Interchange Fees in Payment Networks: Implications for Prices, Profits, and Welfare*

Robert M. Hunt† Konstantinos Serfes‡ Yin Zhang§

March 26, 2026

Abstract

This paper develops a two-sided model of the payment card market, incorporating elastic consumer demand, merchant and network market power, *ad valorem* interchange fees, cardholder rewards, and cash as an alternative payment method. Building on principles from public finance, we introduce the concept of a *credit card tax*—an endogenous wedge between consumer and merchant prices determined by three key factors: (i) interchange fees, (ii) rewards and (iii) the rate of credit card adoption. We analyze how this tax influences equilibrium prices, platform profits, and social welfare. Our results yield an unexpected and policy-relevant finding: contrary to conventional wisdom, capping interchange fees can *increase* equilibrium rewards when consumer demand elasticity is relatively low. This regulatory intervention can improve consumer welfare for all buyers, including cash users. Even in cases where rewards decline, the net consumer benefit may remain positive. Furthermore, we establish conditions under which interchange fee caps enhance allocative efficiency and promote desirable payment choice. The analysis provides novel theoretical insights for the regulation of two-sided payment markets.

Keywords: credit cards, two sided networks, merchant competition, interchange fees, regulation
JEL codes: L13, L40, G28, E42

(Preliminary and Incomplete For Discussion Purposes Only)

*We are grateful to Konstantinos Charistos, Julia Cheney, Pedro J. Cuadros-Solas, Gabe de Bondt, Bernhard Ganglmair, Germain Gaudin, Christos Genakos, Fumiko Hayashi, Eugenio Miravete, Arina Nikandrova, Chris Phillips, Patrick Rey, Oz Shy and Zhu Wang for helpful comments and discussions. We also thank seminar and conference participants at the 2023 Economics of Payments Conference (Federal Reserve Board), Oligo 2024, Dialogues in Competition Law Economics 2024, CRESSE 2024, CRETE 2024, SPR 2024, EEA-ESEM 2024, MaCCI 2025, ICMAIF 2025, EFMA 2025, the American Bankers Association, the Federal Reserve Bank of Philadelphia, the University of Cambridge and ZEW–University of Mannheim for valuable feedback. All errors remain our own. **Disclaimer:** The views expressed in this research paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve Board of Governors, or the Federal Reserve System. Nothing in the text should be construed as an endorsement of any organization or its products or services. Any errors or omissions are the responsibility of the author. No statements here should be treated as legal advice.

†Federal Reserve Bank of Philadelphia, PA 19106 Philadelphia. e-mail: bob.hunt@phil.frb.org.
‡School of Economics, Drexel University, Philadelphia PA 19104. E-mail: ks346@drexel.edu.
§Bates White, Washington DC 20006. E-mail: yin.zhang@bateswhite.com.

1 Introduction

This paper examines the equilibrium level of the wholesale price, the interchange fee, set by payment card networks. This fee fundamentally influences the distribution of acceptance costs and benefits between merchants and consumers. In the United States, these fees typically range between 2% and 3% of the transaction value.[1] Interchange fees have been the subject of intense debate for half a century, particularly since payment cards became the dominant method of payment at the point-of-sale in the 1970s.

On one side of this debate, merchants contend that interchange fees are excessively high, eroding their profit margins and compelling them to pass these costs on to consumers through higher retail prices.[2] This perception has spurred regulatory interventions. In the United States, the Durbin Amendment to the Dodd-Frank Act capped debit card interchange fees for large issuers, mandating that they be "reasonable and proportional" to processing costs.[3] More recently, the proposed Credit Card Competition Act of 2023 aims to increase network competition for credit cards, which could exert downward pressure on fees.[4] Internationally, jurisdictions like the European Union,[5] Australia, and Canada have implemented similar measures, including direct fee caps, to address perceived anti-competitive elements in retail payment systems.

Conversely, card networks and issuing banks argue that such regulations may undermine the economic ecosystem that funds consumer benefits, most notably credit card rewards programs.[6] These rewards, which average approximately 1.45% of transaction value according to Wang (2023),[7] are typically structured to incentivize greater card usage. The core tension, therefore, lies in balancing the goal of reducing costs for merchants against the risk of diminishing the rewards and associated benefits for consumers.

Complicating the analysis is that payment markets represent a classic example of a two-sided platform, where a network intermediates between cardholders and merchants. In such markets, pricing de-

[1] Interchange fees account for most of the merchant discount (on average about 2.25%, according to Wang (2023)) paid by merchants to their acquiring banks for processing card transactions.

[2] See, for example, the recent settlement: thehill.com/business/4556451....

[3] For regulatory details, see: Federal Register (2023).

[4] See S.1838 - Credit Card Competition Act of 2023.

[5] Regulation (EU) 2015/751. Available at: EUR-Lex.

[6] The industry's stance is captured in campaigns such as "Congress: Hands off my rewards!" (Electronic Payments Coalition, 2022).

[7] For a discussion on the potential impact on rewards, see: American Banker (2023).

cisions affecting one side (e.g., cardholder fees or rewards) critically influence participation and pricing decisions on the other side (e.g., merchant acceptance and retail prices). This interdependence is central to understanding the role and consequences of interchange fees.

To address this complexity, we develop a comprehensive model of a two-sided payment card market. Our model endogenizes key features including the interchange fee, cardholder rewards, merchant pricing and the choice of payment instrument (card versus cash). We consider a setting with a monopoly payment network and, to the best of our knowledge, provide the first analysis that jointly incorporates the following realistic elements—each extensively discussed in the subsequent literature review: (i) elastic consumer demand for goods, (ii) merchants possessing market power, (iii) ad valorem (percentage-based) interchange fees and rewards and (iv) the presence of cash as an alternative payment method, which constrains the market power of the card network. This framework enables a rigorous assessment of interchange fee cap regulations. By incorporating more realistic assumptions about market structure and agent behavior, the model yields predictions that align more closely with observed market outcomes than prior, more stylized analyses.

We analyze several questions, with a central focus on three core issues: how interchange fees influence the distribution of costs between merchants and consumers, how networks strategically adjust rewards when fees are capped and whether such caps improve welfare by reducing market inefficiencies and improving payment choices. Alongside these, we also explore related topics such as the impact of merchant and network market power and the conditions under which card users may be effectively subsidized.

We parameterize the retail sector along three dimensions: the (constant) elasticity of demand (k), the number of firms (n), and the conjectural variation (λ) that retailers use when forming expectations about their competitors' responses to quantity decisions.[8] The number of retailers (n) and the conjectural variation (λ) jointly determine the conduct parameter (γ) that characterizes the intensity of competition in the final goods market. This setup allows us to analyze how consumer preferences, captured by the demand elasticity and the degree of differentiation between cash and credit (which affects the network's market power), interact with merchant market power to endogenously determine retail prices, interchange fees and the consumer response to rewards. These outcomes, in turn, affect aggregate consumer demand

[8] As an extension (Section 6.4), and to capture the effect of non-constant elasticity, we also analyze a linear demand specification. Relative to the constant elasticity case, this introduces an additional effect: the *elasticity effect*, whereby aggregate consumption influences the elasticity and, consequently, the price.

and feed back into retail pricing.

The model establishes an equilibrium in which aggregate demand and merchant profits are shaped by what we term "credit card taxes"—a central conceptual contribution of this paper. We emphasize that we do not use this terminology in a normative sense, as if such taxes were inherently undesirable. Rather, our intent is to highlight the close analytical parallel between the economic mechanisms identified here and those found in the theoretical literature on taxation. These "taxes" introduce a wedge between the price paid by consumers, which differs depending on whether they use cash or credit, and the price received by merchants for each unit of the final good. This wedge reflects the network's revenue per dollar of transaction and embodies the distributional consequences of the payment card system.

The magnitude of these credit card taxes is positively associated with interchange fees and the rate of credit card adoption by consumers, yet negatively associated with the generosity of cardholder rewards. These taxes form endogenously through the linkage of product markets with card network pricing and reward structures–a significant interaction that most of the existing literature has failed to address.

Our analysis of a regulatory cap on interchange fees begins by examining payment network's strategic setting of cardholder rewards. This focus is critical due to the dual impact of rewards. First, they directly affect the effective price credit card users pay. Second, they indirectly influence market equilibrium by driving credit card adoption: higher reward implies a higher fraction of consumers who choose to use the credit card. Under the common no surcharging rule, higher adoption increases the proportion of sales subject to the merchant fee, thereby raising the merchants' average cost of a transaction. This elevated cost is ultimately passed through to *all* consumers in the form of higher retail prices.

We find that the regulation's impact on equilibrium rewards is dictated by demand elasticity. While an interchange fee cap regulation unambiguously reduces the profitability of the payment network, its strategic response is less straightforward. Whether the network increases or decreases equilibrium rewards depends on the nature of the relationship between interchange fees and rewards in its profit function–specifically, whether they are substitutes or complements. In the substitutable case, a mandated (modest) reduction in the interchange fee raises the marginal profitability of offering rewards. Conversely, when the two are complementary, the same reduction lowers the marginal profitability of rewards.

This classification is shaped by the elasticity of product demand. When demand elasticity is relatively low, the reduction in the interchange fee does not significantly stimulate aggregate consumption. Under

such conditions, the network finds it profitable to increase rewards, partially offsetting lost revenue by encouraging additional spending, which occurs both through increased purchases by existing cardholders and by enticing cash users to adopt credit. By contrast, when demand elasticity is relatively high, the regulation leads to a substantial increase in consumption, enabling the network to recoup some of its lost revenue by reducing rewards while still benefiting from the resulting growth in transaction volume.

In policy discussions surrounding interchange fees for credit cards, it is often taken as axiomatic that a reduction in interchange fees leads to a reduction in consumer rewards. This implication is also embedded in most theoretical models, where the structure of the model either mechanically links lower interchange fees to lower rewards, or key forces are absent, e.g., an endogenous credit card adoption by consumers, as we further discuss in the Literature Review Section. However, our analysis reveals a novel result: this relationship holds only when demand is relatively elastic. Under less elastic demand, a reduction in interchange fees can, in fact, lead to an increase in rewards, challenging the conventional wisdom and offering new insights into the welfare implications of fee regulation.[9] Importantly, this result is not an artifact of assuming a constant elasticity demand; it holds under linear demand as well.

Next, we examine how the regulation affects the taxes paid by consumers, beginning with cash users. Two forces are at play, which may either offset each other or move in the same direction. On the one hand, lower interchange fees exert a direct downward effect on the tax. On the other hand, there may be an indirect upward effect if credit card adoption rises, which can occur when product demand is relatively inelastic. In that case, as discussed earlier, equilibrium rewards increase, potentially intensifying the cross subsidy embedded in retail prices. The net effect depends critically on the demand structure. Under constant elasticity demand, the direct effect dominates, so the tax unambiguously falls. Under linear demand, however, the indirect channel outweighs the direct one, if and only if demand elasticity is relatively low, leading to a higher tax for cash users. In contrast, when product demand is relatively elastic, both forces work in the same direction: equilibrium rewards and card adoption both decline, reducing the cross subsidy and leading to a clear reduction in the tax paid by cash users. Although equilibrium rewards do not enter the cash user tax directly, they influence it indirectly through their impact on card adoption.

[9] There is some empirical evidence that interchange fee caps do not uniformly reduce credit card rewards. While the Reserve Bank of Australia's reforms were expected to diminish the generosity of rewards programs, banks responded in varied ways. Some introduced new premium cards with higher interchange fees to support more generous rewards, while others enhanced rewards on existing products–particularly through the addition of American Express companion cards, which increased reward rates without raising annual fees (Chan et al. (2012)).

The tax paid by credit card users–over and above the interchange fee and the cost of card adoption–also depends on the reward. Our analysis shows that credit card users consistently experience a reduction in this tax following a regulatory cap. This result holds under both the constant elasticity and linear demand specifications we consider. When equilibrium rewards rise, cardholders gain directly. Even when rewards fall, credit card users still enjoy a net benefit: the positive effects of a lower merchant fee and reduced card adoption outweigh the negative impact of the lower reward.

What really matters for consumer welfare is the price consumers pay. This price is jointly determined by the taxes and the merchant's pricing decision. Under constant elasticity demand or perfect competition, the merchant price does not change in response to the regulation; consequently, any reduction in the tax is fully passed through to consumers in the form of a lower price. These assumptions–constant elasticity demand or perfect competition–therefore serve as useful benchmarks. They not only simplify the analysis, but also allow us to isolate and understand the various effects one at a time.

When we allow for linear demand and imperfect competition, an additional force emerges: the elasticity effect. This introduces imperfect pass-through, meaning that changes in taxes are not fully transmitted to consumer prices. As a result, the interaction between market power, pricing behavior and tax incidence becomes more nuanced than in the benchmark cases. Under a cap regulation that successfully lowers the tax, merchants respond by increasing prices. In some cases, this leads to cash users facing higher prices, while credit card users always benefit by paying a lower price.

As discussed, a primary argument advanced by banks against interchange fee regulation is that it would reduce credit card rewards, thereby harming cardholders. However, our analysis reveals a more nuanced reality. First, such caps can, in some cases, lead to an increase in equilibrium rewards. Second, this very increase can indirectly harm cash users. Finally, and perhaps most critically, even when rewards decrease, the combined effects of the regulation, including lower merchant fees and lower credit card adoption, can result in a net benefit for all consumers.

Under constant elasticity demand, a cap on interchange fees also benefits merchants, except in the limiting case of perfect competition in the product market (where merchant profits are zero). This is because the regulation lowers the equilibrium taxes, and given that the price merchants receive is unaffected by the taxes (due to the constant elasticity demand specification), aggregate consumption increases and hence merchant profits. Consequently, the regulation redistributes surplus from the payment network to

both consumers and merchants. In contrast, under a linear demand specification with perfect competition among merchants and relatively low demand elasticity, the surplus transfer originates from the payment network and *cash users*, flowing exclusively to *cardholders*. When demand elasticity is relatively high, the redistribution of surplus aligns with that under the constant elasticity demand. Under a linear demand and imperfect competition, the regulation redistributes surplus away from the network, and sometimes cash users, towards credit card users and merchants.

This work identifies the key economic forces triggered by an interchange fee cap and evaluates their relative strength under two tractable demand systems: constant elasticity and linear. It is important to note that the precise balance of these forces may change under a different demand specification. Despite this, the analysis reveals a crucial insight arising from the positive relationship between rewards and credit card adoption: the regulation can benefit both consumer groups and it is much less likely to harm both. If credit card users are made worse off, such as when rewards decrease significantly, the resulting market conditions would benefit cash users. Conversely, in scenarios where cash users are worse off, typically when rewards increase, credit card users are better off.

Our model also identifies three key sources of social inefficiency in the credit card payment system: network market power, merchant market power and reward-induced distortions in consumer payment choices. These factors can lead to excessive credit card usage and suboptimal consumption levels. We show that an interchange fee cap can mitigate both inefficiencies when demand elasticity falls within an intermediate range. In this case, the regulation reduces the implicit taxes paid by both cash and credit card users, improves consumption efficiency, and curbs overuse of credit, thereby enhancing overall social welfare.

Schuh et al. (2010) attempt to quantify the size and distributional consequences of implicit monetary transfer to credit card users from non-card, or cash, users that result from interchange fees and card rewards. They show that, on average, each cash-using household pays $149 to card-using households, and each card-using household receives $1,133 from cash users every year. Their calculations suggest this is a regressive transfer from low-income to high-income households. Using more recent data Felt et al. (2023) estimate that the lowest income cohorts in the U.S. (Canada) are paying 20 (60) basis points, respectively, more than the highest income cohorts in order to conduct their purchases.[10]

[10]The calculations in those papers are derived using varying assumptions about the extent of pass through of card acceptance costs to consumers. Pass through is not formally modeled or estimated.

While our model does not explicitly incorporate income heterogeneity, the resulting equilibrium reflects a similar dynamic: cash users effectively face a higher implicit tax than credit card users due to the rewards associated with credit card transactions.

However, the term subsidy is often used imprecisely in the literature. Within our framework, depending on the elasticity of demand, the equilibrium reward may be sufficiently large that credit card users pay a price lower than the amount received by merchants–effectively transforming their tax burden into a subsidy. In some cases, credit card users may even pay a price below marginal cost, leading to socially inefficient outcomes. Our analysis clarifies the conditions under which credit card users are genuinely subsidized, rather than simply benefiting from a lower effective tax than cash users.[11]

Our analysis predicts how merchant market power influences equilibrium taxes and merchant profits. Under a constant elasticity demand, as the product market becomes more competitive, the equilibrium tax for credit card users decreases. For cash users, the equilibrium tax decreases if only if the product demand elasticity is sufficiently low. Aggregate merchant profits exhibit an inverse U-shape relative to merchant market power. Interestingly, there is a region where increased competition among merchants leads to higher equilibrium profits. This occurs because the network, which competes with cash, reduces the tax for credit card users as product market competition intensifies, thereby encouraging higher consumption levels.

Under linear demand, however, competition has the opposite effect on equilibrium taxes. This is due to the elasticity effect, which makes the network more reluctant to reduce taxes as the product market becomes more competitive.

We also derive predictions regarding how consumer preferences for credit versus cash as payment instruments influence equilibrium fees, rewards, and taxes. As the two payment methods become closer substitutes, reducing the network's market power, the effective tax paid by credit card users increases, primarily due to diminished rewards. The tax burden on cash users declines, but only when product demand is relatively elastic; otherwise it rises because interchange fees rise. As anticipated, equilibrium network profits decrease, but merchant profits also decline.

[11] An important qualification to the distributional implications discussed in Schuh et al. (2010) and Felt et al. (2023) is the omission of a potential relationship, both in existence and direction, between interchange fees and credit supply or pricing. While our paper does not examine this possibility directly, it is addressed in other work, which we summarize in the following section.

The remainder of the paper is organized as follows. Section 2 provides a concise review of the relevant literature. Section 3 introduces the core structure of the model and Section 4 presents the equilibrium analysis. Section 5 investigates the socially efficient outcome and evaluates the effects of imposing a cap on interchange fees. In Section 6, we consider four extensions: (i) the implications of differential benefits or costs between cash and credit transactions for consumers, showing that higher benefits associated with credit (relative to cash) lower the network's equilibrium profit; (ii) the impact of prohibiting rewards, demonstrating that banning rewards can make cash users worse off (depending on the demand elasticity); (iii) the consequences of costly cash usage, where an illustrative numerical example shows that if cash entails a transaction cost, the introduction of credit can lower the effective "tax" for the majority of consumers (i.e., credit card users); and (iv) linear demand in the product market, to assess the impact of a non constant elasticity. Finally, Section 7 concludes.

2 Literature review

Over the last three decades, an extensive theoretical literature on two-sided markets has emerged and many of those papers examine the implications for payment systems; see Rochet and Tirole (2006), Rysman (2009) and Jullien, Pavan, and Rysman (2021) for reviews of this literature. Our purpose here is simply to sketch the main findings related to our work and highlight our contributions.

Our modeling framework builds upon and extends the work of Wang (2010) and Shy and Wang (2011), both of which adopt a constant elasticity demand specification with ad valorem interchange fees and rewards. In particular, Wang (2010) assumes a perfectly competitive merchant sector and an exogenously fixed distribution of cash and credit card users. Among other findings, his analysis concludes that capping interchange fees reduces rewards (for demand elasticity greater than one—the case we focus on), lowers retail prices, increases consumption, and thereby enhances consumer welfare.

Our analysis yields several important distinctions relative to Wang (2010). First, we show that a cap can in some circumstances lead to an *increase* in rewards, but the equilibrium tax imposed on both cash and credit card users declines (under a constant elasticity demand). This demonstrates that consumer benefits from regulation can persist even in a more realistic setting characterized by merchant market power and endogenous competition between payment instruments. However, we also establish that under a linear demand specification, a cap can redistribute welfare negatively, potentially harming cash users.

By incorporating merchant market power, an endogenous rate of credit card adoption by consumers and departing from constant elasticity demand, our framework generates several key theoretical predictions and policy implications that either fall outside the scope of Wang (2010) or meaningfully diverge from its conclusions.

Shy and Wang (2011) compare 'proportional' and 'fixed' transaction fees in a model where merchants possess market power, but cash is not available as an alternative payment method. Our primary innovation relative to Shy and Wang (2011) is the explicit inclusion of cash as an alternative payment mode. In our framework, each consumer chooses their preferred payment method, allowing the network to influence both the intensive and extensive margins of payment behavior. In the absence of cash, an indeterminacy arises: the interchange fee and reward cannot be uniquely determined.[12] Consequently, without cash, meaningful comparative statics cannot be derived to analyze how product market competition and other key parameters affect the interchange fee and rewards separately. In particular, the impact of interchange fee regulation cannot be assessed. Any cap imposed on the interchange fee becomes ineffective, as the network can adjust rewards to offset the cap, thereby maintaining the effective tax and thus prices and welfare, at their pre-regulation equilibrium levels.

A key feature of our framework is that regulation affects not only interchange fees and rewards, but also the endogenous adoption of credit cards. This adoption margin subsequently shapes the effective costs for both card and cash users. In the models of Wang (2010) and Shy and Wang (2011), where the share of credit transactions is fixed, or where cash is not a viable alternative, this channel is absent. When a payment network can influence not only the intensive margin of usage, but also the extensive margin of adoption, it has stronger incentives to raise rewards. This is because the same increase in rewards stimulates card transactions more substantially when it also expands the base of card users. Consequently, in a model with a fixed adoption rate (Wang, 2010), or one that excludes cash (Shy and Wang, 2011), as our analysis reveals, a fee cap regulation is less likely to increase equilibrium rewards and thus less likely to generate outcomes that harm cash users.

Baxter (1983) developed the first formal model of interchange fees in a payment scheme. In that

[12] Wang (2010) resolves this indeterminacy by assuming that the network (or issuers collectively) faces a convex cost function for processing card payments, which uniquely determines the interchange fee and rewards, even though merchants accepting cards do not serve cash-paying customers. The assumption of convex transaction processing costs for issuers, however, is increasingly outdated. Advances in technology–such as cloud computing, automation, and scalable infrastructure–have significantly reduced marginal costs. Today, issuers can process large volumes of transactions with minimal incremental cost, making convex cost structures less plausible in practice.

paper, Baxter makes three key assumptions: i) issuers and acquirers make no profit (perfect competition), ii) merchants do not use card acceptance strategically, i.e., to attract consumers from rival merchants who do not accept a card and iii) there is no merchant heterogeneity in the benefit of accepting cards. Schmalensee (2002) develops a model that explores the double marginalization problem that emerges when networks set an interchange fee that is paid to card issuers and the acquiring bank for the merchant separately sets a merchant discount for processing transactions.

In Rochet and Tirole (2002), the consumers receive a different benefit from transacting using cards rather than paying cash. There is a single payment network that sets an interchange fee, but it is not ad valorem. The retail market is a Hotelling model, but consumers face the choice of purchasing a fixed quantity of their preferred good. Rochet and Tirole (2011) develop a similar model, but here there is competition between networks modeled using the Hotelling framework. Wright (2004) uses a model similar to Rochet and Tirole (2002) and relaxes all three assumptions of Baxter (1983). Bedre-Defolie and Calvano (2013) show that networks over subsidize card usage and overtax merchants.

Wang and Wright (2017, 2018) assume Bertrand competition among sellers in a market with many different goods that vary widely in their costs and values. The authors show that ad valorem fees and taxes represent an efficient form of price discrimination relative to uniform fees that disadvantage low-cost, low-value goods. Wang (2023) develops a structural empirical approach to a two-sided market of payments. He finds that interchange fee caps increase welfare by reducing rewards, retail prices, and credit card use. Overall consumer welfare increases because low-income consumers, who rely less on credit, pay lower prices, but high-income consumers get hurt because their rewards decrease. In our model, while there is no income heterogeneity, we predict that any reduction in rewards will not completely offset the interchange fee reduction due to a cap.

Edelman and Wright (2015) shows that an intermediary always chooses to impose price coherence if it has the ability to do so. In the context of payments this is the equivalent of the "no surcharge" rule that precludes merchants from charging customers a different price based on how they choose to pay. An important qualification is that the fees examined in that paper are not ad valorem.

The prevailing methodology in most of the above papers includes one or more of the following assumptions: (1) consumer demand is inelastic, (2) fees are specific rather than ad valorem, (3) Bertrand competition is assumed in the product market, (4) cash transactions are not allowed, or fraction of credit

card users is exogenous, and (5) the benefits and costs experienced by both merchants and consumers due to card usage are independent of each other and of the product price, implying that interchange fees and rewards do not affect retail prices. While these simplifications facilitate analysis, they diverge significantly from real-world observations. Our study aims to bridge this gap by offering a model that encompasses many realistic features of the credit card market.

Nevertheless, and in order to provide a tractable framework for analyzing our core mechanisms, we adopt the following two assumptions. First, we assume perfect competition among banks (Wang, 2010; Shy and Wang, 2011), which isolates the role of the card network's fee structure from effects of banking market power. Second, we model a mature market (Wang, 2010) where all merchants accept cards, allowing us to examine the endogenous outcomes, such as consumer payment choice and retail price adjustments, that follow widespread adoption, rather than the adoption decision itself.

The available empirical research on the effects of payment card regulation documents that reductions in credit card interchange fees are generally passed on to merchants and that generosity of rewards falls, but there is also evidence to the contrary (Chan et al. (2012) for the example of Australia). The evidence of a pass through of merchant cost savings into the prices consumers pay is mixed. For the U.S., this research has focused on the effects of the caps on interchange fees charged by large debit card issuers implemented by Reg ii. Mukharlyamov and Sarin (forthcoming) explore whether gasoline prices fell in geographies where gasoline retailers were relatively more exposed to reductions in debit interchange. They could not quantify a measurable effect. One aspect of of the Reg ii cap was that interchange fees actually rose for small ticket purchases. Wang et al. (2014) report the result of a survey of merchants which showed a higher proportion of merchants experienced an increase in debit acceptance costs as opposed to a decrease and were more likely to increase prices than reduce them.[13] More general surveys of the pass through of cost or tax changes in different markets find intermediate values of pass through (see the Appendix to Evans and Mateus (2011)).

Although not explicitly modeled in this paper, it is plausible that interchange fees and rewards are influenced by the incentives merchants and lenders face in extending credit to consumers. For instance, Chakravorti and To (2007) demonstrate that retailers are willing to accept higher merchant discounts when the proportion of illiquid consumers rises. Similarly, Chatterjee et al. (2025) develops a general

[13] It should be noted that the modal experience reported by merchants was no change in debit acceptance costs and no change in the prices they charged.

equilibrium framework in which interchange fees and rewards contribute to expanding credit limits and lowering interest rates.

In our model there are two means by which the consumers incur costs for payment services–either embedded in retail prices or fees (i.e. negative rewards). In practice there are other costs consumers may incur. For example, one apparent consequence of the introduction of interchange fee caps for debit card transactions in the U.S. was an increase in fees charged on checking accounts and a reduction in the availability of free checking (Manuszak and Wozniak (2017), Kay et al. (2018), Mukharlyamov and Sarin (forthcoming)).

3 Structure of the market

We consider an industry consisting of n firms (merchants), $j = 1, ..., n$, producing a single homogeneous product. The output of firm j is denoted by x_j and the industry output by $X = \sum_{j=1}^{n} x_j$. All the merchants have the same cost structure $C(x) = cx$, where $c > 0$ is a constant marginal cost. The consumer price is given by an inverse demand function $P(X)$, with derivative $P_X(X) < 0$ and elasticity $\varepsilon \equiv \frac{P}{XP_X} < 0$. Consumers, when purchasing goods, can use either cash or a credit card.

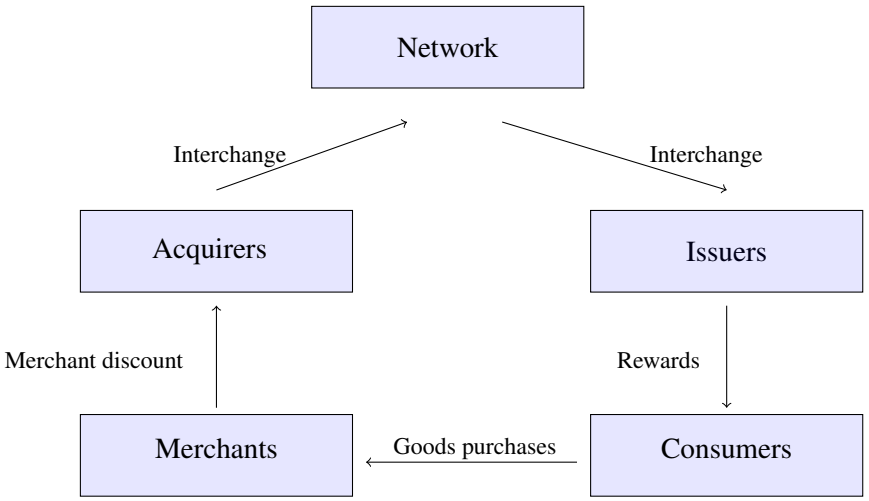


Figure 1: Payment flows in the Network

Payment systems operate on a two-sided model: each time a credit card is used, the merchant incurs a fee, while the consumer may benefit from a reward. We assume that there exists a monopoly payment

network denoted by $\mathcal{N}$, with $N_A < n$ acquiring banks and $N_I < n$ issuing banks.[14] Banks are homogeneous and compete à la Bertrand for merchants and cardholders. In Figure 1, we present the payment flows in the network. Each acquiring bank α chooses its merchant discount m^α to attract merchants and each issuing bank ι chooses the reward $R^\iota \in [\underline{r}, 1]$ to attract users/consumers and to influence the value of consumer transactions. If $R^\iota < 0$, then the reward becomes a fee. The network chooses the interchange fee, $i \in [0, 1]$, which becomes each acquirer's marginal cost. Part of i, denoted by r, goes to the issuing banks to fund the rewards and the rest is kept by the network.[15] For simplicity, all other fees and costs to process a transaction are assumed to be zero.[16]

The reward for the credit card is a percentage of the value of the transaction that a consumer who uses the credit card receives as cashback from the issuing bank. The acquiring bank charges a merchant discount fee, which is a percentage of the value of the transaction that is paid by the merchant to the acquiring bank when a consumer uses the credit card. We assume that the network cannot price discriminate across banks, i.e., all acquirers pay the same interchange fee to the network and all issuers receive the same fraction of the interchange fee from the network.

Consumers exhibit horizontal preferences between cash and credit, as discussed further in Section 4.4. Network fees and rewards affect the value of transactions within the payment network and can incentivize switching between the two payment modes. We assume that merchants do not impose surcharges, meaning they cannot differentiate prices based on the payment method.[17] Furthermore, we consider a mature card market in which the extensive margin of card adoption is largely saturated (e.g., Wang (2010)). As a result, all merchants are assumed to accept both cash and credit.

We analyze a five-stage game with simultaneous and independent moves in each stage. The timing is as follows.

- Stage 1: The payment network sets the interchange fee i and determines the portion r of the fee that is passed to each issuing bank.

[14] American Express operates as a closed-loop network, directly setting both merchant fees and consumer rewards. In contrast, open-loop networks such as Visa and MasterCard influence merchant and consumer pricing indirectly through adjustments to interchange and network fees. Our modeling framework aligns with the structure and pricing mechanisms of these open-loop networks.

[15] Equivalently, the issuing banks receive the interchange fee and pay the network a network fee f, so that $r = i - f$.

[16] See Section 6.3 for an illustration of the implications of a non-zero cost to merchants of accepting cash.

[17] Foster et al. (2024) find that surcharging remains relatively uncommon.

- Stage 2: Each acquiring bank sets its merchant discount m^α and each issuing bank sets its reward R^ι.
- Stage 3: Each merchant chooses its output level x_j, anticipating the share of consumers who will use credit cards.
- Stage 4: Each consumer observes the rewards and the equilibrium retail price, then decides whether to use a credit card or cash.
- Stage 5: Consumers make purchases.

We solve for a subgame-perfect Nash equilibrium in pure strategies.

4 Equilibrium analysis

Although we solve the game by backward induction, the equilibrium in stage 2 is straightforward. Since acquiring and issuing banks compete à la Bertrand and are homogeneous, their equilibrium behavior is uniquely determined by the zero profit condition, regardless of the subsequent choices of merchants and consumers. We therefore present the stage 2 equilibrium first for expositional clarity.

4.1 Acquiring and issuing banks

Acquiring banks compete in merchant fees m^α. The network interchange fee i is each acquiring bank's marginal cost. Given that acquiring banks are homogeneous and compete for merchants à la Bertrand, each acquiring bank sets the same merchant discount $m^\alpha = i$, for all α. Acquiring banks earn zero profits in equilibrium.

Issuing banks compete in rewards R^ι. Each issuing bank receives from the network part of the interchange fee $r < i$. This is the maximum amount, per dollar of transactions, that each issuing bank can give to users as a reward. (If $r < 0$, then the network fee f issuing banks pay to the network is higher than the interchange fee they receive.) Given that issuing banks are homogeneous and compete à la Bertrand for users, each issuing bank sets the same reward $R^\iota = r$, for all ι (it will be a fee for credit card users if $r < 0$). Issuing banks earn zero profits in equilibrium.

Therefore, in the unique equilibrium, each merchant pays a merchant discount i and each consumer receives a reward r per dollar of transaction with a credit card. The network fee, representing the per-dollar profit earned by the network, is given by $i - r$.

4.2 Consumers

Consumers make two sequential decisions. First, each consumer selects a payment mode; subsequently, they determine their level of consumption. The initial decision determines the fraction of consumers, denoted by μ, who choose to make purchases using a credit card. The determination of μ is discussed in detail in Section 4.4. Consumers who pay with cash face a price of P, whereas those who use a credit card pay a discounted price of $P \cdot (1 - r)$ (if $r > 0$).[18]

There are two goods in the model: a consumption good x and a numeraire good y, the price of which is normalized to one. Each consumer's preferences over these goods are represented by the quasi-linear utility function

$$U = \frac{kx(1/x)^{1/k}}{k-1} + y,$$

where $k > 1$. Consumers maximize their utility subject to a budget constraint, yielding a standard constant elasticity demand function with elasticity $\varepsilon = -k$. The demand for good x by a cash-paying consumer is given by $x_{ch} = \frac{1}{P^k}$, while the demand by a credit card user is $x_{cc} = \frac{1}{(P\cdot(1-r))^k}$.

Aggregate demand is then

$$X \equiv \mu x_{cc} + (1 - \mu)x_{ch},$$

and the corresponding aggregate inverse demand function is

$$P(X, r, \mu) = \left(\frac{1}{X}\right)^{1/k} \left(\frac{\mu}{(1-r)^k} + 1 - \mu\right)^{1/k}. \tag{4.1}$$

4.3 Merchant competition

All merchants accept cash and credit. Also, merchants form rational expectations about credit card adoption μ. We assume that each merchant j expects a fraction μ of its sales to be paid with a credit

[18] We use the symbol $\cdot$ to clearly denote multiplication, as in $P \cdot (1 - r)$, to distinguish it from a functional form such as $P(X)$, particularly when the distinction is not immediately clear from context.

card, while a fraction $1-\mu$ to be paid with cash, i.e., there is no expectation of a systematic sorting of cash or credit users to certain merchants, which is a reasonable assumption given that merchants are identical. Each merchant takes the rewards and the interchange fee as given and chooses its output x_j to maximize (average) profits given by

$$\begin{aligned}\pi_j &= (\mu \cdot (1-i)P(X,r,\mu) + (1-\mu)P(X,r,\mu))x_j - cx_j \\ &= (1-\mu i)P(X,r,\mu)x_j - cx_j,\end{aligned} \tag{4.2}$$

where $P(X,r,\mu)$, the price at the register, is given by (4.1).

In selecting its output each merchant j conjectures that other merchants' responses will be such that $\frac{dX}{dx_j} = \lambda$, the conjectural variation λ being taken as a fixed constant throughout. The case $\lambda = 1$ corresponds to the Cournot conjecture. When $\lambda = 0$, conjectures are 'competitive' and we obtain the Bertrand outcome. When $\lambda = n$, each firm believes that all other active firms will behave exactly as it does: tacit collusion among incumbent firms then being perfect (in the sense that aggregate profits are maximized conditional on the number of firms). It will be assumed throughout that $\lambda \in [0,n]$.[19]

The first order condition of the representative merchant is (omitting arguments)

$$\frac{\partial \pi_j}{\partial x_j} = (1-\mu i)\left(P_X \frac{dX}{dx_j} x_j + P\right) - c = 0. \tag{4.3}$$

Restricting attention to symmetric equilibria, this becomes

$$\begin{aligned}(1-\mu i)\left(P_X X\gamma + P\right) = c \quad &\Rightarrow \quad P \cdot \left(1 + \frac{\gamma}{\varepsilon}\right) = \frac{c}{(1-\mu i)} \\ &\Rightarrow \quad P = \frac{c}{\left(1-\frac{\gamma}{k}\right)} \frac{1}{(1-\mu i)},\end{aligned} \tag{4.4}$$

where $\gamma \equiv \frac{\lambda}{n} \in [0,1]$.[20] As γ decreases, the merchant market becomes more competitive and price approaches marginal cost c (adjusted by the $1-\mu i$). From the merchant's perspective, a lower interchange fee has the same effect as a reduction in marginal cost and therefore the same effect on price.

Since in our case the elasticity of the slope of inverse demand, $E \equiv -\frac{P_{XX}X}{P_X}$, is $\frac{1+k}{k} < 2$, the second

[19] See Seade (1980), Bresnahan (1981) and Delipalla and Keen (1992) for similar modeling frameworks.

[20] Note that γ is similar to the conduct parameter θ in Weyl and Fabinger (2013) and Miklos-Thal and Shaffer (2021). Following most of the literature, we assume that the conduct parameter is not a function of aggregate output. For cases where the conduct parameter is a function of aggregate output, see Miklós-Thal and Shaffer (2021).

order condition, which is $2-\gamma E > 0$, is satisfied. The stability condition requires that $1+\gamma\cdot(1-E) > 0$ (see, for example, Seade (1980) and Delipalla and Keen (1992)), which is also satisfied. From the first-order condition (4.4), a non-negative price-cost margin implies that the elasticity must satisfy $\gamma - k < 0$, which also holds.

From equation (4.4), it follows that the average price received by merchants, P^m, is given by $(1-\mu i)P$. At this point, it is useful to introduce the expressions for the implicit tax consumers face. The tax paid by credit card users is defined as $z_{cc} \equiv \frac{1-r}{1-\mu i}$, implying that the price they pay is $P^{cc} = z_{cc}P^m$. For cash users, the tax is $z_{ch} \equiv \frac{1}{1-\mu i}$, resulting in a price of $P^{ch} = z_{ch}P^m$. An increase in the interchange fee i, or in the share of credit card users μ, raises the tax burden for all consumers. In contrast, a higher reward r reduces the tax specifically for credit card users.

The price merchants receive in the presence of a credit card is the same as the price without a credit card (i.e., using (4.4), P^m is not a function of z_{cc} or z_{ch}),

$$P^m = \frac{ck}{k-\gamma} \geq c. \tag{4.5}$$

Therefore, consumers ultimately bear the full burden of a tax or receive the entire benefit of a subsidy. This outcome is a well known property of constant elasticity demand. It also implies that changes in the taxes directly reflect changes in consumer welfare. Under alternative assumptions about demand, however, a portion of the incidence of a credit card tax could fall on merchants.[21]

Efficiency dictates that $P = c$. We can have $P > c$ either because $\gamma > 0$, or $i > 0$, or both. The first source of inefficiency arises when competition in the merchant market is imperfect. The second source of inefficiency is due to the credit card tax levied by the payment network. Since $i > 0$ (otherwise network profit cannot be positive), there is the potential for double-marginalization: the first mark-up is from the merchants when they have market power and the second mark-up is from the payment network (that has market power).[22]

[21] See Weyl and Fabinger (2013) for a more general treatment of this point. The linear demand we consider in Section 6.4 implies an imperfect pass-through.

[22] A more detailed discussion of the first-best outcome is provided in Section 5.1.

4.4 Endogenizing credit card and cash user shares

We assume that credit and cash are 'differentiated'. We model differentiation using the circular model of Salop (1979).[23] In particular, on a unit circumference circle, the network is located at 0, cash is located at $\frac{1}{2}$ and users/consumers are uniformly distributed on the circumference with density one. We assume that each consumer receives a fixed gross benefit $V > 0$, pays a price P^{cc} when making purchases with a credit card, a price P^{ch} from making purchases with cash and incurs a linear per-unit of distance to a payment method transportation cost $t > 0$.[24] The parameter t captures the degree of differentiation between credit and cash and hence is a measure of network market power: a higher t implies higher network market power. For tractability, we assume that V is independent of the volume of transactions. Thus, the consumer located at $\ell \in \left[0, \frac{1}{2}\right]$, if she uses the credit card obtains a net utility $V - P^{cc} - t\ell$ and if she uses cash obtains a net utility $V - P^{ch} - t \cdot \left(\frac{1}{2} - \ell\right)$. The fraction of consumers who use a credit card is given by

$$\mu = \frac{1}{2} + \frac{P^{ch} - P^{cc}}{t} = \frac{1}{2} + \frac{c \cdot (z_{ch} - z_{cc})}{t \cdot \left(1 - \frac{\gamma}{k}\right)}. \tag{4.6}$$

First, note that the second term of μ in (4.6) is zero when r is zero, since $z_{ch} - z_{cc} = \frac{r}{1-\mu i}$. This means rewards are the key to changing payment shares (the interchange fee i also has an effect when $r \neq 0$). Second, the importance of the transportation cost t is relative to the cost of merchandise. Third, the less competitive is the retail market, i.e., higher γ, the more sensitive is μ to changes in t, all else equal and for $r \neq 0$ and that is because product prices are higher.

4.5 Network's decisions

Using (4.1) and (4.4) the equilibrium aggregate output, for any given interchange fee and reward, is

$$\begin{aligned} X \quad &\equiv \quad \mu x_{cc} + (1-\mu) x_{ch} = \frac{(k-\gamma)^k}{(ck)^k} \left(\frac{\mu \cdot (1-\mu i)^k}{(1-r)^k} + (1-\mu)(1-\mu i)^k \right) \\ &\qquad \left(\text{using } z_{cc} \equiv \frac{1-r}{1-\mu i} \text{ and } z_{ch} \equiv \frac{1}{1-\mu i} \right) \\ &= \quad \frac{(k-\gamma)^k}{(ck)^k} \left(\frac{\mu}{z_{cc}^k} + \frac{1-\mu}{z_{ch}^k} \right), \end{aligned} \tag{4.7}$$

[23] See Jeon and Rey (2022) for a similar assumption regarding differentiation between two platforms.

[24] In Section 6.1, we allow the gross benefits of credit and cash to differ.

where μ is given by (4.6).

The network's profit equals the interchange fee minus the rewards paid to consumers (i.e., the margin per dollar of credit card transactions, which is the network fee, $f = i - r$), multiplied by the total value of transactions made by consumers using a credit card. Formally, this is given by $(i - r)P\mu x_{cc}$, where P denotes the price paid by consumers at the register (see (4.4)). We assume that transaction processing costs are zero.[25] The network profit can therefore be expressed as follows

$$\pi_{\mathcal{N}}(i, r; \mu) = \frac{c\mu \cdot (i - r)}{\left(1 - \frac{\gamma}{k}\right)(1 - \mu i)} x_{cc}(i, r). \tag{4.8}$$

Using the credit card taxes $z_{cc} \equiv \frac{1-r}{1-\mu i}$ and $z_{ch} \equiv \frac{1}{1-\mu i}$, we express (4.8) as an explicit function of these taxes

$$\pi_{\mathcal{N}}(z_{cc}, z_{ch}; \mu) = \frac{c \cdot ((\mu z_{cc} + (1 - \mu) z_{ch}) - 1)}{\left(1 - \frac{\gamma}{k}\right)} x_{cc}(z_{cc}). \tag{4.9}$$

Expressing the network's profit in terms of the implicit taxes z_{cc} and z_{ch}, rather than the interchange fee i and reward r, simplifies the analysis by making the economic trade-offs more transparent. Although μ remains endogenous in both formulations, the tax representation directly captures the wedge between consumer prices and merchant receipts, turning the profit function into an intuitive function of a weighted average tax and credit card consumption. This clarifies how changes in these taxes affect the network's margin, consumer adoption and transaction volume without the added algebraic complexity introduced by the interactions between i, r, and μ.

The network profit function, after we substitute (4.6) into (4.9), is given by

$$\pi_{\mathcal{N}}(z_{cc}, z_{ch}) = \frac{k\left(\frac{k-\gamma}{ck}\right)^k c \cdot \left(\left(\left(\frac{z_{cc}}{2} + \frac{z_{ch}}{2} - 1\right) t - c(z_{cc} - z_{ch})^2\right) k - \frac{t\gamma \cdot (z_{cc} + z_{ch} - 2)}{2}\right) z_{cc}^{-k}}{t \cdot (k - \gamma)^2}. \tag{4.10}$$

The network chooses z_{cc} and z_{ch} to maximize (4.10). For example, the network can increase z_{ch} by increasing the interchange fee i and it can keep z_{cc} fixed by simultaneously increasing the reward r. The weighted average tax, $\mu z_{cc} + (1 - \mu) z_{ch}$, affects network profits. Because $z_{ch} > 1$ (because $i > 0$) the network can even subsidize credit card users by choosing $z_{cc} < 1$ and still make positive profits $(\mu z_{cc} + (1 - \mu) z_{ch} > 1)$. Later in this section we show when this will happen in equilibrium.

[25] Including a constant marginal processing cost would not qualitatively affect our results.

Using (4.9) and (4.6), let's understand the complex trade-offs the network faces when it changes the taxes for cash and credit card users. As a benchmark, observe that when there are no credit card users, $\mu = 0$, then $z_{ch} = 1$ and the network profit is zero. First, we consider the trade-off with respect to the tax for cash users. When z_{ch} increases profits increase, holding μ fixed. However, from (4.6), the share of credit card users, μ, also increases, which reduces profits per transaction assuming $r > 0$ ($z_{ch} > z_{cc}$), which is confirmed in equilibrium. Since credit card users 'cost' the network relatively more than cash users, due to the rewards, any increase in the share of credit at the expense of cash lowers profits per transaction, all else equal. Finally, because $z_{ch} \equiv \frac{1}{1-\mu i}$, a higher μ reinforces the increase in z_{ch}. Second, we consider the trade-off with respect to the tax for credit card users. When the network increases z_{cc}, profits increase while holding μ and individual consumption x_{cc} fixed. Additionally, μ decreases, leading to further profit increases per transaction as discussed above. However, individual consumption using a credit card, x_{cc}, decreases, illustrating the usual tradeoff between margin and volume. Finally, a lower μ mitigates the increase in z_{ch}, since $z_{cc} \equiv \frac{1-r}{1-\mu i}$.

Lemma 1 *The network profit function* (4.10) *is quasi-concave in z_{cc} and z_{ch} if and only if $t < \frac{16ck}{k-\gamma}$, or equivalently $t < 16P^m$.*

The condition in Lemma 1 ensures that the first-order conditions (with equality) are sufficient for a maximum and also guarantees $z^*_{cc} > 0$. If this condition is not met, the solution is at a corner and becomes uninteresting. In the rest of the paper, we assume that this condition is always satisfied.

The network profit-maximizing taxes, which are the solutions to the system of first-order conditions with equality, are given by

$$z^*_{cc} = \frac{(16c - t)k + \gamma t}{16c(k-1)} \text{ and } z^*_{ch} = \frac{(16c + 3t)k^2 - (3\gamma + 4)tk + 4\gamma t}{16(k-1)ck}. \tag{4.11}$$

Cash users always pay a higher tax in equilibrium, $z^*_{ch} - z^*_{cc} = \frac{t\cdot(k-\gamma)}{4ck} = \frac{t}{4}P^m > 0$, which implies that, $\mu = \frac{3}{4}$, i.e., 75% of consumers use a credit card and 25% use cash.[26] It also implies that consumers receive rewards in equilibrium, $r^* > 0$.

Using (4.11), (4.6), and $z_{cc} \equiv \frac{1-r}{1-\mu i}$ and $z_{ch} \equiv \frac{1}{1-\mu i}$ we can derive the unique equilibrium interchange

[26] When we allow for different gross benefits between cash and credit the equilibrium share of credit ranges between 0 and 1, see Section 6.1.

fee and reward

$$i^* = \frac{12k^2t - ((12\gamma + 16)t - 64c)k + 16\gamma t}{(48c + 9t)k^2 - (9\gamma + 12)tk + 12\gamma t} \text{ and } r^* = \frac{4(k-\gamma)(k-1)t}{(16c + 3t)k^2 - (3\gamma + 4)tk + 4\gamma t}. \quad (4.12)$$

Without cash as an alternative payment mode (e.g., Shy and Wang (2011)), i and r cannot be uniquely determined.[27] In this scenario, only the tax for credit card users, $z \equiv \frac{1-r}{1-i}$, is pinned down. Consequently, meaningful comparative statics cannot be performed.

The equilibrium network profits, after substituting (4.11) into (4.10), are given by

$$\pi_{\mathcal{N}}(z^*_{cc}, z^*_{ch}) = \frac{((16c - t)k + \gamma t)\left(\frac{16(k-\gamma)}{ck}\right)^k \left(\frac{(16c-t)k+\gamma t}{c(k-1)}\right)^{-k}}{16(k-1)(k-\gamma)}. \quad (4.13)$$

The equilibrium total merchant profit is

$$\begin{aligned} \Pi^m(z^*_{cc}, z^*_{ch}) &= (P^m - c)X = \frac{c\gamma}{k-\gamma}\frac{(k-\gamma)^k}{(ck)^k}\left(\frac{3}{4(z^*_{cc})^k} + \frac{1}{4(z^*_{ch})^k}\right) \\ &= \frac{c\gamma(ck)^{-k}(k-\gamma)^{k-1}\left(3 \cdot 16^k \left(\frac{(16c-t)k+\gamma t}{k-1}\right)^{-k} + \left(\frac{(16c+3t)k^2-(3\gamma+4)tk+4\gamma t}{16(k-1)ck}\right)^{-k}\right)}{4}. \quad (4.14) \end{aligned}$$

4.6 Comparative statics

We analyze how the key parameters of our model, i.e., demand elasticity (k), merchant market power (γ), and network market power (t), influence the equilibrium outcomes, with particular attention to the implicit taxes borne by credit card and cash users. In each of the three propositions presented in this section, we vary one key parameter while holding the others constant, allowing us to isolate and examine its specific effects.

4.6.1 Effect of demand elasticity

We summarize the effect of demand elasticity in the following Proposition.

[27] Alternatively, the values of i and r can be uniquely identified–even in the absence of an endogenous determination of cash and credit card users–under the assumption that issuing banks incur a convex cost when processing transactions, Wang (2010). However, as discussed in footnote 12, the convex cost assumption is increasingly viewed as outdated due to technological advancements.

Proposition 1 *The equilibrium taxes as a function of the demand elasticity k are given as follows:*

a) Relatively low demand elasticity: $1 < k < \gamma + \frac{16c}{t}$.[28] *Both groups of consumers pay a tax. Credit card users receive a reward* $r^* \geq 0$. *Cash users pay a higher tax than credit card users,* $z^*_{ch} > z^*_{cc} > 1$.

- *Credit card and cash users pay a higher price than the price merchants receive:* $P^{ch} = z^*_{ch}P^m = \frac{(16c+3t)k^2-(3\gamma+4)tk+4\gamma t}{16(k-\gamma)(k-1)} > P^{cc} = z^*_{cc}P^m = \frac{(16c-t)k^2+\gamma kt}{16(k-1)(k-\gamma)} > P^m = \frac{ck}{k-\gamma}$.

b) Relatively high demand elasticity: $k > \gamma + \frac{16c}{t}$.[29] *Cash users pay a tax, while credit card users are subsidized by receiving a relatively high reward,* $z^*_{ch} > 1 > z^*_{cc}$.

- *Credit card users pay a lower price than the price merchants receive:* $P^{cc} = z^*_{cc}P^m = \frac{(16c-t)k^2+\gamma kt}{16(k-1)(k-\gamma)} < P^m = \frac{ck}{k-\gamma}$.
- *If* $k > \overline{k}$, *with* $\overline{k} \geq \gamma + \frac{16c}{t}$, *then the price credit card users pay is even below marginal cost,* $P^{cc} < c$.[30]

c) As demand elasticity increases (higher k):

- *the tax credit card users pay,* z^*_{cc}, *decreases,*
- *the difference between the taxes,* $z^*_{ch} - z^*_{cc}$, *increases,*
- *if* $k < \frac{4}{3}$, *the tax cash users pay,* z^*_{ch}, *decreases,*[31]
- *if* $k > \frac{4}{3}$, *the tax cash users pay,* z^*_{ch}, *decreases if and only if* $t < \frac{16k^2c}{3\gamma k^2-8\gamma k+k^2+4\gamma}$.[32]

Some academics and practitioners have argued that cash users effectively subsidize credit card users when the latter receive rewards (see, for example, Schuh et al. (2010) and Felt et al. (2023)). In this paper, however, we adopt a different interpretation of the term. We define the credit card taxes borne by

[28] From the condition in Lemma 1, $\gamma + \frac{16c}{t} > 1$ always holds.

[29] This condition can be rewritten in terms of t as $t > \frac{16c}{k-\gamma}$. Since $\frac{16c}{k-\gamma} < \frac{16ck}{k-\gamma}$, the case we are examining is valid for relatively high values of t, among the permissible values from the condition in Lemma 1.

[30] Where $\overline{k} \equiv \frac{(16c+t)\gamma+16c+\sqrt{256(\gamma+1)^2c^2+32\gamma(\gamma-1)tc+\gamma^2t^2}}{2t}$. The condition in terms of t can be expressed as $t > \frac{16c((1+\gamma)k-\gamma)}{k(k-\gamma)}$. It can be verified that $\frac{16c}{k-\gamma} < \frac{16c((1+\gamma)k-\gamma)}{k(k-\gamma)} < \frac{16ck}{k-\gamma}$. So, among all permissible values from the condition of Lemma 1, this case is valid for the highest values of t.

[31] When the gross benefits of cash and credit differ, as discussed in Section 6.1, this threshold becomes a function of the respective benefits. The same adjustment applies to the subsequent results and propositions. In this section, however, we assume equal gross benefits across payment modes, as allowing them to differ would not qualitatively affect our results.

[32] The threshold is decreasing as γ increases, if and only if $k > 2$.

both card and cash users and show that, when rewards are positive, cash users face a higher effective tax than card users. Proposition 1 demonstrates that when demand elasticity is relatively high, credit card users are subsidized: they pay a lower price than the price received by merchants (even below marginal cost if the elasticity is really high), which corresponds to the equilibrium price in an all-cash economy-implying that merchants bear none of the tax burden. Conversely, when demand elasticity is low, credit usage imposes a tax on both consumer groups, though the burden is lighter for credit card users due to the rewards. This results in a shared, albeit unequal, distribution of the tax burden between cash and credit card users.

4.6.2 Effect of merchant market power

We summarize the effect of merchant market power in the Proposition below.

Proposition 2 *As merchant market power decreases (lower γ):*

a) the equilibrium interchange fee, $i^* > 0$, decreases if and only if product demand elasticity is relatively low, $k < \frac{4}{3}$,

b) the equilibrium reward, $r^* > 0$, increases,

c) the equilibrium tax of credit card users, z^*_{cc}, decreases,

d) the equilibrium tax of cash users, z^*_{ch}, decreases if and only if product demand elasticity is low, $k < \frac{4}{3}$,

e) equilibrium network profits, $\pi_{\mathcal{N}}(z^*_{cc}, z^*_{ch})$, increase and

f) aggregate merchant equilibrium profits, $\Pi^m(z^*_{cc}, z^*_{ch})$, are inverse U-shaped.

The effect of merchant competition arises from the influence of γ on the fraction of credit card users, μ. It can be understood as follows. From (4.6), as γ decreases, μ also decreases for a fixed $z_{ch} - z_{cc} > 0$. Intuitively, in a more competitive merchant market, prices are lower. Since taxes are ad valorem, a tax differential in favor of credit provides a smaller advantage to credit relative to cash. Therefore, the network has an incentive to increase $z_{ch} - z_{cc} > 0$ to restore its market share. This is achieved by decreasing z_{cc} and increasing z_{ch} or decreasing z_{ch} but at a lower rate.

Aggregate merchant profits can increase as competition in the product market intensifies. This occurs because the network, which competes with cash, lowers the tax on credit card users, who, in the base case of equal gross benefits between cash and credit, constitute the majority of consumers, thereby encouraging increased consumption. In other words, there is a range of values for the market conduct parameter in which the double marginalization problem is first order in economic significance. This result contrasts with the findings of Shy and Wang (2011), who show that aggregate merchant profits monotonically decrease as the market becomes more competitive. Higher consumption also benefits the network profits.

4.6.3 Effect of network market power

We summarize the effect of differentiation between cash and credit, which is used as a proxy for network market power, in the Proposition below (most parts of the proof are straightforward and are omitted).

Proposition 3 *As network market power decreases (lower t):*

a) the equilibrium tax cash users pay, z^*_{ch}, decreases if and only if product demand elasticity is above a threshold, $k > \frac{4}{3}$,

b) the equilibrium interchange fee i^* decreases if and only if product demand elasticity is above a threshold, $k > \frac{4}{3}$,

c) the equilibrium tax credit card users pay, z^*_{cc}, increases,

d) the equilibrium reward, r^*, decreases and

e) the difference between the taxes, $z^*_{ch} - z^*_{cc}$, decreases,

f) the equilibrium network profit, $\pi_{\mathcal{N}}(z^*_{cc}, z^*_{ch})$, decreases,

g) aggregate merchant equilibrium profits, $\Pi^m(z^*_{cc}, z^*_{ch})$, decrease.[33]

The finding that the tax for credit card users increases as network market power decreases might seem counterintuitive. One might expect that, as the network faces stronger competition from cash, it

[33] An analytical solution for this result could not be obtained; instead, it is supported by numerical simulations. All parameter configurations explored consistently corroborate the stated outcome.

would increase the rewards offered to credit card users, thereby lowering the tax for these consumers. The intuition is as follows. From (4.6), and given that $z_{ch} > z_{cc}$, a lower t increases the share of credit card users, holding the taxes fixed. As we have already discussed, this leads to lower network profits per transaction (because credit card users cost the network relatively more than cash users). The network lowers the rewards and the difference between the two taxes to mitigate this negative effect. Aggregate consumption falls impacting negatively network and merchant profits.

5 Social efficiency and regulation

5.1 Social efficiency

Relative to the social optimum, our model identifies three primary sources of inefficiency. First, network market power leads to inefficiency, as the monopoly network sets an interchange fee above the marginal cost of processing transactions, which is assumed to be zero. Second, merchant market power introduces further inefficiency. Together, these two sources create a double marginalization problem, both contributing to distorted consumption patterns. Third, consumer behavior may be inefficiently influenced by rewards, which can incentivize the use of a payment method that deviates from the consumer's preferred choice in the absence of cross-subsidization. Specifically, in our model, cash and cards are imperfect substitutes, with consumer preferences represented along a Salop circle. Rewards can induce consumers to incur transportation costs that exceed the socially optimal level.

To address these distortions, a social planner would set $i = r = 0$, or equivalently $z_{cc} = z_{ch} = 1$, resulting in an equal (50%) distribution of consumers between credit and cash, see (4.6).[34] Additionally, the planner would set the price P equal to the marginal cost c.

In what follows, we do not explicitly compare the relative importance of the effects of i and r on consumption to their effects on transportation costs. But we do show how regulation affects them both and show when an interchange fee cap could improve welfare in both dimensions. We emphasize that this is not always the case.

[34] In section 6.1, we allow for different valuations of cash and credit and, in that case, the socially optimal mix of payments could be different from 50%. In addition, if we allow for a differential in the physical cost of accepting cash versus cards, this too would affect the planner's decision. See Section 6.3 for an illustration

5.2 Regulation

We examine a regulatory intervention that imposes a cap, reducing the interchange fee from its private equilibrium level. Our primary focus is on the implications of this policy for consumer welfare and overall social efficiency. Since the regulation caps i directly, we need to express the network's profit function, given in equation (4.10), as a function of i. Using the expression for μ, given by (4.6), and $z_{ch} \equiv \frac{1}{1-\mu i}$ and $z_{cc} \equiv \frac{1-r}{1-\mu i}$, we solve for i and r

$$i = \frac{2t \cdot (z_{ch} - 1)(\gamma - k)}{z_{ch}(t \cdot (\gamma - k) - 2ck \cdot (z_{ch} - z_{cc}))}, \quad r = 1 - \frac{z_{cc}}{z_{ch}}.$$

Next, we invert this system to express z_{ch} and z_{cc} as functions of i and r

$$\begin{aligned} z_{ch} &= \frac{(k-\gamma)(2-i)t - \sqrt{16t \cdot \left(\frac{(2-i)^2(k-\gamma)t}{16} - cikr\right)(k-\gamma)}}{4cikr} \\ z_{cc} &= \frac{(1-r)\left((k-\gamma)(2-i)t - \sqrt{16t \cdot \left(\frac{(2-i)^2(k-\gamma)t}{16} - cikr\right)(k-\gamma)}\right)}{4cikr}. \end{aligned} \tag{5.1}$$

We substitute (5.1) into the network profit function (4.10), which now becomes $\pi_{\mathcal{N}}(i, r)$ (similar to (4.8), but with endogenous μ). From the first-order condition of the network profit with respect to r, we derive the effect of i on the equilibrium reward, evaluated at the equilibrium values given by (4.12): $\frac{dr^*}{di} = -\frac{\frac{\partial^2 \pi_{\mathcal{N}}(i^*, r^*)}{\partial i \partial r}}{\frac{\partial^2 \pi_{\mathcal{N}}(i^*, r^*)}{\partial r^2}}$. Given the quasi-concavity of the profit function, we have $\frac{\partial^2 \pi_{\mathcal{N}}(i^*, r^*)}{\partial r^2} < 0$. Therefore, if $\frac{\partial^2 \pi_{\mathcal{N}}(i^*, r^*)}{\partial i \partial r} > 0$, implying that i and r are complementary in equilibrium, a marginal cap on the interchange fee would lead the network to reduce the reward. Conversely, if $\frac{\partial^2 \pi_{\mathcal{N}}(i^*, r^*)}{\partial i \partial r} < 0$, implying that i and r are substitutable in equilibrium, the network would increase the equilibrium reward. This subtlety turns out to be important when thinking about policy and welfare.

The effect of the regulation on the equilibrium reward constitutes an intermediate step, but it is not synonymous with changes in consumer welfare. To assess welfare implications, it is necessary to examine the impact on equilibrium taxes. We differentiate equations (5.1) with respect to i, yielding the partial derivatives $\frac{\partial z_{ch}^*}{\partial i}$ and $\frac{\partial z_{cc}^*}{\partial i}$. These expressions incorporate the direct effect of i, as well as the indirect effect on the equilibrium reward, captured by $\frac{dr^*}{di}$, and are evaluated at the equilibrium defined by equation (4.12).

Given the complexity of the underlying expressions, closed-form analytical solutions are intractable. As a result, we rely on numerical simulations to characterize the impact of the regulation. To ensure the robustness of our findings, we conduct extensive sensitivity analyses across a wide range of parameter constellations, extending beyond those presented in the main text. In all cases, the qualitative nature of our results remains unchanged. Throughout this section, we refer to our main findings as "results" rather than "propositions," as their validation is based on numerical simulations rather than formal analytical proofs. Accordingly, the statements hold for the parameter values we have examined and while we have undertaken extensive robustness checks, they should be interpreted with this caveat in mind.

We set $t = c = 1$ for the remainder of this section and focus on the roles of merchant market power and demand elasticity. The simulations indicate that $\frac{dr^*}{di} < 0$ if and only if $k < 5.093822$ when $\gamma = 1$; $k < 4.8424$ when $\gamma = 0.5$; and $k < 4.6164$ when $\gamma = 0$.[35] By continuity, the sign of the derivative remains valid in a neighborhood of the equilibrium interval $(\underline{i}, i^*]$ (see Ronnen (1991) for a similar approach).[36] We summarize the findings below.

Result 1 *There exists an interval $(\underline{i}, i^*]$, such that if the interchange fee is set within $(\underline{i}, i^*]$, where i^* is the unregulated equilibrium interchange fee, the equilibrium rewards will increase if and only if demand elasticity is lower than a threshold.*

In policy debates over credit card interchange fees, the claim that reducing these fees necessarily reduces cardholder rewards is often treated as axiomatic. In our model, however, we find that this holds only when demand is relatively elastic. When demand is less elastic, a reduction in interchange fees can actually lead to an increase in rewards. This result is not limited to the case of constant elasticity demand; it also holds under a linear demand specification (see Section 6.4).

The core intuition is as follows. A forced reduction in the interchange fee, i, lowers the network's margin, $i - r$. The network can respond by either lowering the reward, r, to partially restore the margin or increasing r to stimulate consumption. When demand elasticity is relatively low, a reduction in i does not lead to a significant increase in consumption. In this case, the network finds it profitable to offset part of the lost revenue by encouraging additional spending, from existing and 'new' cardholders, through

[35] We developed a MATLAB script (available upon request) that iterates over numerous configurations of (t, c, γ). For each configuration, we identified a unique threshold, greater than one, with respect to k.

[36] This continuity property applies to all results presented in this section, although the lower bound $\underline{i}$ may vary across different results.

an increase in r.[37] By contrast, when demand elasticity is high, consumption responds strongly to the regulation. This allows the network to recoup some of its lost revenue by lowering r to partially restore its margin, while still benefiting from the higher transaction volume.

The impact of an interchange fee on consumption depends critically on the equilibrium credit card taxes borne by consumers (recall that the price merchants receive is not affected by the tax, so any tax change is passed-on to consumers). These taxes, $z_{ch} \equiv \frac{1}{1-\mu i}$ and $z_{cc} \equiv \frac{1-r}{1-\mu i}$, are determined by three variables: i and μ for all consumers and additionally r for credit card users. The regulation directly reduces i; the central question is how μ and r adjust in response. A decrease in μ–that is, a reduction in the proportion of consumers using credit cards–unambiguously lowers the tax burden on *all* consumers. Moreover, μ and r are positively related, all else equal: a lower reward decreases the fraction of credit card users. Consequently, when equilibrium rewards decrease, cash users are unambiguously better off (because both i and μ decrease and hence $1-\mu i$ increases), whereas credit card users face two opposing forces: higher $1-r$, but also higher $1-\mu i$. Conversely, when equilibrium rewards increase, both types of users experience conflicting effects, since i and μ move in opposite directions; however, credit card users are more likely to benefit. In the following section, we examine the net effect of the regulation in greater detail.

First, we examine the sign of the partial derivatives $\frac{\partial z^*_{cc}}{\partial i}$ and $\frac{\partial z^*_{ch}}{\partial i}$ for $\gamma \in \{0, 0.5, 1\}$. For all these parameter values, the sign is positive for all $k > 1$, indicating that an interchange fee cap reduces taxes across all levels of demand elasticity and for all consumer types. Thus, even when the equilibrium reward decreases, the equilibrium taxes for credit card users decline.[38]

Result 2 *There exists an interval $(\underline{i}, i^*]$, such that if the interchange fee is set within $(\underline{i}, i^*]$, where i^* is the unregulated equilibrium interchange fee, the equilibrium taxes for credit and cash users decrease.*

Thus, even for the case (high elasticity) when a reduction in interchange fees reduces rewards, credit card users are not made worse off because, on net, the credit card tax falls, reducing the prices they ultimately pay.[39]

[37] This relationship holds under a modest reduction in the interchange fee. In contrast, a substantial decrease would erode the margin and necessitate a downward adjustment of the reward by the network. Despite this, the qualitative result–that a lower interchange fee may lead to higher rewards–continues to hold even for sizeable reductions, such as 10% (see Section 6.4).

[38] As with Result 1, we developed a MATLAB script (available upon request) that iterates over numerous configurations of (t, c, γ, k). For each configuration, $\frac{\partial z^*_{ch}}{\partial i}$ and $\frac{\partial z^*_{cc}}{\partial i}$, evaluated at $i = i^*$, are strictly positive.

[39] In our main model, the magnitude of consumer benefit from an interchange fee cap depends significantly, but not entirely,

Although the economic forces we have identified are general, their magnitude depends on the demand specification. In the case of linear demand, as opposed to constant elasticity, post regulation outcomes for cash users can deteriorate when demand is relatively inelastic. The mechanism is a considerable increase in the equilibrium reward, which triggers an expansion in credit card adoption. This expansion is large enough to offset the initial reduction in the interchange fee i, thereby depressing the value of $1 - \mu i$ (see Section 6.4 for more details and intuition).

The regulation induces a reallocation of surplus from the payment network to consumers and merchants. Because the net price received by merchants is independent of the credit card tax, the tax reduction does not affect their per-unit revenue. However, by lowering the final price faced by consumers, the regulation stimulates the transaction volume. This increase in aggregate consumption raises total merchant profits whenever $\gamma > 0$. Thus, surplus is transferred from the network to both sides of the market. In the limiting case where $\gamma = 0$, merchant profits are zero and the transfer accrues solely to consumers.

Next, we turn to the effect of an interchange fee cap on optimal payment choice. To do that, we compare the magnitudes of $\frac{\partial z^*_{cc}}{\partial i}$ and $\frac{\partial z^*_{ch}}{\partial i}$. We find that $\frac{\partial z^*_{cc}}{\partial i} > \frac{\partial z^*_{ch}}{\partial i}$ if and only if $k < 4.5543$ when $\gamma = 1$, $k < 4.2637$ when $\gamma = 0.5$, and $k < 4$ when $\gamma = 0$. This comparison is important because it determines whether the gap $z^*_{ch} - z^*_{cc}$ narrows or widens following the regulation. This gap, in turn, affects μ, which–as discussed in Section 5.1–is socially excessive. Under the assumptions of the model, the gap narrows if and only if demand elasticity exceeds a threshold. This outcome arises primarily because, as noted above, the reward decreases, making the card less attractive relative to cash. Therefore, for high values of demand elasticity, the regulation helps mitigate the inefficiency associated with excessive card usage. In contrast, when elasticity is low, the gap widens, thereby exacerbating the social inefficiency.

Result 3 *There exists an interval $(\underline{i}, i^*]$, such that if the interchange fee is set within $(\underline{i}, i^*]$, where i^* is the unregulated equilibrium interchange fee, the inefficiency associated with excessive credit card usage is mitigated if and only if demand elasticity exceeds a threshold.*

When the equilibrium taxes fall, the effect on consumption efficiency is, in general, ambiguous. For cash users, since $z^*_{ch} > 1$, consumption moves closer to the first-best outcome (which, as discussed in

on the complete pass–through to prices that is implied by the constant elasticity demand specification. When demand elasticity varies with output, pass–through tends to be imperfect, allowing merchants to retain a portion of the interchange fee reduction. In Section 6.4, we consider a linear demand which features a non constant elasticity and hence implies an imperfect pass-through.

Section 5.1, is $z_{cc} = z_{ch} = 1$ and $P = c$). For credit card users, the effect depends on the level of z^*_{cc}, which can be either greater or less than 1 (see Proposition 1). If $z^*_{cc} > 1$–a situation that arises when demand elasticity is relatively low ($k < \gamma + \frac{16c}{t}$)–consumption moves toward the efficient outcome. Conversely, if $z^*_{cc} < 1$, which occurs when demand elasticity is relatively high ($k > \gamma + \frac{16c}{t}$), a lower tax for credit card users may lead to a deviation from the first-best. As discussed in Proposition 1, for values of k even higher than $\gamma + \frac{16c}{t}$, that is $k > \overline{k}$, credit card users are subsidized to the extent that they pay a price below the marginal cost of production c, so their consumption is inefficiently high. A lower tax in this case exacerbates this inefficiency.

Result 4 *There exists an interval $(\underline{i}, i^*]$, such that if the interchange fee is set within $(\underline{i}, i^*]$, where i^* is the unregulated equilibrium interchange fee, the consumption inefficiency is mitigated if demand elasticity is relatively low, that is $k < \overline{k}$.*

According to Result 3, the social inefficiency arising from excessive credit card usage is alleviated under an interchange fee cap regulation when the demand elasticity is relatively high. Conversely, Result 4 demonstrates that the inefficiency associated with suboptimal consumption levels is mitigated when the demand elasticity is relatively low. This raises the question of whether a single regulatory intervention can simultaneously address both types of inefficiencies. The answer is affirmative: there exists a range of values for k within which an interchange fee cap regulation mitigates both forms of social inefficiency. Specifically, when $\gamma = 0$, the relevant interval is $k \in (4, 16)$; when $\gamma = 0.5$, the interval expands to $k \in (4.2637, 24.17)$; and when $\gamma = 1$, it further extends to $k \in (4.5543, 32.5)$. Hence, for demand elasticity within these ranges, the regulation not only benefits all consumers by reducing credit card taxes, but also enhances overall social welfare by promoting consumption closer to the first-best level and curbing excessive card usage.

6 Extensions

6.1 Cash and credit have different benefits

We extend our baseline model to account for differential benefits (or costs) associated with cash and credit transactions. Specifically, we assume that a consumer derives a gross benefit of V_{ch} when purchas-

ing with cash, and a gross benefit of V_{cc} when purchasing with credit. Define the difference in benefits as $\Delta \equiv V_{cc} - V_{ch}$. When credit offers greater non-pecuniary benefits than cash, $\Delta > 0$; conversely, when cash is more beneficial, $\Delta < 0$. The mass of consumers who choose to use a credit card, as originally defined in equation (4.6), is now re-expressed as follows:

$$\mu = \frac{1}{2} + \frac{\Delta + P^{ch} - P^{cc}}{t} = \frac{1}{2} + \frac{\Delta}{t} + \frac{c \cdot (z_{ch} - z_{cc})}{t \cdot \left(1 - \frac{\gamma}{k}\right)}.$$

There are no changes in the other parts of the model. The profit function of the network is quasi-concave if and only if $\Delta > \frac{t}{2} - \frac{2\sqrt{ctk}}{\sqrt{k-\gamma}}$. The equilibrium taxes are as follows

$$\begin{aligned} z_{cc}^* &= \frac{(-k+\gamma)t^2 + 4((4c+\Delta)k - \Delta\gamma)t - 4\Delta^2(k-\gamma)}{16(k-1)ct} \\ z_{ch}^* &= \frac{(3t^2 + (16c - 4\Delta)t - 4\Delta^2)k^2 - 3(t-2\Delta)\left(\left(\gamma + \frac{4}{3}\right)t + \frac{2\Delta\gamma}{3}\right)k + 4t\gamma \cdot (t - 2\Delta)}{16(k-1)tck}. \end{aligned}$$

The equilibrium share of credit is $\mu = \frac{\Delta}{2t} + \frac{3}{4}$. Note that $\mu \in (0,1)$ if and only if $\Delta \in \left(-\frac{3t}{2}, \frac{t}{2}\right)$. The difference in equilibrium taxes is $z_{ch}^* - z_{cc}^* = \frac{(k-\gamma)(t-2\Delta)}{4ck}$ and is decreasing in Δ.

The equilibrium interchange fee and rewards are are follows

$$\begin{aligned} i^* &= \frac{4t[(\gamma-k)(2\Delta-t)(2k\Delta+3kt-4t)+16ckt]}{(2\Delta+3t)[(\gamma-k)(2\Delta-t)(2k\Delta+3kt-4t)+16ck^2t]} \\ r^* &= \frac{4t(k-1)(k-\gamma)(t-2\Delta)}{(\gamma-k)(2\Delta-t)(2k\Delta+3kt-4t)+16ck^2t}. \end{aligned}$$

The effect of Δ on the equilibrium taxes is given by

$$\frac{dz_{cc}^*}{d\Delta} = \frac{(k-\gamma)(t-2\Delta)}{4(k-1)ct} \text{ and } \frac{dz_{ch}^*}{d\Delta} = -\frac{((t+2\Delta)k - 2t)(k-\gamma)}{4(k-1)tck}.$$

Credit card users pay a higher tax when the credit benefits increase relative to cash benefits, $\frac{dz_{cc}^*}{d\Delta} > 0$. In addition, the equilibrium reward, r^*, is decreasing in Δ and at $\Delta = \frac{t}{2}$, $r^* = 0$. When all consumers use credit, the equilibrium reward is zero. For cash users, $\frac{dz_{ch}^*}{d\Delta} > 0$ if and only if $k < 2 + \frac{t}{\Delta}$. When demand elasticity is relatively low, the network raises the interchange fee as the benefits of credit increase (or the cost of cash increases). The reverse is true when demand elasticity is relatively high.

The equilibrium profit of the network is given by

$$\pi_{\mathcal{N}} = \frac{16^k \left(\left(-\frac{t^2}{16} + \left(c + \frac{\Delta}{4}\right) t - \frac{\Delta^2}{4}\right) k + \frac{\gamma(t-2\Delta)^2}{16}\right) \left(\frac{(-4\Delta^2+4\Delta t+16ct-t^2)k+\gamma(t-2\Delta)^2)}{(k-1)ct}\right)^{-k} \left(\frac{k-\gamma}{ck}\right)^k}{t(k-\gamma)(k-1)}.$$

It can verified that $\pi_{\mathcal{N}}$ is decreasing in Δ, for all $\Delta \in \left(\frac{t}{2} - \frac{2\sqrt{ctk}}{\sqrt{k-\gamma}}, \frac{t}{2}\right)$.[40] When the benefit of credit, relative to cash, increases, the network is worse-off.

We summarize in the following Proposition.

Proposition 4 *As the benefits to consumers of credit, relative to cash, increase (or the cost of cash relative to credit increases), i.e., Δ increases, the following hold:*

a) The equilibrium share of credit card users, $\mu = \frac{\Delta}{2t} + \frac{3}{4}$, increases,

b) the equilibrium tax for credit card users, z^*_{cc}, increases,

c) the equilibrium tax for cash users, z^*_{ch}, increases if and only if demand elasticity is relatively low, i.e., $k < 2 + \frac{t}{\Delta}$,

d) the tax differential $z^*_{ch} - z^*_{cc}$ is decreasing and becomes zero when all consumers use credit, that is when $\Delta = \frac{t}{2}$,

e) the equilibrium reward, r^*, is monotonically decreasing and becomes zero when all consumers use credit, that is when $\Delta = \frac{t}{2}$,

f) when demand elasticity is higher, $k > 2 + \frac{t}{\Delta}$, the equilibrium interchange fee, i^*, is monotonically decreasing; otherwise, it can be inverse U-shaped and

g) the equilibrium profit of the network is decreasing.

Equilibrium taxes on credit card users increase for two primary reasons. First, as the benefits associated with credit usage rise, the payment network faces reduced competitive pressure from cash transactions, allowing it to lower the rewards offered to users. Second, a higher proportion of credit card users leads to an increase in the effective tax borne by all consumers. While the network may attempt

[40] Specifically, $\pi_{\mathcal{N}}$ is convex in Δ, is minimized at $\Delta = \frac{t}{2}$ and has an asymptote at $\Delta = \frac{t}{2} - \frac{2\sqrt{ctk}}{\sqrt{k-\gamma}}$.

to mitigate this effect by reducing the interchange fee, such adjustments are insufficient to fully counterbalance the underlying forces. A key implication of higher taxes on credit card users–who constitute the majority of consumers–is a decline in aggregate consumption, which in turn reduces the network's overall profitability.

6.2 Would cash users be better off if rewards were prohibited?

Many scholars and industry experts argue that interchange fees are elevated, in part because they help finance rewards for credit card users. Consequently, consumers who pay with cash may be disadvantaged. To explore this issue, we force $r = 0$, which implies a single tax $z \equiv \frac{1}{1-\mu i}$ that applies uniformly to all consumers. From equation (4.6), it follows that $\mu = \frac{1}{2}$. Substituting into equation (4.9), the network's profit function becomes

$$\pi_{\mathcal{N}} = \frac{(k-\gamma)^k (z-1)c}{(ck)^k z^k \left(1 - \frac{\gamma}{k}\right)},$$

which is maximized at

$$z^* = \frac{k}{k-1}.$$

Next, using the equilibrium tax for cash users when rewards are permitted, as given by equation (4.11), we find:

$$z^*_{ch} - z^* = \frac{(k-\gamma)t \cdot (3k-4)}{16(k-1)ck}, \tag{6.1}$$

which is positive if and only if $k > \frac{4}{3}$.[41] We summarize in the Proposition below.

Proposition 5 *Suppose rewards for credit card transactions are prohibited, i.e., $r = 0$. In that case, the equilibrium tax on cash users decreases, relative to when rewards are allowed, if and only if demand elasticity is relatively high, that is, $k > \frac{4}{3}$.*

Under conditions of relatively high demand elasticity, our model predicts that prohibiting credit card rewards results in a lower effective tax on cash users, thereby improving their welfare. Conversely,

[41] As in Proposition 2, the threshold is not fixed at $\frac{4}{3}$; rather, it becomes a function of the gross benefits when the two payment modes differ in this respect, as modeled in Section 6.1. In this section, we deliberately abstract from such differences in gross benefits, as allowing for them would not qualitatively alter our results.

when demand elasticity is low, cash users are better off when rewards are permitted. The underlying mechanism is as follows: banning rewards reduces the prevalence of credit card usage (from $\mu = \frac{3}{4}$ to $\mu = \frac{1}{2}$), which, all else equal, lowers the cross-subsidization burden borne by cash users. However, in response to the decline in transaction volume, the card network may raise the interchange fee to recoup lost revenue. If the increase in the interchange fee is limited, the net effect is a reduction in the equilibrium tax on cash users. When demand elasticity is high, the network faces stronger disincentives to raise fees substantially, as doing so would significantly suppress aggregate consumption.

6.3 Merchants incur a cash cost – An Illustration

In our baseline model, it becomes apparent that the introduction of credit reduces welfare by exposing consumers to taxation. This conclusion, however, rests on two key simplifying assumptions in our model: (i) the use of cash incurs no cost and (ii) the gross benefit from both payment methods, denoted by V, is identical. In this extension, we relax the first assumption.

Suppose now that merchants incur a cost τ associated with accepting, safeguarding and transporting cash, including potential losses due to employee theft. Accordingly, the profit function of merchant j, originally given in (4.2), is modified as follows

$$\begin{aligned}\pi_j &= \left(\mu(1-i)P(X,r,\mu) + (1-\mu)(1-\tau)P(X,r,\mu)\right)x_j - cx_j \\ &= \left(1-\mu i-(1-\mu)\tau\right)P(X,r,\mu)x_j - cx_j.\end{aligned}$$

Following the same steps as in the analysis in Section 4.3, the tax expressions become

$$z_{ch} \equiv \frac{1}{1-\mu(1-i)-(1-\mu)\tau}, \quad z_{cc} \equiv \frac{1-r}{1-\mu(1-i)-(1-\mu)\tau}. \tag{6.2}$$

In the absence of credit (i.e., $\mu = 0$), cash users face a tax of $z \equiv \frac{1}{1-\tau}$. When $\tau > 0$, the network's profit function becomes significantly more complex than in (4.10), precluding closed-form solutions. Consequently, we rely on numerical simulations. Setting $\tau = 0.03$ implies a 3% cost of using cash, resulting in a tax of $z^* = 1.0309$ in the absence of credit. We further assume $t = 0.2$, $c = 2$, $\gamma = 0.5$, and $k = 35$. In equilibrium, we find $\mu = 0.9107$, indicating that approximately 91% of consumers adopt credit. The corresponding equilibrium tax rates are $z^*_{cc} = 1.028702$ and $z^*_{ch} = 1.0692$. Thus, following

the introduction of credit, 91% of consumers experience a lower tax burden. Naturally, for alternative parameter values–particularly when k is low–the introduction of credit may render all consumers worse off.

The purpose of this extension is to demonstrate that, within our model, the introduction of credit can make the vast majority of consumers better off when cash handling entails a reasonable cost, even if the gross benefit of the two payment modes is identical. We do not assess the plausibility of the parameter values under which this outcome arises.

6.4 Linear demand

While the assumption of constant elasticity simplifies the analysis and facilitates the derivation of several novel results, it is not without its limitations. A well known drawback is that it leads to a one-for-one increase in the price at the register in response to a rise in credit card taxes. This implies perfect pass-through and that the entire tax incidence falls on consumers, a result that may limit the policy relevance of our findings.

To address this concern, albeit at the expense of some clarity and tractability, we adopt an alternative framework with linear product demand. This specification offers a balance of analytical tractability and a demand elasticity that varies with aggregate output. This variation introduces an *elasticity effect*, the direction of which is determined by the sign of the derivative of the demand elasticity with respect to output. Under linear demand, this sign is positive; that is, market demand becomes less elastic as aggregate output increases. In this environment, the incidence of a credit card cost is shared between merchants and consumers, except in the case of a perfectly competitive merchant market.

This result aligns with the taxonomy of demand functions proposed by Mrázová and Neary (2019), who define a demand function as *subconvex* if $\log p$ is concave in $\log x$. This is equivalent to a demand function that is less convex than a constant elasticity function and one whose (absolute) elasticity of demand decreases with output, precisely the case for linear demand. Subconvexity is sometimes called "Marshall's Second Law of Demand". In contrast, *superconvex* demand, where elasticity increases with output, can lead to tax over-shifting, whereby consumers bear more than 100% of a tax. Such an outcome is often considered less plausible, further justifying our focus on the linear, subconvex case for a more

realistic analysis of pass-through.[42]

Our main focus in this extension is on the welfare implications, particularly the counterparts of Results 1 and 2 in Section 5.2, which examine the effects of an interchange fee cap regulation.

Suppose individual demand is linear $x = 1 - P$. The inverse aggregate demand, accounting for rewards r and the share of credit card users μ, is given by

$$P(X, \mu, r) = \frac{1 - X}{\mu \cdot (1 - r) + (1 - \mu)}. \tag{6.3}$$

Combining (6.3) with (4.4) we derive the equilibrium aggregate consumption as a function of the two taxes

$$X(z_{ch}, z_{cc}; \mu) \equiv \mu x_{cc} + (1 - \mu) x_{ch} = \frac{1 - c z_{ch} \cdot (1 - \mu) - c z_{cc} \mu}{1 + \gamma}. \tag{6.4}$$

Using (4.6) and the demand elasticity, $\varepsilon = -\frac{1-X}{X}$, we derive the share of credit card users[43]

$$\mu(z_{ch}, z_{cc}; X) = \frac{1}{2} + \frac{P^{ch} - P^{cc}}{t} = \frac{1}{2} + \frac{c \cdot (z_{ch} - z_{cc})}{1 - \frac{\gamma X}{1 - X}}. \tag{6.5}$$

We solve (6.5) and (6.4) simultaneously with respect to X and μ, to obtain $X(z_{ch}, z_{cc})$ and $\mu(z_{ch}, z_{cc})$, either analytically (when $\gamma = 0$) or numerically (when $\gamma > 0$). The profit of the network, from (4.9), is

$$\pi_{\mathcal{N}}(z_{ch}, z_{cc}; X, \mu) = \frac{c \cdot ((\mu z_{cc} + (1 - \mu) z_{ch}) - 1)}{\left(1 - \frac{\gamma X}{1 - X}\right)} x_{cc}(z_{cc}), \tag{6.6}$$

where $x_{cc} = 1 - c z_{cc}$. We substitute $X(z_{ch}, z_{cc})$ and $\mu(z_{ch}, z_{cc})$ into (6.6). Note that c must be less than 1. Since $X < 1$, the following inequality holds, $z_{ch}(1 - \mu) + z_{cc}\mu < \frac{1}{c}$, which is satisfied when network profits are positive.

[42] Mrázová and Neary (2017, 2019) provide a comprehensive discussion of how superconvexity and subconvexity determine competition effects and relative pass-through.

[43] For simplicity, and without significant loss of generality, we set the transportation cost parameter to $t = 1$.

6.4.1 Equilibrium in the absence of a credit card

We digress briefly in this subsection to derive the equilibrium for the benchmark case without credit. By setting $\mu = 0$ and $z_{ch} = 1$ in (6.4), we obtain the following expressions

$$\widetilde{X} = \frac{1-c}{1+\gamma}, \qquad \widetilde{P} = \frac{\gamma+c}{1+\gamma}, \qquad \widetilde{\Pi} = \frac{\gamma(1-c)^2}{(1+\gamma)^2} \tag{6.7}$$

where $\widetilde{X}$ is the aggregate equilibrium output, $\widetilde{P}$ is the equilibrium price all consumers pay and all merchants receive and $\widetilde{\Pi}$ is the equilibrium merchant aggregate profit, in the absence of the payment network. This benchmark equilibrium will serve as a point of comparison when we later compute the tax incidence due to the credit card.

6.4.2 Perfectly competitive merchant market, $\gamma = 0$

When the merchant market is perfectly competitive ($\gamma = 0$), the model yields tractable closed-form solutions, which we present as a useful benchmark (but still results about the effects of the regulation are numerical). In this limiting case, $\frac{\gamma}{\varepsilon} = \frac{\gamma X}{1-X} = 0$ and the price received by merchants remains unaffected by the tax. Consequently, the final price paid by consumers reflects a one-to-one pass-through of the tax. Even in this limiting setting, however, our framework yields a new result, relative to constant elasticity demand, regarding the effect of a regulation on the equilibrium taxes, that we discuss below. In the next subsection, we generalize the analysis by allowing for imperfect competition ($\gamma > 0$) and hence imperfect pass-through.

The simultaneous solution of (6.5) and (6.4) yields

$$\begin{aligned} X(z_{ch}, z_{cc}) &= 1 + c^2(z_{ch} - z_{cc})^2 - \frac{c \cdot (z_{ch} + z_{cc})}{2} \\ \mu(z_{ch}, z_{cc}) &= \frac{1}{2} + c \cdot (z_{ch} - z_{cc}). \end{aligned} \tag{6.8}$$

After substituting (6.8) into (6.6) we derive the network profit as a function of the two taxes

$$\pi_{\mathcal{N}}(z_{cc}, z_{ch}) = \frac{c \cdot (2(z_{ch} - z_{cc})^2 c - z_{ch} - z_{cc} + 2)(cz_{cc} - 1)}{2}. \tag{6.9}$$

We differentiate $\pi_{\mathcal{N}}(z_{cc}, z_{ch})$ with respect to z_{cc} and z_{ch}. The following three sets of roots satisfy the first order conditions (with equality)

$$\begin{aligned} z_{cc} &= \frac{1}{c} \text{ and } z_{ch} = \frac{5 \pm \sqrt{17 - 16c}}{4c} \\ z_{cc} &= \frac{16c + 15}{32c} \text{ and } z_{ch} = \frac{16c + 23}{32c}. \end{aligned}$$

Only the last set of roots also satisfies the second order conditions. Hence, the taxes that maximize the network profit, subject to $X(z_{ch}, z_{cc}) \in (0, 1)$ and $\mu(z_{ch}, z_{cc}) \in [0, 1]$, are given by

$$z^*_{cc} = \frac{16c + 15}{32c} \text{ and } z^*_{ch} = \frac{16c + 23}{32c}. \tag{6.10}$$

The equilibrium aggregate consumption is $X^* = \frac{15}{32} - \frac{c}{2}$, so the interior solution is valid for $X^* > 0 \Rightarrow c < \frac{15}{16}$.

Using (6.10), together with $z_{ch} \equiv \frac{1}{1-\mu i}$ and $z_{cc} \equiv \frac{1-r}{1-\mu i}$, we derive the equilibrium interchange fee and rewards

$$i^* = \frac{4(23 - 16c)}{3(23 + 16c)} \text{ and } r^* = \frac{8}{23 + 16c}. \tag{6.11}$$

An interchange fee cap regulation. We will now express the network profit function in terms of i and r. We use $\mu(z_{ch}, z_{cc})$, from (6.8), together with $z_{ch} \equiv \frac{1}{1-\mu i}$ and $z_{cc} \equiv \frac{1-r}{1-\mu i}$ and solve with respect to z_{ch} and z_{cc}. We obtain the following two sets of solutions

$$\begin{aligned} z^1_{ch}(i, r) &= \frac{-i + 2 + \sqrt{-16cir + i^2 + 4(1 - i)}}{4cir} \\ z^1_{cc}(i, r) &= -\frac{4\left(\left(\frac{i}{8} - \frac{1}{4}\right)\sqrt{-16cir + i^2 + 4(1 - i)} + cir - \frac{i^2}{8} + \frac{i}{2} - \frac{1}{2}\right)(-1 + r)}{cir(i - 2 - \sqrt{-16cir + i^2 + 4(1 - i)})} \end{aligned}$$

and

$$\begin{aligned} z^2_{ch}(i, r) &= \frac{-i + 2 - \sqrt{-16cir + i^2 + 4(1 - i)}}{4cir} \\ z^2_{cc}(i, r) &= -\frac{4\left(\left(-\frac{i}{8} + \frac{1}{4}\right)\sqrt{-16cir + i^2 + 4(1 - i)} + cir - \frac{i^2}{8} + \frac{i}{2} - \frac{1}{2}\right)(-1 + r)}{cir(i - 2 + \sqrt{-16cir + i^2 + 4(1 - i)})}. \end{aligned}$$

We begin by substituting the two candidate solution sets into Equation (6.9). To simulate the effect of an interchange fee cap, we exogenously reduce the interchange fee i by 10% from its benchmark equilibrium level defined in Equation (6.11). We then solve the model by maximizing profit with respect to the reward r.[44] Our analysis identifies the second solution set as the economically relevant one.

Given the complexity of the resulting expressions, we proceed numerically. We examine how the equilibrium changes by varying the cost parameter c. This parameter influences aggregate consumption X, which in turn determines the demand elasticity. It is through this channel that c ultimately governs the direction of the change in the equilibrium reward.

We present illustrative results for three different values of c in Table 1. Most of these results are consistent with those derived under the constant elasticity demand, with one notable exception. Under a linear demand, when elasticity is relatively low (i.e., $c = 0.4$), an interchange fee cap can make cash users worse off by increasing the tax they pay.[45]

It is important to note that the numerical results reported in Table 1 should be interpreted qualitatively rather than quantitatively. In particular, the equilibrium values of i and r are substantially higher than those observed in practice. Our model abstracts from real world factors such as competition among networks, regulatory constraints (including the possibility of triggering regulation) and behavioral considerations, which further limit the quantitative accuracy of the predictions. In other words, the model represents a setting with an upstream monopolist that competes only against cash and the no-purchase option, both of which have fixed prices. Accordingly, the results should be viewed as illustrating the direction of effects and relative impacts rather than their precise magnitudes.

Key Observations and Intuition from Table 1

- Effect on Rewards (r): The network's strategic response to the cap on i depends on demand elasticity.
  - Relatively low Elasticity ($c = 0.4$, $\varepsilon \approx -2.72$ and $c = 0.45$, $\varepsilon \approx -3.10$): A lower i does not significantly boost total consumption. The network finds it more profitable to increase r to

[44] We confirmed that our qualitative conclusions are robust to smaller reductions in the interchange fee, such as 1% or 0.1%.

[45] The simulations underlying Table 1 were produced with MATLAB code. The code computes equilibrium outcomes across a wide range of parameter configurations and all results are qualitatively similar to those presented in the table. The code is available from the authors upon request.

Table 1: Equilibrium Effects of a 90% Interchange Fee Cap under Linear Demand and Perfect Competition

	$c = 0.4$		$c = 0.45$		$c = 0.5$	
Variables	Unregulated (i^*)	Regulated ($0.9i^*$)	Unregulated (i^*)	Regulated ($0.9i^*$)	Unregulated (i^*)	Regulated ($0.9i^*$)
Interchange Fee (i)	0.7528	0.6776	0.6976	0.6278	0.6452	0.5806
Reward (r)	0.2721	0.3715	0.2649	0.2979	0.2581	0.2538
Cash Tax (z_{ch})	2.2969	2.3613	2.0972	1.9008	1.9375	1.7141
Credit Tax (z_{cc})	1.6719	1.4841	1.5417	1.3345	1.4375	1.2791
Credit Share (μ)	0.7500	0.8508	0.7500	0.7548	0.7500	0.7175
Network Profit ($\pi_{\mathcal{N}}$)	0.1097	0.1000	0.0938	0.0851	0.0791	0.0725
Consumption (X)	0.2690	0.3540	0.2440	0.3370	0.2188	0.2990
Elasticity (ε)	-2.72	-1.82	-3.10	-1.97	-3.57	-2.34

stimulate spending, partially offsetting the lost revenue from the lower fee.

– Relatively high Elasticity ($c = 0.5$, $\varepsilon \approx -3.57$): A lower i triggers a substantial increase in total consumption. The network lowers r to recoup some margin, relying on the volume increase to protect its profits.

These results align with the findings and intuition under the constant elasticity demand specification (see Result 1).

- Effect on Credit Users: The price for credit users is determined by the tax $z_{cc} \equiv \frac{1-r}{1-\mu i}$.[46] A change in z_{cc} is driven by changes in i, r and μ:

  – Direct Effect of $\downarrow i$: A lower interchange fee directly reduces the denominator $(1 - \mu i)$, which lowers z_{cc} (a beneficial force).

  – Indirect Effect via r:

    * If r increases (low elasticity), the numerator $(1 - r)$ decreases, which further lowers z_{cc} (reinforcing the benefit).

    * If r decreases (high elasticity), the numerator $(1 - r)$ increases, which raises z_{cc} (a countervailing force).

  – Indirect Effect via μ: A change in the credit share μ affects the denominator.

[46] The price paid by credit card users is given by $P^{cc} = z_{cc}P^m$. Since $P^m = c$, and hence is unaffected by z_{cc}, it is sufficient to focus on z_{cc} to determine the direction of the change in the price paid by credit card users. The same reasoning applies to cash users.

* If μ increases (low elasticity), due primarily to higher rewards, it diminishes the beneficial effect of lower i in the denominator.
* If μ decreases (high elasticity), due primarily to lower rewards, it amplifies the beneficial effect of lower i in the denominator.

Net Effect: The results show that the direct, beneficial, effect of the lower interchange fee dominates the potentially countervailing movements in r and μ in all cases, leading to a lower final tax z_{cc} and thus a lower price for credit users. These results are consistent with the ones derived under the constant elasticity demand (see Result 2).

- Effect on Cash Users: The price for cash users is determined by the tax $z_{ch} \equiv \frac{1}{1-\mu i}$. A change in z_{ch} is driven by changes in i and μ:
  - Direct Effect of $\downarrow i$: A lower interchange fee directly reduces the denominator $(1 - \mu i)$, which lowers z_{ch} (a beneficial force).
  - Indirect Effect via μ: The change in credit share μ is critical.
    * If μ increases substantially (low elasticity, $c = 0.4$), due to higher rewards, it decreases the denominator $(1 - \mu i)$, which raises z_{ch}. The regulation harms the cash users.
    * If μ is stable or decreases (higher elasticity, $c = 0.45, 0.5$), due to lower rewards, the countervailing force through μ is weak or becomes reinforcing. This allows the direct benefit of lower i to dominate, lowering z_{ch} and benefiting cash users.

Under the constant elasticity demand specification, we find no evidence that cash users incur a welfare loss following the regulation (see Result 2). In contrast, under linear demand, such losses can occur. The intuition is as follows.

Under linear demand, as aggregate consumption rises due to a lower weighted-average tax, demand becomes less elastic. This does not occur under constant elasticity demand, where elasticity remains unchanged. Consequently, the payment network, seeking to boost consumption when the initial demand elasticity is relatively low, increases the reward more substantially under linear demand to counteract the reduction in elasticity. This relatively larger reward promotes greater credit card adoption and can leave cash users worse off.

- Effect on Credit Share (μ): The share of credit users, $\mu = \frac{1}{2} + c \cdot (z_{ch} - z_{cc})$, changes primarily with the tax differential $(z_{ch} - z_{cc})$.
  - When r increases (low elasticity), the gap $(z_{ch} - z_{cc})$ widens, making credit more attractive and increasing μ.
  - When r decreases (high elasticity), the gap $(z_{ch} - z_{cc})$ narrows, making credit less attractive and decreasing μ.
- Network Profit ($\pi_{\mathcal{N}}$): As expected, the interchange fee cap reduces the network's profit in all scenarios, as it directly constrains a key instrument and forces a suboptimal (from the network's perspective) adjustment of r and μ.

As in the constant elasticity demand model, a regulation that caps the interchange fee impacts the three endogenous variables—i, r, and μ—which in turn influences the effects of both forms of taxation. Our analysis under linear demand reveals that, while such a regulation can improve welfare for both consumer groups (as was the case under constant elasticity), it cannot worsen welfare for both. This latter result is new; in the constant elasticity model, both groups were always made better off.

The outcome highlights a clear trade-off. A deterioration for cash users—which occurs when rewards rise under relatively inelastic demand, exacerbating the cross-subsidization—typically benefits credit card users. Conversely, a deterioration for credit card users, which happens when rewards fall under relatively elastic demand, mitigates the cross-subsidization, ultimately benefiting cash users.

In summary, given that merchant profits remain at zero, the regulation reallocates surplus away from the payment network. This surplus is transferred either to all consumers, if they are collectively better off, or, in scenarios where cash users are made worse off, to credit card users, effectively resulting in a transfer from both the network and cash users to cardholders.

6.4.3 Imperfectly competitive merchant market, $\gamma > 0$

The payment network chooses z_{cc} and z_{ch} to maximize its profit, as given by equation (6.6). Under imperfect competition, incomplete pass-through implies that the incidence of an interchange fee cap depends on the intensity of product market competition. Inspecting (6.6) reveals how merchant market power, captured by γ, interacts with demand elasticity through the price. Specifically, $-\frac{\gamma}{\varepsilon} = -\frac{\gamma X}{1-X}$,

indicating that the impact of elasticity becomes more pronounced as γ increases—at $\gamma = 0$, this effect vanishes.[47] This introduces a novel *elasticity effect* not present under constant elasticity demand or linear demand with perfect competition.

When $\varepsilon' > 0$ (as with linear, subconvex demand), a decrease in aggregate output X makes demand more elastic. In this setting, in addition to the effects already identified by our analysis so far, an increase in the network's taxes z_{cc} and z_{ch} generates a negative effect on marginal profitability: a higher tax reduces aggregate output X, which in turn increases demand elasticity, leading merchants to lower their prices and thereby reducing the network's revenue. A lower γ–reflecting stronger product market competition–weakens this effect. Consequently, the network optimally responds by increasing its tax.[48]

To measure the extent to which consumers bear the burden of the credit card "tax," we define the tax incidence for each payment group. Let $\widetilde{P}$ denote the equilibrium price in an all-cash economy (the benchmark without credit cards, see (6.7)). The incidence for credit card and cash users is then given by

$$\tau_{cc} \equiv \frac{P^{cc} - \widetilde{P}}{P^{cc} - P^{m}}, \qquad \tau_{ch} \equiv \frac{P^{ch} - \widetilde{P}}{P^{ch} - P^{m}}.$$

A value of τ near one implies that consumers bear nearly all of the tax burden, meaning the credit card tax is passed through almost entirely to consumers. Conversely, a value near zero implies that merchants bear most of the burden.

As in the perfect competition case, our results are numerical. Following the structure of Table 1 for the case of $\gamma = 0$, we present, in Tables 2 and 3 below, equilibrium values–both unregulated and regulated–for different levels of γ, in order to assess the impact of merchant market power. The numbers in these tables are not correct. Once we have the new Matlab code we will put the correct numbers.[49]

The main findings obtained under perfect competition, regarding the impact of an interchange fee cap regulation, carry over when we allow for imperfect competition. In particular, an interchange fee

[47] This is intuitive. In a perfectly competitive market, demand elasticity has no effect on pricing, as price is driven to marginal cost regardless of elasticity. In contrast, for a monopolist, market elasticity plays a crucial role in determining price.

[48] This mechanism is closely related to the effect of competition on markups discussed in Mrázová and Neary (2017). As the authors note on page 3840: "Hence, if globalization reduces incumbent firms' sales in their home markets, it is associated with a higher elasticity and so a lower markup if and only if demand is subconvex."

[49] These tables are computed using the same MATLAB code as Table 1. For space considerations, we only report results for two representative values of γ and all additional simulations yield qualitatively similar outcomes. The code is available from the authors upon request.

Table 2: Equilibrium Effects of a 90% Interchange Fee Cap under Linear Demand and Imperfect Competition, $\gamma = 0.1$

	$c = 0.4$		$c = 0.45$		$c = 0.5$	
Variables	Unregulated (i^*)	Regulated ($0.9i^*$)	Unregulated (i^*)	Regulated ($0.9i^*$)	Unregulated (i^*)	Regulated ($0.9i^*$)
Interchange Fee (i)	0.7416	0.6675	0.6877	0.6190	0.6368	0.5731
Reward (r)	0.2666	0.3466	0.2599	0.2904	0.2536	0.2507
Cash Tax (z_{ch})	2.2533	2.2388	2.0654	1.8887	1.9142	1.7088
Credit Tax (z_{cc})	1.6525	1.4628	1.5285	1.3401	1.4287	1.2805
Credit Share (μ)	0.7500	0.8290	0.7500	0.7601	0.7500	0.7238
Network Profit ($\pi_{\mathcal{N}}$)	0.1132	0.1047	0.0963	0.0888	0.0809	0.0750
Consumption (X)	0.2789	0.3289	0.2517	0.3070	0.2250	0.2733
Elasticity (ε)	-2.58	-2.04	-2.97	-2.25	-3.44	-2.66
Agg. Mer. Profit (Π)	0.0045	0.0068	0.0039	0.0064	0.0034	0.0053
Price cc users pay (P^{cc})	0.6876	0.6153	0.7118	0.6310	0.7357	0.6653
Price ch users pay (P^{ch})	0.9376	0.9417	0.9618	0.8892	0.9857	0.8878
Tax Incidence, cc (τ_{cc})	0.8584	0.8257	0.8605	0.8180	0.8618	0.8222
Tax Incidence, ch (τ_{ch})	0.9263	0.9349	0.9308	0.9303	0.9352	0.9296

Table 3: Equilibrium Effects of a 90% Interchange Fee Cap under Linear Demand and Imperfect Competition, $\gamma = 0.5$

	$c = 0.4$		$c = 0.45$		$c = 0.5$	
Variables	Unregulated (i^*)	Regulated ($0.9i^*$)	Unregulated (i^*)	Regulated ($0.9i^*$)	Unregulated (i^*)	Regulated ($0.9i^*$)
Interchange Fee (i)	0.6585	0.5927	0.6263	0.5637	0.5895	0.5305
Reward (r)	0.2321	0.2473	0.2318	0.2391	0.2303	0.2247
Cash Tax (z_{ch})	1.9760	1.8509	1.7924	1.7494	1.7924	1.6459
Credit Tax (z_{cc})	1.5172	1.3932	1.4486	1.3312	1.3796	1.2761
Credit Share (μ)	0.7500	0.7756	0.7500	0.7600	0.7500	0.7397
Network Profit ($\pi_{\mathcal{N}}$)	0.1354	0.1322	0.1111	0.1076	0.0907	0.0874
Consumption (X)	0.3472	-	0.2989	-	0.2586	-
Elasticity (ε)	-1.88	-	-2.34	-	-2.87	-
Agg. Mer. Profit (Π)	0.0045	0.0068	0.0039	0.0064	0.0034	0.0053
Price cc users pay (P^{cc})	0.6876	0.6153	0.7118	0.6310	0.7357	0.6653
Price ch users pay (P^{ch})	0.9376	0.9417	0.9618	0.8892	0.9857	0.8878
Tax Incidence, cc (τ_{cc})	0.8584	0.8257	0.8605	0.8180	0.8618	0.8222
Tax Incidence, ch (τ_{ch})	0.9263	0.9349	0.9308	0.9303	0.9352	0.9296

cap continues to raise the equilibrium reward when demand elasticity is relatively low (i.e., $c = 0.4$ and $c = 0.45$) and to reduce the reward when elasticity is relatively high (i.e., $c = 0.5$). Moreover, across all cases, all consumers face lower credit card taxes following the regulation.[50]

[50]When $\gamma = 0$, cash users pay a higher tax after regulation when elasticity is low ($c = 0.4$), due to a substantial increase in

There are, however, new results stemming from imperfect pass-through. When the network lowers its taxes, merchants respond by raising their price, an optimal response to the less elastic demand, making the net effect on consumer prices, in general, ambiguous. From Table 2, we see that credit card users pay a lower price, but cash users can pay a higher price when $c = 0.4$, even though the taxes decline in all instances. Cash users are worse off when card adoption increases significantly after the regulation, such that z_{ch} does not decrease sufficiently relative to the increase in P^m.

Merchants are better off after the regulation, as they benefit from both a higher markup, $P^m - c$, and higher aggregate consumption X resulting from the decline in taxes. Thus, under imperfect competition, a regulation redistributes surplus away from the network—and sometimes from cash users—toward credit card users and merchants. Although the regulation makes both credit card users and merchants better off, the tax burden for credit card users decreases. For cash users, the tax burden decreases only when the price they pay decreases (for $c = 0.45$ and $c = 0.5$).

Also, as expected, the tax incidence for consumers decreases as the merchant market becomes less competitive (when $\gamma = 0$, $\tau_{cc} = \tau_{ch} = 1$, consumers bear all the tax burden).

The effect of stronger product market competition (i.e., lower γ) is higher equilibrium taxes. This result contrasts with the constant elasticity case (see Proposition 2), where more intense product market competition lowers the tax paid by credit card users and also lowers the tax paid by cash users whenever demand elasticity is relatively low. The difference arises from the elasticity effect, which dominates the effect we identified in the constant elasticity case—an effect that is also at work here. Under linear demand, increased product market competition weakens the elasticity effect. Consequently, as we have discussed above, the network becomes less reluctant to lower taxes as competition intensifies, leading to higher equilibrium taxes in more competitive product markets.

7 Conclusion

The model developed in this paper features a two-sided market with a monopoly payment network, interchange fees in the form of an ad valorem markup, consumers who can choose to pay with either credit or

credit card adoption, μ. This is not the case in Tables 2 and 3. Under imperfect competition, the increase in credit card adoption is more modest (when $c = 0.4$). This is because elasticity is lower due to higher output when the merchant market is more competitive and the network has weaker incentives to increase the reward.

cash, merchants with market power and elastic demand. Credit card taxes emerge from the interchange fees, rewards and the rate of credit adoption, creating a wedge between the prices consumers pay and the prices merchants receive, all endogenously determined. The model establishes an equilibrium where aggregate demand and merchant profits are influenced by these credit card taxes.

Merchants argue that the high fees for processing credit card transactions erode their profits and force them to raise prices, negatively affecting consumers. Those concerns have prompted governments worldwide to introduce regulations aimed at curbing these fees. Networks and banks, on the other hand, contend that lowering interchange fees would compel them to reduce the rewards offered to credit card users, thereby disadvantaging their customers.

Under the assumptions of our model, we show that capping interchange fees can enhance welfare for both cardholders and cash users by mitigating the implicit taxes embedded in cash and credit card transactions–even when rewards decline. Given endogenous pricing and constant elasticity of demand, the reduction in rewards, which occurs when product demand elasticity is relatively high, is consistently smaller than the decrease in interchange fees, yielding a net gain for all consumers. This effect is driven by the network's strategic response to competition from cash as an alternative payment method. Contrary to conventional wisdom, when product demand elasticity is relatively low, a fee cap induces the network to increase equilibrium rewards–an outcome that significantly amplifies cardholders' welfare, but under a linear demand it may hurt cash users.

The imposition of a fee cap diminishes the profit of the payment network. The resulting surplus redistribution depends on the underlying demand specification: under constant elasticity demand, the surplus is reallocated to merchants and consumers collectively; under linear demand, when cash users experience a welfare loss, the redistribution favors credit card users and merchants. In this latter case, the transfer occurs from the payment network and cash users to cardholders and merchants.

More broadly, our framework identifies three central sources of social inefficiency in the credit card payment system: network market power, merchant market power, and reward-induced distortions in consumer payment behavior. In our baseline model, these mechanisms contribute to excessive reliance on credit cards and suboptimal consumption. We show that a cap on interchange fees can simultaneously mitigate both inefficiencies when demand elasticity lies within an intermediate range. In such cases, the regulation lowers implicit taxes for both cash and credit card users, improves consumption efficiency,

and curtails overuse of credit.

As in other models, we find that credit card users face a lower effective tax than cash users, due to the rewards offered for card-based purchases. Under certain conditions, these rewards can be so generous that the tax on credit card transactions effectively becomes a subsidy–cardholders end up paying less, on net, than the amount merchants receive, even below marginal cost in some cases.

As competition in the product market intensifies, the equilibrium tax paid by credit card users decreases under constant demand elasticity but increases under linear demand. For cash users, the tax decreases only when demand elasticity is relatively high, yet it also rises under linear demand. This divergence stems from the elasticity effect, which emerges once we depart from the constant elasticity specification. As this effect becomes dominant, it disincentivizes the network from lowering the taxes.

Finally, under a constant elasticity demand, aggregate merchant profits can increase with intensified product market competition due to the significant double marginalization problem within a certain range of the market conduct parameter. More intense product market competition prompts the network, which competes with cash, to lower the tax for credit card users. This, in turn, significantly increases their consumption, offsetting the negative impact of competition on merchant profits. In models that do not include cash as an alternative payment mode, merchant profits strictly decline as competition increases.

Our results relate to the well-known *waterbed effect*, whereby regulation of one price margin induces offsetting adjustments elsewhere. In mobile telephony, (Genakos and Valletti, 2011, 2015) empirically document that reductions in regulated termination rates were partially offset by higher retail prices. Policymakers, academics and industry participants in payment markets are well aware that interchange fee caps can trigger rebalancing. Our contribution is to characterize the *direction* of the response: in a two-sided payment market, the network's adjustment of equilibrium rewards and the associated effect on credit card adoption, depends on demand elasticity.

The model generates testable implications about the effect of product market competition, the degree of differentiation between cash and credit and the elasticity of demand on the taxes for credit and cash users.

We have made significant progress in developing a more realistic model of a payment network. However, there remain at least two simplifying assumptions that we aim to relax in future research. First, in our model, we assume the presence of a single monopoly payment network. In reality, multiple

payment card networks typically exist. Second, we have assumed Bertrand competition among issuing and acquiring banks. At the expense of substantial complication, introducing market power at this layer of the network would likely yield additional effects and insights.

References

Baxter, William F (1983), "Bank interchange of transactional paper: Legal and economic perspectives." *Journal of Law and Economics*, 26, 541–588.

Bedre-Defolie, Özlem and Emilio Calvano (2013), "Pricing payment cards." *American Economic Journal: Microeconomics*, 5, 206–231.

Bresnahan, Timothy F (1981), "Duopoly models with consistent conjectures." *American Economic Review*, 71, 934–945.

Chakravorti, Sujit and Ted To (2007), "A theory of credit cards." *International Journal of Industrial Organization*, 25, 583–595.

Chan, Iris, Sophia Chong, and Stephen Mitchell (2012), "The personal credit card market in Australia: Pricing over the past decadel bulletin–March 2012." *Reserve Bank of Australia*.

Chatterjee, Satyajit, Dean Corbae, Kyle P Dempsey, and José-Víctor Ríos-Rull (2025), "Credit scores and inequality across the life cycle." Technical report, National Bureau of Economic Research.

Delipalla, Sofia and Michael Keen (1992), "The comparison between ad valorem and specific taxation under imperfect competition." *Journal of Public Economics*, 49, 351–367.

Edelman, Benjamin and Julian Wright (2015), "Price coherence and excessive intermediation." *Quarterly Journal of Economics*, 130, 1283–1328.

Evans, David S and Abel M Mateus (2011), "How changes in payment card interchange fees affect consumers fees and merchant prices: An economic analysis with applications to the european union." *Available at SSRN 1878735*.

Felt, Marie-Hélène, Fumiko Hayashi, Joanna Stavins, and Angelika Welte (2023), "Regressive effects of payment card pricing and merchant cost pass-through in the United States and Canada." *Journal of Banking & Finance*, 154, 106968.

Foster, Kevin, Claire Greene, and Joanna Stavins (2024), "2023 survey and diary of consumer payment choice: Summary results." *Federal Reserve Bank of Atlanta Research Data Report*.

Genakos, Christos and Tommaso Valletti (2011), "Testing the waterbed effect in mobile telephony." *Journal of the European Economic Association*, 9, 1114–1142.

Genakos, Christos and Tommaso Valletti (2015), "Evaluating a decade of mobile termination rate regulation." *The Economic Journal*, 125, F31–F48.

Jeon, Doh-Shin and Patrick Rey (2022), "Platform competition and app development." *Unpublished manuscript, Toulouse Sch. Econ., Toulouse.*

Jullien, Bruno, Alessandro Pavan, and Marc Rysman (2021), "Two-sided markets, pricing, and network effects." In *Handbook of Industrial Organization*, volume 4, 485–592, Elsevier.

Kay, Benjamin S, Mark D Manuszak, and Cindy M Vojtech (2018), "Competition and complementarities in retail banking: Evidence from debit card interchange regulation." *Journal of Financial Intermediation*, 34, 91–108.

Manuszak, Mark D and Krzysztof Wozniak (2017), "The impact of price controls in two-sided markets: evidence from US debit card interchange fee regulation." *Finance and Economics Discussion Series 2017-074*.

Miklós-Thal, Jeanine and Greg Shaffer (2021), "Input price discrimination by resale market." *RAND Journal of Economics*, 52, 727–757.

Miklos-Thal, Jeanine and Greg Shaffer (2021), "Pass-through as an economic tool: on exogenous competition, social incidence, and price discrimination." *Journal of Political Economy*, 129, 323–335.

Mrázová, Monika and J Peter Neary (2017), "Not so demanding: Demand structure and firm behavior." *American Economic Review*, 107, 3835–3874.

Mrázová, Monika and J Peter Neary (2019), "Selection effects with heterogeneous firms." *Journal of the European Economic Association*, 17, 1294–1334.

Mukharlyamov, Vladimir and Natasha Sarin (forthcoming), "Price regulation in two-sided markets: Empirical evidence from debit cards." *Journal of Financial Economics*.

Rochet, Jean-Charles and Jean Tirole (2002), "Cooperation among competitors: Some economics of payment card associations." *Rand Journal of Economics*, 549–570.

Rochet, Jean-Charles and Jean Tirole (2006), "Two-sided markets: a progress report." *RAND Journal of Economics*, 37, 645–667.

Rochet, Jean-Charles and Jean Tirole (2011), "Must-take cards: Merchant discounts and avoided costs." *Journal of the European Economic Association*, 9, 462–495.

Ronnen, Uri (1991), "Minimum quality standards, fixed costs, and competition." *The RAND Journal of Economics*, 490–504.

Rysman, Marc (2009), "The economics of two-sided markets." *Journal of Economic Perspectives*, 23, 125–143.

Salop, Steven C (1979), "Monopolistic competition with outside goods." *Bell Journal of Economics*, 141–156.

Schmalensee, Richard (2002), "Payment systems and interchange fees." *Journal of Industrial Economics*, 50, 103–122.

Schuh, Scott, Oz Shy, and Joanna Stavins (2010), *Who gains and who loses from credit card payments?: Theory and calibrations*. DIANE Publishing.

Seade, Jesus (1980), "On the effects of entry." *Econometrica*, 479–489.

Shy, Oz and Zhu Wang (2011), "Why do payment card networks charge proportional fees?" *American Economic Review*, 101, 1575–1590.

Wang, Lulu (2023), "Regulating competing payment networks." *URL: https://luluywang. github. io/PaperRepository/payment_jmp. pdf (last accessed on 10 November 2023)*.

Wang, Zhu (2010), "Market structure and payment card pricing: What drives the interchange?" *International Journal of Industrial Organization*, 28, 86–98.

Wang, Zhu, Scarlett Schwartz, and Neil Mitchell (2014), "The impact of the durbin amendment on merchants: A survey study." *Economic Quarterly*, 183–208.

Wang, Zhu and Julian Wright (2017), "Ad valorem platform fees, indirect taxes, and efficient price discrimination." *RAND Journal of Economics*, 48, 467–484.

Wang, Zhu and Julian Wright (2018), "Should platforms be allowed to charge ad valorem fees?" *Journal of Industrial Economics*, 66, 739–760.

Weyl, E Glen and Michal Fabinger (2013), "Pass-through as an economic tool: Principles of incidence under imperfect competition." *Journal of Political Economy*, 121, 528–583.

Wright, Julian (2004), "The determinants of optimal interchange fees in payment systems." *Journal of Industrial Economics*, 52, 1–26.

A Appendix: Proofs

A.1 Proof of Lemma 1

The second derivative of the network profit, (4.10), with respect to z_{cc} is given by

$$\frac{\partial^2 \pi_{\mathcal{N}}}{\partial z_{cc}^2} = -\frac{A}{t \cdot (k-\gamma)^2},$$

where $A \equiv c z_{cc}^{-2-k} \left(\frac{k-\gamma}{ck}\right)^k k^2 (c \cdot (k-1)(k-2) z_{cc}^2 - 2((c z_{ch} + t/4)k - \gamma \cdot (t/4))(k-1) z_{cc} + (1+k)((c z_{ch}^2 - (1/2) t z_{ch} + t)k + t\gamma \cdot (z_{ch} - 2)/2))$. When we evaluate the second derivative at the solutions to the first order conditions, (4.11), it is negative if and only if $t < \frac{16ck}{k-\gamma}$. The second derivative of the network profit with respect to z_{ch} is given by,

$$\frac{\partial^2 \pi_{\mathcal{N}}}{\partial z_{ch}^2} = -\frac{\left(\frac{k-\gamma}{ck}\right)^k z_{cc}^{-k} k^2 2c^2}{t(k-\gamma)^2} < 0.$$

The determinant of the Hessian matrix of $\pi_{\mathcal{N}}(z_{cc}, z_{ch})$ is given by

$$\frac{\left(k(z_{cc} - z_{ch})^2 (k-1)c^2 + \frac{3\left(\left(z_{cc} - \frac{z_{ch}}{3} - \frac{2}{3}\right)k - z_{cc} + \frac{z_{ch}}{3} - \frac{2}{3}\right) t \cdot (k-\gamma) c}{2} + \frac{t^2 (k-\gamma)^2}{8}\right) 2k^4 \left(\frac{k-\gamma}{ck}\right)^{2k} c^2 z_{cc}^{-2(1+k)}}{t^2 (k-\gamma)^4},$$

which when evaluated at the solutions to the first order conditions, (4.11), is positive if and only if $t < \frac{16ck}{k-\gamma}$.

A.2 Proof of Proposition 1

Note that $z^*_{cc} < 1$ if and only if $k > \gamma + \frac{16c}{t}$, or $t > \frac{16c}{k-\gamma}$. The effect of demand elasticity on the taxes is: $\frac{\partial z^*_{cc}}{\partial k} < 0 \Leftrightarrow t < \frac{16c}{1-\gamma}$, which is always satisfied since $\frac{16c}{1-\gamma} > \frac{16ck}{k-\gamma}$ (from the second order condition), $\frac{\partial z^*_{ch}}{\partial k} < 0 \Leftrightarrow t < \frac{16k^2c}{3\gamma k^2-8\gamma k+k^2+4\gamma}$ and $\frac{\partial (z^*_{ch}-z^*_{cc})}{\partial k} > 0$. Moreover, $\frac{16k^2c}{3\gamma k^2-8\gamma k+k^2+4\gamma} < \frac{16c}{1-\gamma}$ and $\frac{16k^2c}{3\gamma k^2-8\gamma k+k^2+4\gamma} > \frac{16ck}{k-\gamma} \Leftrightarrow k < \frac{4}{3}$. Therefore, the tax credit card users pay is decreasing, and the difference between the taxes is increasing, as demand becomes more elastic. The effect of elasticity on the tax cash users pay is ambiguous. When $k < \frac{4}{3}$, z_{ch} is decreasing as k increases, while when $k > \frac{4}{3}$ it is decreasing if and only if $t < \frac{16k^2c}{3\gamma k^2-8\gamma k+k^2+4\gamma}$.

The equilibrium price credit card users pay is $P^{cc} = z^*_{cc}P^m = \frac{(16c-t)k^2+\gamma kt}{16(k-1)(k-\gamma)} \lesseqgtr P^m$ and the equilibrium price cash card users pay is $P^{ch} = z^*_{ch}P^m = \frac{(16c+3t)k^2-(3\gamma+4)tk+4\gamma t}{16(k-1)(k-\gamma)} > P^m$. It can be verified that both prices are increasing in γ.

A.3 Proof of Proposition 2

The effect of merchant market power on the taxes is: $\frac{\partial z^*_{cc}}{\partial \gamma} > 0$, $\frac{\partial z^*_{ch}}{\partial \gamma} > 0 \Leftrightarrow k < \frac{4}{3}$ and $\frac{\partial (z^*_{ch}-z^*_{cc})}{\partial \gamma} < 0$. The effect of the degree of differentiation between cash and credit on the taxes is: $\frac{\partial z^*_{cc}}{\partial t} < 0$, $\frac{\partial z^*_{ch}}{\partial t} > 0 \Leftrightarrow k > \frac{4}{3}$ and $\frac{\partial (z^*_{ch}-z^*_{cc})}{\partial t} > 0$.

From (4.12), it can be verified that $\frac{\partial i^*}{\partial \gamma} > 0 \Leftrightarrow k < \frac{4}{3}$ and $\frac{\partial r^*_1}{\partial \gamma} < 0$.

It can be verified, using (4.13), that $\frac{\partial \pi_{\mathcal{N}}}{\partial \gamma} < 0$, if $t < \frac{16ck}{k-\gamma}$.

Using (4.14), while $\frac{\partial \Pi^m}{\partial \gamma} > 0$ for low γ, because $\Pi^m = 0$ at $\gamma = 0$, we have shown numerically that for a wide range of permissible parameter values Π^m is inverse U-shaped in γ.

A.4 Proof of Proposition 3

We differentiate (4.13) with respect to t to derive

$$\frac{\partial \pi_{\mathcal{N}}(z^*_{cc}, z^*_{ch})}{\partial t} = \frac{\left(\frac{16(k-\gamma)}{ck}\right)^k \left(\frac{(16c-t)k+t\gamma}{c\cdot(k-1)}\right)^{-k}}{16} > 0,$$

following from the condition in Lemma 1. Hence, a higher t, which serves as a proxy for a higher network market power, increases the equilibrium profits of the network.