the joint distribution of first and second choices, that is*

$$P\left(\left(u_{ijt} = \max_{l\in J} u_{ikt}\right) \cap \left(u_{ikt} = \max_{l\in J\setminus\{j\}} u_{ilt}\right)\right) = P\left(p = j, s = k\right)$$

Proof. Fix i. The probability of choosing j is

$$q_i\left(j\right) = \frac{\exp\left(\delta_i\right)}{\sum_{l\in J}\exp\left(\delta_l\right)}$$

The joint distribution of first and second choices comes from a standard result on logit choice probabilities:

$$P\left(\left(u_{ijt} = \max_{l \in J} u_{ikt}\right) \cap \left(u_{ikt} = \max_{l \in J \setminus \{j\}} u_{ilt}\right)\right) = q_i(j) \times \frac{q_i(k)}{\sum_{l \neq j} q_i(l)} \tag{33}$$

Next we calculate the joint stationary distribution of the wallets w_t. Denote this stationary distribution with P_i. Fix the wallet $w_{t+1} = (p_{t+1}, s_{t+1})$ at time $t+1$. For this to have occured, there are two possibilities for the wallet at time t. In the first case, the wallet did not change and $w_{t+1} = w_t$. This happens with probability $q_i(p_{t+1}) P_i(w_{t+1})$. In the second case, a new primary card was chosen at time $t+1$ such that the primary card is p_{t+1} and the secondary card was s_{t+1}. This happens with probability

$$\begin{aligned} q_i(p_{t+1}) \sum_{k=1}^{n} P(w_t = (s_{t+1}, k)) &= q_i(p_{t+1}) q_i(s_{t+1}) \sum_{w_{t-1} \in S^2} P_i(w_{t-1}) \\ &= q_i(p_{t+1}) q_i(s_{t+1}) \end{aligned}$$

We can then drop time subscripts, and the stationary distribution P_i must then be determined by

$$\begin{aligned} P_i(w) &= q_i(p) P_i(w) + q_i(p) q_i(s) \\ P_i(w) &= \frac{q_i(p) q_i(s)}{1 - q_i(p)} \\ &= q_i(p) \times \frac{q_i(s)}{\sum_{l \neq p} q_i(l)} \end{aligned}$$

Which is the same as 33. Conditioning down on i then gives the desired result. $\square$

F.2 Discussion

This works because an IIA assumption holds conditional on i. For a given i, if a particular card p is the primary card, then the probability a different card is the second choice is determined by just dividing the probabilities.

The assumption that the primary card changes only if the new primary card is a different brand helps map the thought experiment to my empirical work. In my empirical work, the secondary card counts any card brand with any amount of positive spending. Therefore if a Visa/Mastercard multihomer decides to add a new Visa card to her wallet, as long as she puts some positive spending on Mastercard I will count her secondary card as Mastercard. Therefore adding a new card does not change primary/secondary status if the new card has the same brand as the old primary card.

A key behavioral assumption is that the consumer does not choose the new primary card based on the characteristics of the secondary card. This is reasonable at the brand level. This is primarily because variation in merchant adoption of payments is primarily vertical, rather than horizontal. While different cards might offer different rewards, Visa, Mastercard, and American Express all offer rewards programs across major sectors. Therefore I'm assuming that if a consumer has a travel card, and is looking to get a card with grocery rewards, then her preferences over brands would be the same as if she already had a grocery card and is looking to get a travel card.

The microfoundation means that if a consumer has a primary Visa credit card card and a secondary Mastercard credit card as a secondary card, then in an alternative world where Visa did not exist the consumer would have chosen to use a Mastercard credit card as their primary card. This is only a reasonable assumption if Visa and Mastercard credit cards offer similar services. For example, if one thought that Visa and American Express were primarily accepted on the west coast and that Mastercard was primarily accepted on the east coast, then a west coast person who travels occasionally to the east coast might choose to have a primary Visa card and a secondary Mastercard card to cover their needs. If Visa did not exist, such a person would likely choose American Express as the primary card to cover their west coast needs.