Proportional vs. Unit Fees: Welfare and Incentive

Dingwei Gu, Zhiyong Yao, and Wen Zhou*

November 29, 2019

Abstract

We compare social welfare and business profitability between proportional and unit fees that are widely used in platform markets, retailing, technology licensing, and taxation. If the demand is sub-convex or constant elastic, then proportional fees weakly improve welfare. If the demand is super-convex, then proportional fees improve welfare if there is sufficient competition among networks. Allowing merchants to multi-home on competing networks further increases the likelihood of welfare improvement. The conditions for welfare deterioration are also identified. For profitability, while a monopoly network always benefits from proportional fees, competing networks may be hurt. If network competition is weak, the supply chain (i.e., all networks and merchants combined) gains if and only if consumers also gain. If network competition is strong, the supply chain always loses after switching to proportional fees. In that case, networks and merchants may each gain or lose, but they can never both gain.

Keywords: proportional fee, unit fee, agency model, wholesale model

JEL Codes: D2, D4, L1, L4

1 Introduction

When a payment card is used to settle transactions, the card network collects an interchange fee from the merchants. For a long time, merchants have been complaining that the high fees had inflated the transaction costs. A particular concern is that the fees are usually proportional to the transaction value (henceforth referred to as proportional fees), even though the cost of executing each transaction does not seem to vary much with the price. More reasonable, argue the merchants, would be a fixed, per-unit transaction fee (referred to as unit fees) (Shy and Wang 2011).

*Gu: School of Economics, Fudan University; Shanghai Institute of International Finance and Economics, Dingwei.gu@gmail.com; Yao: School of Management, Fudan University, yzy@fudan.edu.cn; Zhou: Faculty of Business and Economics, University of Hong Kong, wzhou@hku.hk. We thank Gary Biglaiser, Yongmin Chen, Kim-sau Chung, Andrei Hagiu, Rongzhu Ke, Robin Lee, Melody Pei-yu Lo, Matthew Mitchell, Jeffrey Zwiebel, and participants of 2019 WEAI, 2019 IIOC and 2019 HKUST IO conference for helpful comments and suggestions. This research is supported by China National Natural Science Foundation (71873036).

Beyond payment card and other platform markets, proportional and unit fees are also used in a wide range of economic activities. For example, the most common business models in retailing and supply chains are the agency and wholesale models, the key difference of which is that the agency model adopts proportional fees while the wholesale model uses unit fees (Gaudin and White 2014a; Johnson 2017). Government levies either an ad-valorem tax (proportional to the value) or a specific tax (a fixed amount per unit of the good or transaction). Technology licensing royalty is calculated based on either the value of the whole product or the component specific to the technology, which are equivalent to ad valorem and per-unit royalty rates respectively (Llobet and Padilla 2016).

In recent years, the agency model has become increasingly popular among platforms for the sale of digital contents such as music, electronic books, hotel and airline booking, and cell phone applications, and has been at the center of some high profile antitrust lawsuits. The ebook industry started with the wholesale model on Amazon's Kindle, but Apple entered the competition by signing agency contracts with five major publishers, who in turn forced Amazon to adopt the agency model. In 2013, Apple was ordered by Department of Justice (DOJ) to abandon the agency agreements.[1] More recently, Apple's practice of charging 30 per cent commission to software sold in its Apple Store has been challenged by consumers and is currently considered by Supreme Court. In technology licensing, recent court rulings show a trend of supporting per-unit royalty.[2]

These rulings are in sharp contrast with theoretical findings, which unanimously support the proportional fee regime. For payment cards, Shy and Wang (2011) demonstrate that proportional fees benefit a monopoly network and social welfare at the expense of merchants. For the ebook industry and retailing, respectively, Gaudin and White (2014a) and Johnson (2017) find that agency model generates larger social welfare. For patent licensing, Llobet and Padilla (2016) prove that ad valorem royalty welfare-dominate per-unit royalty. How to understand the contrast between court rulings and theoretical findings? Is there any mistake in the studies or decisions, or the two can be reconciled in a more general framework where each may be right but only under certain conditions?

In this paper, we revisit the comparison between proportional and unit fees. For ease of understanding we will imagine the underlying industry as payment cards and therefore talk about networks and merchants, but the results apply directly to other settings such as internet and telecom platforms, retailing, taxation, and licensing. All networks are assumed to charge the same type of fees, either unit or proportional, to both consumers and merchants. They compete by choosing independently their fee levels (for an exogenously given fee type), after which single-homing or multi-homing merchants carry out imperfect competition. We

[1] More details of the case are discussed in Section 6.

[2] As cited by Llobet and Padilla (2016), in the 2014 ruling of *Ericsson Inc. v. D-Link Systems Inc.*, the United States Court of Appeals for the Federal Circuit suggested that ad valorem royalties be avoided (http://www.cafc.uscourts.gov/sites/default/files/opinions-orders/13-1625.Opinion.12-2-2014.1.PDF). In 2015, DOJ praised the Institute of Electrical and Electronics Engineers (IEEE) for incorporating the per-unit royalty in the organization's updated patent policy (Hesse 2015).

compare the two fee schemes' impacts on social welfare and the profits of networks and merchants.

The welfare consequence is found to depend crucially on the properties of the demand. For certain demands, the degree of network competition and merchant multi-homing also play important roles. If the demand is sub-convex or constant elastic, then proportional fees (weakly) improve welfare. If the demand is super-convex, then proportional fees improve welfare if there is sufficient competition among networks (as quantified by the number of networks or degree of network substitutability). Finally, allowing merchants to multi-home on competing networks further increases the likelihood of welfare improvement. These findings suggest that court rulings and theoretical literature do not necessarily contradict each other; they may both be correct, albeit under different conditions. The flip side is that neither holds unconditionally. Establishing and understanding these conditions, then, becomes the major task of this research.

The findings mentioned above can be understood as follows. A network needs to strike a balance between the size of the pie and its share of the pie vis-a-vis its merchants, as a larger share distorts merchants' output choices more and hence reduces the total profit of the two parties. This is true for both fee schemes, but the distortion is smaller under proportional fees than unit fees.[3] When a network switches from unit fees to proportional fees, therefore, its profitability frontier expands, which in turn suggests the feasibility of welfare improvement. In equilibrium, the network may settle at the original share and allow its merchants to indeed expand the output, which raises welfare. Alternatively, however, the network may also do the opposite by raising its share, even to the point where the output decreases, in which case the welfare drops. That is why the welfare consequence of proportional fees is ambiguous in general.

Demand properties affect a network's tradeoff. A sub-convex demand becomes less elastic when output increases,[4] which means a small reduction in its share can increase the total profit a lot. In equilibrium, then, output always increases after all the networks simultaneously switch to proportional fees. This holds regardless of the degree of network competition or merchants' homing feasibility.

If the demand is super-convex, network competition becomes crucial for the welfare consequence. Compared with a monopoly network, output expansion on a competing network (with single-homing merchants) is more beneficial because it encroaches upon competing networks' market shares. Such a business-stealing incentive exists for both fee schemes, but it

[3]When a merchant expands its output, the transaction price will drop. This reduces its payment to the network under proportional fees but not under unit fees. In other words, proportional fees give merchants an additional incentive to expand and thus mitigate the distortions caused by merchant markup. This effect has also been pointed out by Shy and Wang (2011) and Johnson (2017).

[4]A demand is sub-convex if and only if the demand becomes less elastic when the price is lower, or equivalently the pass-through rate is less than 100%, i.e., the percentage increase of price is smaller than the percentage increase of cost (Mrázová and Neary, 2013; 2017). The opposite holds for super-convex demands. A constant elastic demand is a borderline case between sub- and super-convex demands.

is stronger under proportional fees. As a result, network competition increases the likelihood of welfare improvement. If network competition is sufficiently strong, welfare always rises.

Merchants' multi-homing further strengthens proportional fees' tendency to improve welfare on competing networks, through two additional channels. First, when a merchant expands its transaction on one network, its sale price on a competing network will drop, as the two sales are imperfect substitutes. This will reduce the merchant's payment to the second network under proportional fees but not under unit fees, and therefore gives the first network an additional incentive to lower its proportional fee when merchants multi-home. Second, merchants' multi-homing strengthens network competition. When a network lowers its fee, each of its merchants will reallocate its total transactions across networks in favor of the lower-priced network.[5] The intensified network competition happens under both fee schemes, but the stimulus to output expansion is again stronger under proportional fees. As a result, multi-homing strengthens each competing network' output expansion incentive under proportional fees.

In sum, network competition intensifies steadily when the market structure changes from monopoly network to competitive networks, and then from single-homing merchants to multi-homing merchants. A consistent message, therefore, is that intensified network competition is conducive to welfare improvement by proportional fees. The driving force can again be traced to the fundamental function of proportional fees to expand a network's profitability frontier, which allows a network to lower its fee and encourage its merchants to expand. In network competition, such feasibility is increasingly turned into equilibrium behavior, as networks compete by giving cost advantages to their own merchants in competition with merchants affiliated with other networks. Such competitive incentive strengthens when network competition intensifies.

For proportional fees' impacts on business profitability, we find that a monopoly network always benefits from a switch to proportional fees, but competing networks may be hurt. Even so, proportional fee may be a dominant strategy, leading to a Prisoners' Dilemma if networks choose the fee type independently. If network competition is weak, the supply chain (i.e., all networks and merchants combined) gains if and only if consumers and welfare also gain. Therefore, when consumers lose, merchants also lose, but networks gain. When consumers gain, networks and merchants may both gain. If, on the other hand, network competition is strong, the supply chain always loses. Networks and merchants may each gain or lose, but they can never both gain.

These findings are established in a two-sided market where each network collects fees from

[5] This is best understood when network services are homogeneous. Suppose a network charges a lower fee than other networks. If merchants single-home, other networks would still have businesses because their single-homing merchants, despite competitive disadvantages, would have no choice but to carry out some transactions on their locked-in networks. By contrast, when merchants multi-home on all networks, every merchant will move all its trasactions onto the lowest charging network; the remaining networks will have zero business. The fierce competition pushes all networks to lower their fees, in fact all the way down to their marginal cost in equilibrium.

both consumers and merchants. The two fees are assumed to be of the same type, and the resulting price neutrality (Rochet and Tirole 2002) implies that consumer fees can be set to zero without any loss of generality. As a result, the setting is equivalent to a vertical market structure, and the findings are directly applicable to retailing, licensing, and taxation. We adopt the two-sided setting not only because it is more general, but, more importantly, it helps us understand the crucial role of price neutrality, which turns out to depend on the innocuous assumption that the two sides pay the same type of fees. As shown in an extension, price neutrality breaks down if the two fees differ in type, and the comparison between unit and proportional fee regimes will be different.

The research highlights the possibility for proportional fees to reduce welfare, which happens only when the demand is super-convex. The necessary condition is also found to be the same regardless of network competition and merchant multi-homing. Whether a demand is sub- or super-convex is apparently an empirical question. Although sub-convex demands have been commonly used in theoretical analysis, the super-convex demand is attracting more and more attention. Several recent studies in industrial organization and international trade show that super-convex demands cannot be ruled out theoretically (Amir and Lambson, 2000; Chen and Riordan, 2007; Zhelobodko et al., 2012; Mrázová and Neary, 2013; 2017). Empirical studies have also identified super-convex demands where entry or economic integration leads to higher markups (Ward et al., 2002; Badinger, 2007). Therefore, what happens under super-convex demands is an important question for welfare comparison between proportional and unit fees.

Most studies of proportional fees focus on a setting of monopoly network with various additional restrictions, such as zero network cost or monopoly merchant. The general finding is that proportional fees improve social welfare if the demand is sub-convex or constant elastic (Shy and Wang, 2011; Johnson, 2017; and Llobet and Padilla, 2016). Our results are stronger because our setting is more general, which allows us to demonstrate that as long as the demand is sub-convex or constant elastic, no other conditions are needed for proportional fees to (weakly) improve welfare. Although some researches find that proportional fees may reduce social welfare if the demand is super-convex (Gaudin and White, 2014a, 2014b), they only establish super-convexity as a necessary condition. By contrast, we characterize the equilibrium comparison more completely by establishing conditions that are both necessary and sufficient, which encompass all the underlying parameters. This can be useful to guide both empirical studies and antitrust policies.[6]

Perhaps the biggest contribution of our paper is to go beyond monopoly by investigating network competition and merchant multi-homing, and to demonstrate that both conditions are conducive to welfare-improvement of proportional fees. In the context of fee scheme comparison, the literature rarely touches on network competition. The only exception is Johnson (2017). However, due to different methodologies in modeling network competition,

[6]Policy implications will be elaborated in more details in Section 6.

we arrive at quite different predictions in both welfare and profits.[7] Johnson (2017) adopts a constant conduct-parameter approach by assuming that a player's conduct is dictated by an exogenous parameter that is invariant to the business models under comparison. This setting seems to forego a major force from the outset, as different fees schemes do not alter competition intensity in either layer of the business. By contrast, we have taken a standard successive oligopoly approach, which fully endogenizes firms' behavior as well as competition intensity. An additional advantage is that our approach enables a natural modeling of multi-homing merchants, and the finding deepens our understanding of the driving force.

The paper is organized as follows. The analysis is carried out progressively in three models: monopoly network, competing networks with single-homing merchants, and competing networks with multi-homing merchants. Section 5 presents a series of extensions including Bertrand competition, two-sided markets, and mixed fees. Section 6 presents more details of antitrust cases and elaborates policy implications of our findings. Proofs are collected in the appendix in Section 7.

2 Monopoly Network

Suppose that $m < \infty$ merchants compete à la Cournot by selling a homogeneous good to consumers. All merchants have the same constant marginal cost of production, κ_M. The inverse demand for the good, $p(Q)$, is twice continuously differentiable and satisfies $p(Q) > 0$ and $p'(Q) < 0$, where Q is the total sales quantity of all the merchants. Assume that the elasticity of demand $\varepsilon \equiv -\frac{p(Q)}{p'(Q)Q} \geq 1$ so that marginal revenue is non-negative; and the concavity of demand $\rho \equiv \frac{p''(Q)Q}{p'(Q)} > -2$ so that marginal revenue decreases with output.[8]

Consumers use payment cards to buy the good, and all the cards belong to a monopoly card network, which incurs a constant marginal cost of κ_N for processing the transaction of each unit of the good. The network charges service fees to both consumers and merchants. A fee can take two forms: it can be proportional to the transaction price, hereby referred to as a proportional fee, or a fixed amount for each transacted unit, thus referred to as a unit fee. Denote the proportional coefficient as τ_C or τ_M, and the unit fee as t_C or t_M, where the subscripts indicate consumers (C) or merchants (M).

Assume that the network uses the same type of fees for both consumers and merchants (mixed fees will be analyzed later). By definition, the demand, $p(Q)$, is a consumer's per-unit gross expenditure inclusive of the fee paid to the network. Under unit fees, the network receives t_C from consumers and t_M from merchants, so its average revenue is $t_C + t_M$, and the merchants receive the remaining $p(Q) - (t_C + t_M)$. Under proportional fees, the transaction price between consumers and merchants is $\frac{p(Q)}{1+\tau_C}$. The network receives $\frac{\tau_C}{1+\tau_C}p(Q)$ from consumers and $\frac{\tau_M}{1+\tau_C}p(Q)$ from merchants, so its average revenue is $\frac{\tau_C+\tau_M}{1+\tau_C}p(Q)$, and

[7] More detailed comparison will be discussed in the text.

[8] Both ε and ρ are local properties depending on the output Q and do not have to be constant. The two properties are assumed only in the neighborhood of any relevant equilibrium.

the merchants receive the remaining $p(Q) - \frac{\tau_C + \tau_M}{1+\tau_C} p(Q) = \frac{1-\tau_M}{1+\tau_C} p(Q)$. Although the network has two instruments (t_M, t_C) or (τ_C, τ_M), only their combination (i.e., $t_C + t_M$ and $\frac{\tau_C+\tau_M}{1+\tau_C}$) matters for the equilibrium outcome and each party's profit. This is price neutrality (Rochet and Tirole, 2002). Let $t \equiv t_C + t_M$ and $\tau \equiv \frac{\tau_C+\tau_M}{1+\tau_C}$.

The game proceeds as follows. For any given fee scheme, the network chooses τ (under proportional fees) or t (under unit fees). Taking either fee as given, each merchant chooses its output in Cournot competition, and consumers purchase the good. Finally the network collects fees from merchants and consumers based on the transaction price and quantity.

2.1 Unit fees

When the network charges unit fees, the profit of merchant i is $\pi_i^U = [p(Q) - t - \kappa_M] q_i$ if it produces q_i. The merchant first-order condition (FOC) leads to:[9]

$$\frac{1}{m} p'(Q)Q + p(Q) = \kappa_M + t. \tag{1}$$

As a result, the network's average revenue is:[10]

$$\begin{aligned} D^U(Q) &= t \\ &= p(Q) - \kappa_M + p'(Q)\frac{Q}{m}. \end{aligned} \tag{2}$$

Given the average revenue, the network's profit is:

$$\begin{aligned} \Pi_N^U(Q) &= (t - \kappa_N)\, Q \\ &= [p(Q) - \kappa_M - \kappa_N]\, Q + \frac{1}{m} p'(Q) Q^2. \end{aligned} \tag{3}$$

Since choosing t is equivalent to choosing Q, we will treat Q as the network's choice variable and express its average revenue, t, as a dependent variable. Assuming $\frac{\partial^2 \Pi_N^U(Q)}{\partial Q^2} < 0$,[11] the network's FOC:

$$\kappa_M + \kappa_N = p'(Q)Q + p(Q) + \frac{1}{m}(\rho + 2)\, p'(Q)Q, \tag{4}$$

will determine a unique optimal output, Q^U, based on which other endogenous variables can be calculated.

[9] $\varepsilon \geq 1$ and $\rho > -2$ are sufficient to guarantee merchants' profit functions behave well so that the equilibirum is unique and stable.

[10] $D^U(Q)$ links the network's per-unit revenue to the total output, Q, and can therefore be regarded as its (inverse) demand function. However, since the revenue is collected from both consumers and merchants, it is unclear who is demanding the service. For this reason, we will use the term *average revenue* rather than inverse demand.

[11] This condition relates to the third derivative of the final demand function ($p'''(Q)$) and is very mild; there is a similar condition for proportional fees. Later we will distinguish between sub- and super-convex demands, which are related to the first and second derivatives of the demand function and are therefore independent of the second-order condition assumed here.

2.2 Proportional fees

When the network charges proportional fees, the profit of merchant i is $\pi_i^P = [(1-\tau)p(Q) - \kappa_M]\, q_i$. The merchant FOC results in:

$$(1-\tau)\left[\frac{1}{m}p'(Q)Q + p(Q)\right] = \kappa_M. \tag{5}$$

As a result, the network's average revenue under proportional fees is:

$$\begin{aligned} D^P(Q) &= \tau p(Q) \\ &= p(Q) - \kappa_M + \frac{\kappa_M}{p'(Q)\frac{Q}{m} + p(Q)} p'(Q)\frac{Q}{m}. \end{aligned} \tag{6}$$

The network's profit is:

$$\begin{aligned} \Pi_N^P(Q) &= \left[D^P(Q) - \kappa_N\right] Q \\ &= [p(Q) - \kappa_M - \kappa_N]\, Q + \frac{\kappa_M}{p'(Q)Q + mp(Q)} p'(Q)Q^2 \end{aligned} \tag{7}$$

Again treating Q as the choice variable, we have the network's FOC:

$$\kappa_M + \kappa_N = p'(Q)Q + p(Q) - \kappa_M \frac{m\left[(\rho+2)\,\varepsilon + 1\right] - 1}{(m\varepsilon - 1)^2}. \tag{8}$$

A unique optimal output, Q^P, can be determined from condition (8) if $\frac{\partial^2 \Pi_N^P(Q)}{\partial Q^2} < 0$, which is assumed.

2.3 The discount effect

The network's average revenue under unit fees has been given by (2). In that expression, the term $p'(Q)\frac{Q}{m} < 0$ captures merchants' markup. It is a distortion in the network's view and is caused by merchants' market power. Under proportional fees, the network's average revenue in (6) can be rewritten as

$$D^P(Q) = p(Q) - \kappa_M + (1-\tau)\, p'(Q)\frac{Q}{m}, \tag{9}$$

where merchants' markup is $(1-\tau)\, p'(Q)\frac{Q}{m}$.

Although the individual fees levied on merchants (τ_M) or consumers (τ_C) can be negative (i.e., a subsidy to one side), the composite fee τ must be positive in equilibrium, as it is the coefficient of the network's total revenue (in fact $\tau \in (0,1)$). As a result,

$$D^P(Q) - D^U(Q) = \underbrace{-\tau(Q)p'(Q)\frac{Q}{m}}_{\text{Discount Effect }(+)} > 0.$$

That is, *for any given* Q, the network receives a larger average revenue under proportional fees than under unit fees.[12] Note that we are treating Q as the choice variable and therefore τ is expressed as a function of Q.

[12] $D^P(Q) = D^U(Q)$ only when $Q = 0$ or $p = 0$ (see Figure 1), which never happens in equilibrium. For this reason, subsequent discussions ignore the possibility of $D^P(Q) = D^U(Q)$.

To understand why $D^P(Q) > D^U(Q)$, note that each merchant's output choice balances its marginal revenue (from the final demand) with its marginal cost, which includes payments to the network. Under unit fees, if a merchant sells more, a one dollar drop in the final sales price is fully shouldered by the merchant itself. Under proportional fees, however, a one-dollar drop of the price will result in only a $1 - \tau(Q) < 1$ dollar loss for the merchant. Fixing the output level, therefore, proportional fees reduce a merchant's marginal cost without affecting its marginal revenue. The saved marginal cost, which equals the proportional ratio, $\tau(Q)$, multiplied by the merchant's markup, $-p'(Q)\frac{Q}{m}$, can then be taken away by the network without reducing the merchant's output.

Therefore, proportional fees give the network a larger average revenue for any given total output or, alternatively and equivalently, a larger total output for any given average revenue (see Figure 1). We will refer to this property as *the discount effect* of proportional fees. A few features of the discount effect are worth mentioning. First, the discount effect is valid for an arbitrary Q and is therefore independent of equilibrium characterization. Second, it is not due to competition among merchants (i.e., a horizontal externality), and therefore exists even when there is only one merchant or when merchant competition is Bertrand (with differentiated products), as shown in an extension later. Third, the driving force is the existence of merchant markup (i.e., a vertical externality or distortion) and the fact that proportional fees mitigate the distortion. As a result, the discount effect arises if and only if individual merchants have market power, i.e., $-p'(Q)\frac{Q}{m} \neq 0$, which holds whenever $m \neq \infty$, as is assumed in the model.[13]

Figure 1 demonstrates the network's average revenue in the two fee schemes, $D^P(Q)$ and $D^U(Q)$. They are constructed from the consumer final demand, $p(Q)$, which is moved down first by κ_M to take account of merchants' marginal production cost, and then further down due to merchants' market power.

2.4 Consumer surplus and social welfare

Since merchants have identical cost and produce a homogeneous product, consumer surplus and social welfare are completely characterized by the equilibrium total output.

At any given output, if the network switches to proportional fee, its total profit will increase due to the discount effect. Consider how a marginal increase in the output would affect the *additional profit*, which equals the proportional share, $\tau(Q)$, times merchants' total earnings.[14] Fixing the share, a larger output increases merchant earnings (as will be explained later) and therefore tends to raise the network's additional profit. This is an *earnings effect*. Fixing the earnings, a larger output requires a larger share to merchants and hence a smaller

[13] Under perfect competition among merchants (i.e., $\lim_{m\to\infty} -p'(Q)\frac{Q}{m} = 0$), the effect disappears completely, and the two fee schemes will lead to exactly the same outcome in equilibrium.

[14] The additional total profit equals the total output multiplied by the additional average revenue (i.e., the discount effect). As mentioned earlier, the discount effect equals $\tau(Q)$ multiplied by merchants' markup, $-p'(Q)\frac{Q}{m}$, so merchants' total earning/profit is $-\frac{1}{m}p'(Q)Q^2$.

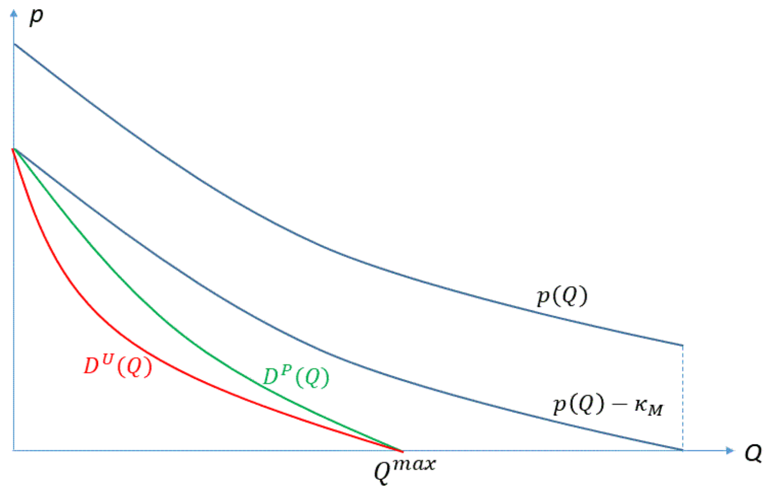

Figure 1: The network's average revenue under unit and proportional fees

share to the network. This tends to reduce the network's additional profit and is referred to as a *concession effect*. These two effects can be expressed explicitly as follows (see Appendix 7.1):

$$
\begin{aligned}
& sign\left\{\frac{\partial \Pi^P}{\partial Q}-\frac{\partial \Pi^U}{\partial Q}\right\} \qquad (10)\\
= \quad & sign\left\{\underbrace{\tau(\rho+2)}_{\text{Earnings Effect }(+)}+\underbrace{-(1-\tau)\,\tfrac{\rho+m+1}{m\varepsilon-1}}_{\text{Concession Effect }(-)}\right\},
\end{aligned}
$$

where $\frac{\partial \Pi^P}{\partial Q}-\frac{\partial \Pi^U}{\partial Q}$ is the network's additional marginal revenues due to proportional fees.

The decomposition is valid for any arbitrary output, not just the equilibrium level. Whether the equilibrium output is raised by proportional fees depends on how the additional profit changes with Q evaluated at the unit fee equilibrium Q^U. Since the earnings and concession effects always move in opposite directions, to determine the net effect, we solve κ_M from (4) and substitute it into (8). Then $Q^P > Q^U$ if and only if (at the equilibrium Q^U):

$$\kappa_N > -p(Q^U)\frac{[\varepsilon(\rho+1)+1](\rho+m+1)}{\varepsilon\{m[\varepsilon(\rho+2)+1]-1\}} \equiv \widehat{\kappa}_N. \qquad (11)$$

Notice that Q^U, ε and ρ are endogenous and therefore dependent on m, κ_N and κ_M. In addition, $p(Q^U)$, $\rho+m+1$ and $\varepsilon\{m[\varepsilon(\rho+2)+1]-1\}$ are all positive, while the sign of $\varepsilon(\rho+1)+1$ is ambiguous. According to Mrázová and Neary (2013, 2017), a demand function is sub-convex if $\varepsilon(\rho+1)+1>0$, constant-elastic if $\varepsilon(\rho+1)+1=0$, and super-convex if $\varepsilon(\rho+1)+1<0$.

If the demand is sub-convex, $\widehat{\kappa}_N<0$ and thus $Q^P>Q^U$. If the demand has constant elasticity, $\widehat{\kappa}_N=0$ and thus $Q^P\geq Q^U$, with equality if and only if $\kappa_N=0$. If the demand is super-convex, $\widehat{\kappa}_N > 0$, therefore $Q^P > Q^U$ if and only if the (11) holds. In that case, we can show (see Appendix 7.2) that $Q^P > Q^U$ is more likely when κ_M is smaller, or κ_N is larger when ρ is constant. The impact of m is more complex: $Q^P > Q^U$ is more likely for *larger* m if the demand is mildly convex, but for *smaller* m if the demand is very convex.

In sum, proportional fees improve consumer surplus and social welfare (as compared to unit fees) if and only if the total output increases (i.e., $Q^P > Q^U$). In addition,

Proposition 1 *On a monopoly network,*

(i) If the demand is sub-convex, then $Q^P > Q^U$.

(ii) If the demand is constant-elastic, then $Q^P > Q^U$ *if* $\kappa_N > 0$, *and* $Q^P = Q^U$ *if* $\kappa_N = 0$.

(iii) If the demand is super-convex, then

- *(a) if* $\kappa_N = 0$, *then* $Q^P < Q^U$;
- *(b) if* $\kappa_N > 0$ *and* $\kappa_M = 0$, *then* $Q^P > Q^U$;
- *(c) if* $\kappa_N > 0$ *and* $\kappa_M > 0$, *then* $Q^P > Q^U$ *is more likely if*
  - *(c1)* κ_M *is smaller;*
  - *(c2)* m *is smaller when* $\rho \in \left(-2, -\frac{m(3\varepsilon+2)-1}{2m\varepsilon}\right)$, *and* m *is larger when* $\rho \in \left(-\frac{m(3\varepsilon+2)-1}{2m\varepsilon}, -\frac{\varepsilon+1}{\varepsilon}\right)$;
  - *(c3)* κ_N *is larger when* ρ *is constant.*

Result (i) says that demand sub-convexity is sufficient for proportional fees to strictly improve welfare. The intuition is the following. In order to expand the total output, the network has to lower the fees, waiting for the merchants to pass (part of) the concessions to final consumers. A sub-convex demand becomes less elastic when Q increases (Mrázová and Neary, 2017), which means the final price will become increasingly sensitive to changes in the merchants' effective marginal cost.[15] Since a given increase in Q can be achieved through a smaller concession, the net effect tends to be positive.

If the demand is super-convex (result (iii)), the opposite is true: output expansion under proportional fees requires a larger concession to merchants than that under unit fees, so the net effect is ambiguous. The equilibrium comparison will then depend on the remaining three parameters: κ_M, m, and κ_N. Note that the tradeoff between the earnings and concession effects depends crucially on the endogenous $\tau(Q)$ under proportional fees. For any given Q, a larger $\tau(Q)$ means the network receives a larger share of merchants' earnings and therefore has a stronger incentive to expand.

[15] Roughly, $\frac{p-c}{p} = \frac{1}{\varepsilon}$, where c is a merchant's effective marginal cost, which includes its own cost κ_M and the fees paid to the network. Then $p = \left(1 + \frac{1}{\varepsilon-1}\right) c$. When ε is smaller, the coefficient $1 + \frac{1}{\varepsilon-1}$ becomes larger, meaning that a small drop in c can lead to a large drop in p.

The impacts of κ_M and m can be established for *any given* Q and therefore do not involve equilibrium characterization.[16] When κ_M is smaller, the network is able to collect a larger fee for any given Q (see (5)), which strengthens the earnings effect and weakens the concession effect.[17] As for m, on one hand, a larger m reduces merchants' markup and therefore allows for a larger share; on the other hand, a larger m also means a faster drop of the proportional fee is needed to increase the output, which strengthens the concession effect. These two move in opposite directions. A more convex demand indicates a larger pass-through, i.e., any change of the fees will be mostly passed through to final consumers, so the merchants' profit is less sensitive to the network's fees. This affects the relative strength of the two effects.

The impact of the network's cost, κ_N, must be evaluated in *equilibrium*. A larger κ_N reduces the equilibrium output, which in return indicates that the network obtains a larger share of additional profit (i.e., a larger $\tau(Q)$). At the same time, the smaller output raises demand elasticity (due to super-convexity), which further reduces merchants' share of the additional profit (given constant ρ). Both effects tend to strengthen the earnings effect and reduce the concession effect.[18]

Proposition 1 relates proportional fees' welfare consequence to all the underlying parameters in terms of demand (as captured by ε and ρ, which are themselves inter-related), costs (as captured by κ_N and κ_M), and the degree of merchants competition (as captured by m). Result (i) has been known and well documented by Bishop (1968), Gaudin and White (2014a), Llobet and Padilla (2014), and Johnson (2017). Our contribution is to show that the result continues to hold even with merchant competition and positive network marginal cost, whereas existing studies assume either single merchant or zero network cost. Result (ii) is what Shy and Wang (2011) got, as they used constant-elastic demands. Although Gaudin and White (2014a, 2014b) realized that super-convexity is *necessary* for proportional fees to reduce social welfare, our result (iii) goes further by providing a *sufficient and necessary* condition, which encompasses all the underlying parameters.

[16]The equilibrium output is determined from the intersection between the network's marginal cost (κ_N) and its marginal revnue ($\frac{\partial \Pi^P}{\partial Q}$ and $\frac{\partial \Pi^U}{\partial Q}$ in (10), which are for any arbitrary Q) under either fee scheme. These two marginal revenues depend on κ_M and m but not on κ_N.

[17]An alternative way to understand the role of κ_M is the following. Under proportional fees, a merchant maximizes $[(1-\tau)p(Q)-\kappa_M]q_i = (1-\tau)\left[p(Q)-\frac{\kappa_M}{1-\tau}\right]q_i$. Since the merchant's true cost is actually κ_M, revenue-sharing with the network inflates the merchant's cost and distorts (from the viewpoint of the network) the output choice. When κ_M is smaller, the distortion created by a given τ is also smaller or, equivalently, the network can implement a given Q with a larger τ without worrying too much about the distortions brought by the large τ. If $\kappa_M = 0$, for example, the cost distortion disappears no matter how large τ is. By contrast, under unit fees, a merchant maximizes $[p(Q)-t-\kappa_M]q_i$. Even when $\kappa_M = 0$, the presence of t still creates substantial cost distortion in the merchant's output choice. That is why when κ_M is smaller, proportional fees are more likely to raise output in equilibrium.

[18]Note that although κ_N and κ_M both affect the equilibrium Q in the same way (i.e., a higher cost leads to a lower output), they have opposite impacts on the comparison between Q^P and Q^U. In particular, the benign welfare effect of proportional fees is more likely for a larger network cost (κ_N), but for a smaller merchant cost (κ_M). The network's cost (κ_N) mainly determines how large the equilibrium output is, while the merchants' cost (κ_M) determines how much the merchants' market power distorts the output choice.

In the taxation literature, the general conclusion is that ad-valorem tax always raises welfare, and is a Pareto improvement if merchant is monopoly (Suits and Musgrave 1953, Bishop 1968, Delipalla and Keen 1992, Skeath and Trandel 1994, and Anderson et al. 2001).[19] Note that a government (equivalent to our network) seeks to maximize social welfare for a given target amount of tax revenue rather than maximizing the tax revenue itself. Then the conclusion is an immediate corollary to our discount effect, which is independent of equilibrium characterizations and hence the monopolist's objective. If a government, on the other hand, does seek to maximize tax revenue (the so-called "Leviathan hypothesis", see Gaudin and White 2014b), then our result (iii) highlights the possibility for proportional fees to hurt social welfare.

2.5 Network and merchant profits

For the network, proportional fees' impact is straightforward and unambiguous:

$$\begin{aligned} \Pi_N^P(Q^P) &\geq \Pi_N^P(Q^U) \quad \text{(as } Q^P \text{ is chosen optimally to maximize } \Pi_N^P(Q)) \\ &> \Pi_N^U(Q^U) \quad \text{(as } D^P(Q) > D^U(Q) \text{ due to the discount effect)} \end{aligned}$$

Therefore, a monopoly network always benefits from a switching to proportional fees. Fixing any given Q, proportional fees give the network a larger profit due to the discount effect. When Q is allowed to differ between the two schemes (i.e., in equilibrium), by revealed preference, the network must benefit even more.

Now turn to the merchants. Under unit fees, the total profit of all merchants is $\Pi_M^U(Q) \equiv (p(Q) - t(Q) - \kappa_M)\, Q$. Given the network FOC in (4), $p(Q) - t(Q) - \kappa_M = -\frac{1}{m}p'(Q)Q$ at the equilibrium Q^U, so

$$\Pi_M^U(Q^U) = -\frac{1}{m}p'(Q^U)(Q^U)^2.$$

Under proportional fees, the total profit is $\Pi_M^P(Q) \equiv [(1 - \tau(Q))\, p(Q) - \kappa_M]\, Q$. By equation (5), $(1 - \tau(Q)) \left[\frac{1}{m}p'(Q)Q + p(Q)\right] = \kappa_M$. Then at the equilibrium Q^P,

$$\Pi_M^P(Q^P) = m\pi_i^P = -\left[1 - \tau(Q^P)\right] \frac{1}{m}p'(Q^P)(Q^P)^2.$$

Fixing any Q, proportional fees must reduce merchants' profits: $\Pi_M^P(Q) \leq \Pi_M^U(Q)$. This is the mirror relation of the network's profit being higher under proportional fees, as the two parties' profits add up to the supply chain's total profit, which is invariant to the fee structure for any given Q: $\Pi(Q) = \Pi_M^P(Q) + \Pi_N^P(Q) = \Pi_M^U(Q) + \Pi_N^U(Q)$. However, proportional fee also changes the equilibrium Q. It is easy to show $\frac{\partial \Pi_M^P}{\partial Q} > 0$ and $\frac{\partial \Pi_M^U}{\partial Q} > 0$, meaning that given a fee scheme, the merchants' profit always increases with the total output.[20] Therefore,

[19] When there are multiple merchants, Skeath and Trandel (1994) show that ad-valorem tax may not Pareto-dominate specific tax.

[20] To a merchant, the fee paid to the network is a marginal cost. Total output increases only when the fee is smaller, in which case a merchant's marginal cost decreases and therefore its markup increases. A merchant's profit, which equals the markup multiplied by its output, must therefore increase with its output.

when the network shifts from unit fees to proportional fees, the merchants' profit changes through two channels. First, fixing Q^U, the network gains, implying that the merchants lose. Second, moving from the equilibrium Q^U to the equilibrium Q^P, the merchants may gain or lose depending on whether the output increases or decreases. If $Q^P < Q^U$, as shown in the left panel of Figure 2, the two effects move in the same direction, and merchants are unambiguously worse off under proportional fees. If $Q^P > Q^U$, as shown in the right panel of Figure 2, the two effects move in opposite directions, and the impacts on merchants are ambiguous.[21]

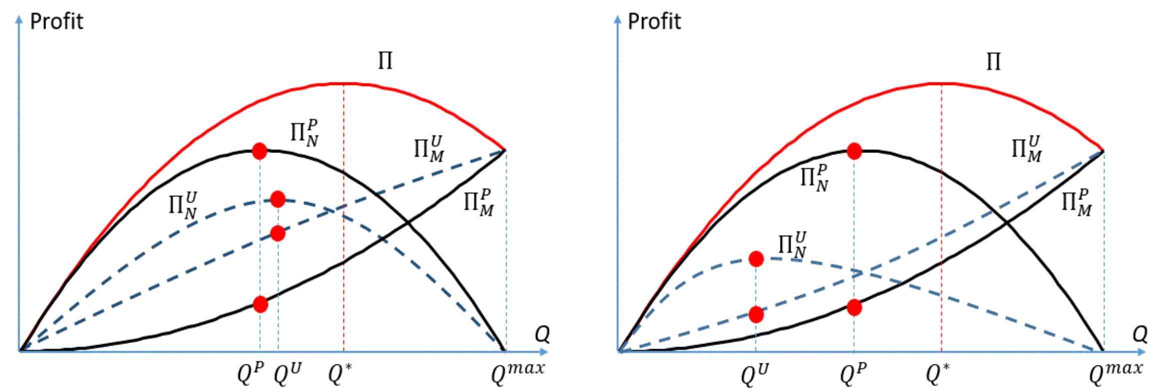

Figure 2: Profits under the two fee schemes on a monopoly network

To further evaluate the net effect, it helps to look at the joint profit of the whole supply chain (i.e., the network and all merchants): $\Pi(Q) = (p(Q) - \kappa_M - \kappa_N)\, Q$, which is maximized by Q^* such that

$$\kappa_M + \kappa_N = p'(Q^*)Q^* + p(Q^*). \tag{12}$$

Compare (12) with (4). The extra term in the right-hand side of (4) indicates how the network's incentive under unit fees differs from joint profit maximization. The term is negative, meaning that the network produces too little from the viewpoint of joint profit maximization, i.e., $Q^U < Q^*$. That is because when Q increases, the merchants will benefit, but the network ignores such externality (Gu, et al., 2019). By the same logic, $Q^P < Q^*$, as can be seen from the comparison between (12) and (8).

For any $Q < Q^*$, the joint profit of the network and merchants increases with Q, as shown in Figure 2. Given $Q^U < Q^*$ and $Q^P < Q^*$, we immediately conclude that proportional fees raise the joint profit if and only if $Q^P > Q^U$.

To summarize,

Proposition 2 *When a monopoly network switches from unit fees to proportional fees,*

[21] Johnson (2017 footnote 18) provides a numerical example in bilateral monopoly where the merchant may benefit from proportional fees if the demand is sub-convex.

(i) the network strictly gains;

(ii) the supply chain gains if and only if the equilibrium output increases. In particular,

(a) if $Q^P > Q^U$, then the supply chain gains; the merchants may gain or lose;

(b) if $Q^P = Q^U$, then the supply chain neither gains nor loses; the merchants lose;

(c) if $Q^P < Q^U$, then the supply chain lose; the merchants also lose.

For a monopoly network, result (i) was also obtained by Shy and Wang (2011), but only for constant-elastic demands. We demonstrate that it holds unconditionally for any demand (as long as merchant competition is imperfect). For merchants, Shy and Wang (2011) and Johnson (2017) find that merchants are always worse off under proportional fees if the demand is log-linear or constant-elastic,[22] which is part of our result (ii)(a). By contrast, we show that the interests of merchants and consumers do not always contradict if the demand is sub-convex.

Propositions 1 and 2 suggest that the supply chain's joint profit is perfectly synchronized with consumer surplus and social welfare. For the three parties individually, the network's incentive in adopting proportional fees is stronger than that of social welfare, while the merchants' incentive is weaker; both are biased. Welfare improvement is necessary but insufficient for merchants to gain. If merchants are better off, then consumers must also be better off; if consumers are worse off, then merchants must also be worse off. It is also possible that all three parties, i.e., the network, merchants, and consumers, are all better off under proportional fees. The following table summarizes the impacts.

Table 1: Impacts of switching from unit fees to proportional fees

	network	consumer	merchant	network+merchant	welfare
If $Q^P > Q^U$	$\uparrow$	$\uparrow$	$\uparrow$ or $\downarrow$	$\uparrow$	$\uparrow$
If $Q^P = Q^U$	$\uparrow$	$=$	$\downarrow$	$=$	$=$
If $Q^P < Q^U$	$\uparrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$

3 Network Competition: Single-Homing Merchants

We now consider network competition. This section focuses on single-homing merchants, i.e., each merchant joins only one network. Multi-homing will be analyzed in the next section. Suppose there are $n \geq 2$ networks, each affiliated with $m < \infty$ non-overlapping merchants (so the total measure of merchants is $n \times m$). Network services are heterogeneous, i.e., consumers have differential preferences for transactions processed by different networks. The inverse demand for goods sold on network j can be represented by $p_j(Q_j, \mathbf{Q}_{-j})$, where $\mathbf{Q}_{-j}$

[22] Log-linear demands have $\rho = -1$, which is a subset of sub-convex demands. Therefore, $Q^P > Q^U$ when the demand is log-linear or constant-elastic (with positive network cost in Johnson's model).

is the vector of the transaction quantities processed by all other networks.[23] For simplicity and also comparability with the monopoly case, we assume:

$$p_j(Q_j, \mathbf{Q}_{-j}) = p(Q_j + \alpha \sum_{-j} Q_{-j}),$$

where $\alpha \in [0,1]$.[24] The assumption is convenient because the definitions of ρ and ε can be directly applied. When $\alpha = 0$, the setting is equivalent to multiple independent monopoly networks; when $n = 1$, it is a single monopoly network. In both cases, it can be checked that all the expressions in this section degenerate into those in the previous monopoly network model. Other usual assumptions to ensure stable equilibrium are omitted for brevity.

Assume that all networks use the same type of fees on both consumers and merchants.[25] Under unit fees, network $j \in \{1, .., n\}$ collects $t_{j,C}$ from consumers and $t_{j,M}$ from merchants. Under proportional fees, j collects $\tau_{j,C}$ from consumers and $\tau_{j,M}$ from merchants. Figure 3 summarizes the payment flows among networks, consumers, and single-homing merchants, where p_j denotes a consumer's per-unit gross expenditure on network j, *inclusive of* the fee paid to the network. Price neutrality implies that, again, the equilibrium is completely characterized by composite fees: $t_j \equiv t_{j,C} + t_{j,M}$ for unit fees, and $\tau_j \equiv \frac{\tau_{j,C}+\tau_{j,M}}{1+\tau_{j,C}}$ for proportional fees.

The game proceeds as follows. For any given fee scheme, all networks simultaneously choose their fee levels: τ_j (under proportional fees) or t_j (under unit fees) by network j. Taking all the networks' fees as given, each merchant chooses its output in Cournot competition, and consumers make their purchases. Finally each network collects fees from parties that use its service based on the transaction price and quantity.

3.1 Unit fees

Under unit fees, for every unit sold, network j receives t_j, while merchant $i \in \{1, 2, .., m\}$ on this network receives $p_j - t_j$. Merchant i's profit is $\pi_{ij}^U = (p_j - t_j - \kappa_M)\, q_{ij}$. Its first-order condition (FOC), $\frac{\partial \pi_{ij}^U}{\partial q_{ij}} = 0$, leads to:

$$\frac{1}{m} p_j' Q_j + p_j = \kappa_M + t_j, \tag{13}$$

where $Q_j = \sum_{i=1}^{m} q_{ij}$ is the total output of all merchants on network j. The system of n FOCs, one for each network, characterizes a unique mapping from all networks' unit fees, $\mathbf{t}$, to each network's output $Q_j(\mathbf{t})$.

[23] Consumers differentiate between networks, not merchants. Products sold on the same network (by different merchants) are regarded as perfect substitutes, while products sold on different networks (by the same or different merchants) are regarded as imperfect substitutes.

[24] Although the setting here allows α to be 0, as the results are identical to that under monopoly network, subsequence discussions ignore the case of $\alpha = 0$.

[25] Therefore, we do not explicitly consider each network's individual incentive in adopting a particular fee type. That question is interesting but challenging and is beyond the scope of this paper. We nevertheless consider such individual choices in a numerical example in Appendix 7.6.

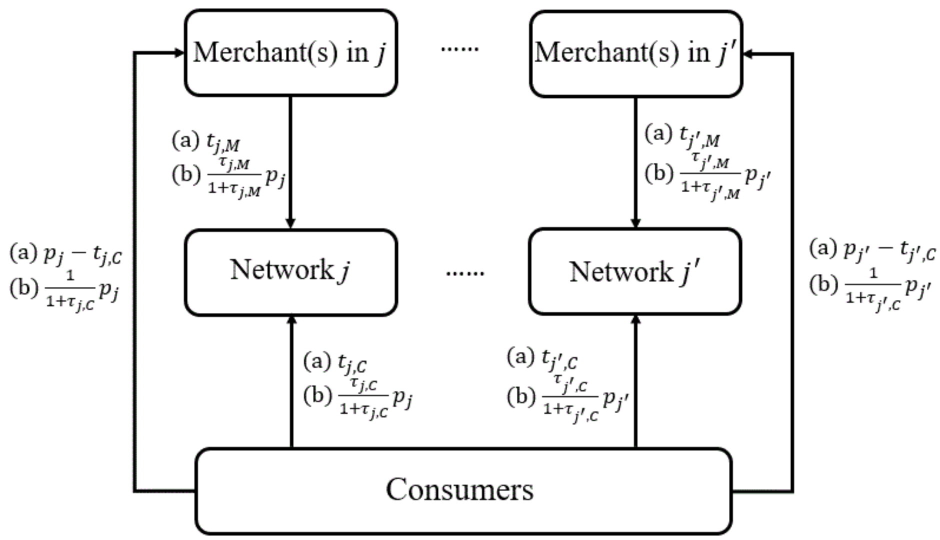

Figure 3: Payment flows when merchants single-home: (a) unit fees; (b) proportional fees

From (13), network j's average revenue is:

$$\begin{aligned} D_j^U &= t_j \\ &= p_j - \kappa_M + \frac{1}{m} p_j' Q_j. \end{aligned} \tag{14}$$

Its profit is $\Pi_j^U(t_j) = (t_j - \kappa_N) Q_j$. Network j choose t_j to maximize its profit, taking other networks' fees, t_{-j}, as given. Assuming negative semidefinite Hessian matrix,[26] network j's FOC gives rise to:

$$Q_j + \frac{\partial Q_j}{\partial t_j}(t_j - \kappa_N) = 0. \tag{15}$$

Condition (15) represents network j's best response to other networks' fees. The equilibrium is then solved from a system of similar FOCs for all networks. To characterize the equilibrium, we take derivative of (13) and other merchants' FOCs with respect to t_j to obtain expressions of $\frac{\partial Q_j}{\partial t_j}$ and $\frac{\partial Q_{-j}}{\partial t_j}$, which, together with the symmetry condition, $Q_j = Q_{-j}$ for any $j, -j \in \{1, 2, .., n\}$, can be substituted into networks' FOCs.

For convenience of comparison, we will characterize the equilibrium in terms of a virtual output, $Q^U = Q_j^U + \alpha \sum_{-j} Q_{-j}^U$, rather than the equilibrium fee t_j. Given the symmetry in equilibrium ($Q_j^U = Q_{-j}^U$), $Q^U = [1 + (n-1)\alpha] Q_j^U = k Q_j^U$, where

$$k \equiv 1 + (n-1)\alpha$$

captures the degree of network competition, with $k = 1$ for monopoly network, and k increases

[26] This property again relates to the third derivative of the demand function ($p'''(Q)$) and is very mild. It is independent of the first and second derivatives of the demand function, which are relevant for the sub- and super-convexity of the demand.

when network competition intensifies due to a larger number of networks (n is larger) or a closer substitution among network services (α is larger).

In equilibrium, Q^U must satisfy (see Appendix 7.3 for details):

$$\frac{1}{\partial Q_j/\partial t_j}\frac{Q}{k}+t-\kappa_N=0. \tag{16}$$

3.2 Proportional fees

Under proportional fees, for every unit sold, network j receives $\tau_j p_j$, while merchant i on this network receives $(1-\tau_j)\,p_j$. The merchant's profit is $\pi_{ij}^P=[(1-\tau_j)\,p_j-\kappa_M]\,q_{ij}$. Merchant FOC leads to:

$$(1-\tau_j)\left(\frac{1}{m}p_j'Q_j+p_j\right)=\kappa_M. \tag{17}$$

The system of n FOCs characterizes a unique mapping from all networks' proportional fees, $\boldsymbol{\tau}$, to each network's output $Q_j(\boldsymbol{\tau})$.

Network j's average revenue under proportional fees is:

$$\begin{aligned} D_j^P &= \tau_j p_j \\ &= p_j-\kappa_M+(1-\tau_j)\,\frac{1}{m}p_j'Q_j. \end{aligned} \tag{18}$$

It chooses τ_j to maximize its profit, $\Pi_j^P=(\tau_j p_j-\kappa_N)\,Q_j$, taking other networks' fees, τ_{-j}, as given. Again, assuming negative semidefinite Hessian matrix, j's FOC gives rise to:

$$\left\{p_j+\tau_j p_j'\left[\frac{\partial Q_j}{\partial \tau_j}+(k-1)\,\frac{\partial Q_{-j}}{\partial \tau_j}\right]\right\}Q_j+(\tau_j p_j-\kappa_N)\,\frac{\partial Q_j}{\partial \tau_j}=0. \tag{19}$$

Using similar method as in unit fees, we are able to characterize the equilibrium under proportional fees. In particular, the equilibrium virtual output is $Q^P=kQ_j^P$, where Q_j^P is network j's equilibrium output under proportional fees, and Q^P must satisfy (see Appendix 7.3):

$$\frac{\{p+\tau p'\left[\partial Q_j/\partial \tau_j+(k-1)\,\partial Q_{-j}/\partial \tau_j\right]\}}{\partial Q_j/\partial \tau_j}\frac{Q}{k}+\tau p-\kappa_N=0. \tag{20}$$

3.3 The discount effect

Given any total output by its merchants, Q_j, network j's average revenue under unit fees and proportional fees are given by (14) and (18) respectively, and therefore

$$D_j^P=D_j^U+\underbrace{\left(-\tau_j\frac{1}{m}p_j'Q_j\right)}_{\text{Discount Effect }(+)}, \tag{21}$$

where τ_j is a function of Q_j and Q_{-j}, and the argument is omitted for brevity.

For given Q_j and Q_{-j} (i.e., output of network j and all networks other than j), if network j switches from unit fees to proportional fees, its average revenue will increase. This is the same discount effect as in the monopoly network case. It measures how much more an individual network can extract from its merchants without reducing their outputs, fixing the outputs of other networks. To merchants that are affiliated with network j, other networks matter only through their respective merchants' outputs. Since these outputs are fixed, the competing networks' fee schemes are irrelevant for the discount effect that network j is experiencing.

3.4 Consumer surplus and social welfare

The welfare effect is again fully captured by the equilibrium total output. Network j's additional profit under proportional fees is:

$$\Delta\Pi_j = \Pi_j^P - \Pi_j^U = -\tau_j \frac{1}{m} p_j' Q_j^2,$$

which is a function of τ_j. Same as before, j's additional profit equals τ_j multiplied by the total earnings of its merchants: $\tau_j \frac{1}{m} p_j'(Q)[Q_j(\tau_j)]^2$. We can show (see Appendix 7.4):

$$\begin{aligned} & sign\left\{ \frac{\partial \Pi_j^P / \partial \tau_j}{\partial Q_j / \partial \tau_j} - \frac{\partial \Pi_j^U / \partial t_j}{\partial Q_j / \partial t_j} \right\} \qquad (22) \\ = \; & sign\left\{ \frac{\partial \Delta\Pi_j / \partial \tau_j}{\partial Q_j / \partial \tau_j} \right\} \\ = \; & sign\left\{ \underbrace{\tau(\rho+2)\frac{\partial Q / \partial \tau_j}{\partial Q_j / \partial \tau_j}}_{\text{Earnings Effect } (+)} + \underbrace{\frac{Q}{\partial Q_j / \partial \tau_j}}_{\text{Concession Effect } (-)} + \underbrace{2\tau\,(k-1)\left(1 - \frac{\partial Q_{-j} / \partial \tau_j}{\partial Q_j / \partial \tau_j}\right)}_{\text{Business-stealing Effect } (+)} \right\}, \end{aligned}$$

where $\tau = \tau_j$ is the symmetric proportional fee. In the expression, $\frac{\partial \Pi_j^P / \partial \tau_j}{\partial Q_j / \partial \tau_j} = 0$ and $\frac{\partial \Pi_j^U / \partial t_j}{\partial Q_j / \partial t_j} = 0$ are a network's best response functions (in term of output choices) under the two fee schemes, so $\frac{\partial \Pi_j^P / \partial \tau_j}{\partial Q_j / \partial \tau_j} - \frac{\partial \Pi_j^U / \partial t_j}{\partial Q_j / \partial t_j}$ represents the additional incentive of output expansion due to proportional fees.

Let j expands its output by lowering τ_j while fixing other networks' τ.[27] Fixing its own output, Q_j, a lower τ_j reduces the additional profit; this is the concession effect. Because Q_j is fixed, other networks do not matter, and the effect is the same as in monopoly. On the other hand, fixing τ_j, a greater Q_j affects the merchants' earnings through two channels. If competing networks' total output (Q_{-j}) is fixed, an expansion of Q_j will increase the total earnings of j's merchants and therefore j's own additional profit. This is the same earnings effect as in monopoly. If, in addition, other networks are allowed to respond, they will cut their outputs given the expansion in Q_j, which in turn further raises the total earnings of j's merchants. This is a *business-stealing effect* of proportional fees on competitive networks.[28]

[27] On a monopoly network, there is a one-to-one mapping between τ and Q, so lowering τ and raising Q is equivalent. When networks compete with their independent τ_j, we must look at the effect of lowering τ_j; the corresponding response in Q_j comes from both the change in τ_j and the responses from competing networks.

[28] Business-stealing exists on competitive networks under both fee schemes, but here we are talking about

Therefore, network competition brings an additional, business-stealing effect. The positive sign of the business-stealing effect indicates that, compared to monopoly, network competition makes it more likely for proportional fees to increase the equilibrium output. Note that (22) (for competitive networks) and (10) (for monopoly network) are both established for any given Q and therefore are comparable despite different equilibrium outputs between the two settings.

Given the decomposition in (22) for arbitrary Q, the equilibrium output comparison is determined by the sign of (22) evaluated at the equilibrium Q^U. To investigate the total effect, plug (16) into (20) to establish that $Q^P > Q^U$ if and only if:

$$\kappa_N > -\Psi \cdot \left\{ [\varepsilon(\rho+1)+1] + \frac{(k-1)(km\varepsilon-1)}{(1-\alpha)(\rho+mk)+k} \right\} \equiv \overline{\kappa}_N, \tag{23}$$

where $\Psi = \frac{p(Q^U)(\rho+mk+1)[(1-\alpha)(\rho+mk)+k]^2}{\varepsilon[(k-\alpha)(\rho+mk)+k]\{\{mk[\varepsilon(\rho+2)+1]-1\}[(1-\alpha)(\rho+mk)+1]+(k-1)(2km\varepsilon-1)(\rho+mk+1)\}} > 0$, and all variables are again evaluated at Q^U. Given $\varepsilon \geq 1$ and $\rho > 2$, the second term in $\overline{\kappa}_N$ is positive if and only if

$$\varepsilon(\rho+1)+1 > -\frac{(k-1)(km\varepsilon-1)}{(1-\alpha)(\rho+mk)+k}. \tag{24}$$

Note that if $\kappa_N = 0$, (24) is sufficient and necessary for $Q^P > Q^U$; if $\kappa_N > 0$, it is only sufficient but not necessary. The right hand side of (24) is always negative. Therefore, when the demand is sub-convex (i.e., $\varepsilon(\rho+1)+1 > 0$) or constant-elastic (i.e., $\varepsilon(\rho+1)+1 = 0$), the left hand side of (24) is non-negative, then (24) always holds, which results in $\kappa_N \geq 0 > \overline{\kappa}_N$, and proportional fees always improve social welfare.

When the demand is super-convex (i.e., $\varepsilon(\rho+1)+1 < 0$), the left hand side of (24) is also negative. It can then be shown that the right hand side is monotonically decreasing with n or α. When $n = 1$ or $\alpha = 0$ (i.e., a monopoly network), (24) cannot hold given the demand is super-convex. When $n \to \infty$ or $\alpha = 1$, it is easy to verify that (24) always holds for any super-convex demand. As a result, for a given super-convex demand and at a given output (so that ε and ρ remain unchanged), there always exists a sufficiently large n or α for (24) to hold. In that case, $\kappa_N > \overline{\kappa}_N$, and proportional fees improve welfare.

Proposition 3 *On competing networks with single-homing merchants, proportional fees improve welfare (as compared to unit fees) if and only if (23) holds. In particular,*

(i) if the demand is sub-convex or constant-elastic, then (23) always holds.

(ii) if the demand is super-convex, then (23) holds if network competition is sufficiently strong, i.e., there are many networks (n is large) or network services are sufficiently similar (α is large).

the *additional* effect due to proportional fees. Also note that the effect vanishes for monopoly network (i.e., $k = 1$, which is the case if $n = 1$ or $\alpha = 0$, and it increases with the degree of network competition (as captured by k).

Compared with monopoly network, network competition (with single-homing merchants) makes it more likely for proportional fees to improve welfare. This is because competition among networks generates a business-stealing effect under proportional fees, which encourages all networks to further expand output compared to the monopoly setting.

Johnson (2017) also finds that, under bilateral imperfect competition and sub-convex demand, the agency model (equivalent to proportional fees) always improves social welfare.[29] Our analysis goes beyond that by investigating super-convex demands and demonstrating the possibility of welfare damage. The sufficient and necessary condition thus established can facilitate further investigation of the roles of other parameters. More importantly, network competition, which plays no role in sub-convex demand (and thus in Johnson's study), turns out to be crucial for super-convex demands and, therefore, lies at the heart of our model.

So the welfare consequence depends crucially on the degree of network competition and network differentiation. Figure 4 draws the curve $\overline{\kappa}_N = 0$ in the space of $\{\varepsilon,\ \rho\}$.[30] When $\kappa_N = 0$, the curve $\overline{\kappa}_N = 0$ is the iso-welfare curve. As can be seen from the graph, the area of welfare improvements expands as α or n increases. The iso-welfare curves for a monopoly network corresponds to $\alpha = 0$ and $n = 1$, and network competition makes it more likely for proportional fees to improve social welfare. Moreover, as the number of competing networks increases or network services become closer substitutes, welfare improvement is also more likely. If $\alpha = 1$, the iso-welfare curve appears at infinite negative (fixing m and n), meaning that proportional fees always improve welfare when network services are homogeneous.

3.5 Profits

Under unit fees, the total profits of merchants and networks are, respectively:

$$\begin{aligned}
\Pi_M^U(Q^U) &= -\frac{n}{mk^2}p'(Q^U)(Q^U)^2; \\
\Pi_N^U(Q^U) &= \frac{n}{k}[p(Q^U) - \kappa_N - \kappa_M]Q^U + \frac{n}{mk^2}p'(Q^U)(Q^U)^2.
\end{aligned}$$

[29] In the agency model of retailing, a retailer sets the commission rate, and suppliers set retail prices; this is similar to proportional fees where the retailer acts like a card network and suppliers act like merchants. In the wholesale model, suppliers set the unit prices at which products or contents are sold to the retailer, who in turn sets the final retail prices. This corresponds to unit fees with a twist that the two parties' roles are swapped. Nevertheless, it can be shown that the identity of the fee-setter does not matter for the equilibrium output. Therefore, welfare in the two arrangements are directly comparable, although not the profits of the two parties.

[30] The benefit of looking at the $\{\varepsilon,\ \rho\}$ space rather than a specific parameter is that we do not need to specify the details of the equilibrium. Any point in the $\{\varepsilon,\ \rho\}$ space represents an equilibrium under particular type of demand function, and proportional fees improve welfare if and only if the point falls into the lower right corner of the corresponding iso-welfare boundary curve. For example, linear demands are characterized by $\rho = 0$, log-linear demands by $\rho = -1$, and constant elastic demands by $\varepsilon(\rho + 1) + 1 = 0$. Therefore, for a linear demand, the equilibrium is represented by the point $(0, 2)$ if the demand elasticity is 2 at that equilibrium. Notice that $\varepsilon(\rho + 1) + 1 = 0$ coincides with the monopoly iso-welfare boundary curve in both panels (given zero network cost): $\alpha = 0$ in the left panel, and $n = 1$ in the right panel.

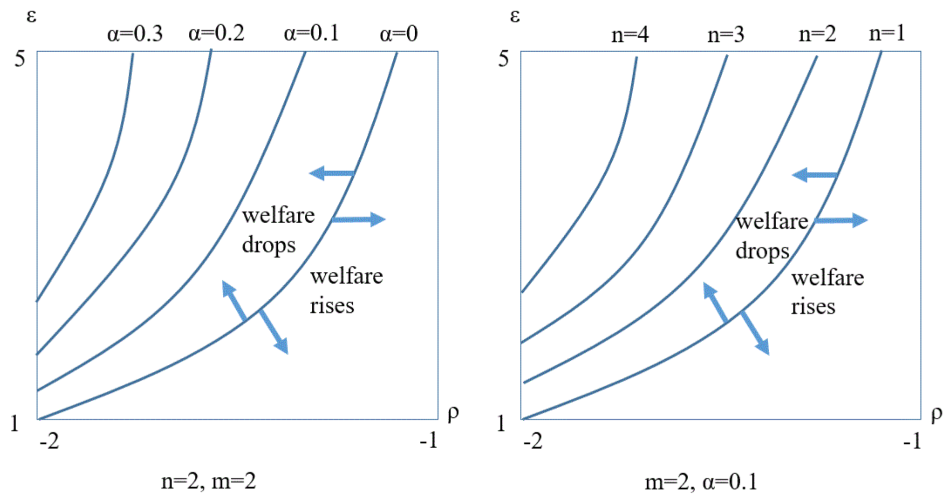

Figure 4: Impacts of network competition (with single-homing merchants): α (left panel) or n (right panel)

Under proportional fees, they are

$$\begin{aligned}\Pi_M^P(Q^P) &= -\left[1-\tau(Q^P)\right]\frac{n}{mk^2}p'(Q^P)(Q^P)^2;\\ \Pi_N^P(Q^P) &= \frac{n}{k}[p(Q^P)-\kappa_N-\kappa_M]Q^P+\left[1-\tau(Q^P)\right]\frac{n}{mk^2}p'(Q^P)(Q^P)^2.\end{aligned}$$

It can be shown that for any given fee scheme, merchants' total profit increases with Q, and networks' total profit decreases with Q for Q greater than the equilibrium level. Fixing Q, proportional fees benefit networks at the expense of merchants. However, a larger equilibrium output hurts networks and benefits merchants. In general, proportional fees' impact on either party is ambiguous.

As in the case of monopoly, here we may gain more insights by looking at the joint profit of all merchants and networks: $\Pi = \sum_{j=1}^{n}[p_j - \kappa_M - \kappa_N]Q_j$. Again, Π is maximized by Q^*, which is characterized by

$$\kappa_M + \kappa_N = p'(Q^*)Q^* + p(Q^*),$$

where $Q^* = kQ_j^*$ is also a virtual output, and Q_j^* indicates each network's quantity that maximizing the joint profit. Whether proportional fee raises the supply chain's joint profit depends on two factors: whether the supply chain over-produces in equilibrium under either free scheme (i.e., whether $Q^U > Q^*$ or $Q^P > Q^*$), and whether proportional fees raise the equilibrium output (i.e., whether $Q^P > Q^U$).

When networks are very different (i.e., α is small), their competition is weak, and the equilibrium output can be smaller than Q^* under either fee scheme. Then it is similar to the monopoly network case: Proportional fees increase the joint profit if and only if it raises the equilibrium output. In addition, when $Q^P \leq Q^U$, proportional fees benefit the networks at the expense of merchants and consumers. The left panel of Figure 5 shows such a case.[31] When $Q^P > Q^U$, proportional fees may benefit both parties; an example is provided in Appendix 7.5.

Conversely, when network services are very similar (i.e., α is large), the equilibrium outputs in both fee schemes can exceed Q^* unless the number of merchants or networks is extremely small.[32] Combined with the finding that proportional fees expand the output for large α, this implies that proportional fees push the equilibrium further away from joint profit maximization, so the supply chain's joint profit must drop. The impacts on either party are ambiguous. The right panel of Figure 5 illustrates an example of profit comparison with $Q^U > Q^*$.

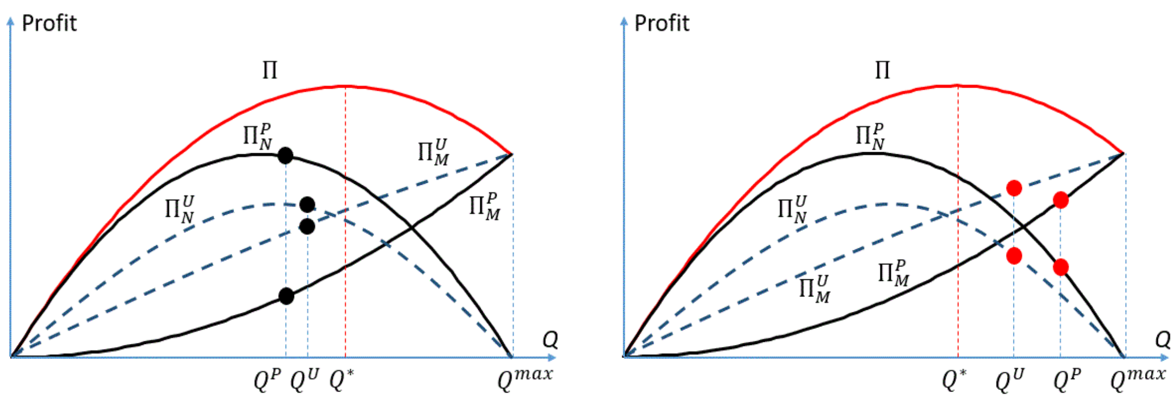

Figure 5: Network competition with single-homing merchants: Equilibrium profit comparison for small α (left panel) and large α (right panel)

Proposition 4 *On competing networks with single-homing merchants,*

(i) When networks are sufficiently differentiated (i.e. α is small), proportional fees raise networks' and merchants' joint profit if and only if equilibrium output increases. Furthermore, if $Q^P \leq Q^U$, then networks gain, but merchants lose; if $Q^P > Q^U$, then the two parties may both gain.

(ii) When networks are sufficiently similar (i.e. α is large), proportional fees reduce the supply chain's joint profit (unless the numbers of networks and merchant are extremely small). Networks and merchants may each gain or lose, but cannot both gain.

[31] Unlike in monopoly where Q^U maximizes Π_N^U and Q^P maximizes Π_N^P, here the equilibrium Q^U is greater than the level that would maximize Π_N^U due to network competition. Ditto for Q^P.

[32] When $\alpha = 1$, the two outputs are smaller than Q^* only if $m = 1$ and $n = 2$, where merchant monopoly generates strong double markup. In that case, merchants under-produce under unit fees, and the output expansion brought by proportional fees may mitigate the under-production and therefore increase the joint profit. Appendix 7.6 shows such an example.

Result (ii) implies that when network services are close substitutes, networks and merchants will have conflicting interest in adopting proportional fees: whenever one party is better off, the other must be worse off. Appendix 7.5 investigates the effect of all parameters with constant-elastic demand. Simulations show that merchants prefer proportional fees when α, ε, m, or n is large, or κ_N is small, while the networks prefer proportional fees under the opposite conditions. When α is not so large, it is possible for all three parties, i.e., networks, merchants, and consumers, to benefit from proportional fees simultaneously.

Result (ii) highlights the possibility for proportional fees to hurt networks and benefit merchants. Appendix 7.5 shows several such examples with constant-elastic demand. Such a result contrasts with Johnson (2017), who shows that revenue-sharing (equivalent to proportional fees) always benefits first-movers (equivalent to our networks) but hurts second-movers (equivalent to merchants) when the demand is either log-linear (a subset of sub-convex demand) or constant-elastic. As explained in Introduction, the different predictions are driven by different methodologies in modeling network competition.

For a monopoly network, the unconditional higher profit directly implies the incentive to adopt proportional fees. When there are competing networks, they may be collectively worse off, as argued above. Even so, each network may still have an individual incentive to adopt proportional fees. Appendix 7.6 shows an examples in which proportional fee is each network's dominant strategy even though they may be collectively worse off. In other words, networks may face a Prisoners' Dilemma when choosing between the two fee schemes.

4 Network Competition: Multi-Homing Merchants

Now consider network competition with multi-homing merchants. In particular, there are $n \geq 2$ networks and $m < \infty$ merchants, and every merchant joins all networks. All the remaining setting of the game is the same as in single-homing merchants. The usual assumptions to ensure stable equilibrium are omitted for brevity. Figure 6 summarizes the payment flows among networks, merchants and consumers under multi-homing merchants.

Note that to make multi-homing comparable with single-homing, in both settings the number of merchants that transact on each network is set at the same m. That is why the total number of merchants is $n \times m$ in single-homing, but m in multi-homing.[33] Since the equilibrium output is determined by networks' choices, we want a setting in which the interface between a network and its affiliated merchants remains unchanged, and the only difference is that a multi-homing merchant cares about how its output choice on one network affects its profits from other networks, whereas a single-homing merchant does not have

[33] If we have $m \times n$ merchants in the multi-homing setting, all the qualitative results continue to hold, and the increased number of merchants only makes welfare improvement even more likely, as a larger m makes condition (30) more likely to hold. Since our finding is that multi-homing is conducive to proportional fees' favorable welfare outcome, the current setting of m merchants instead of $n \times m$ is an assumption against the finding.

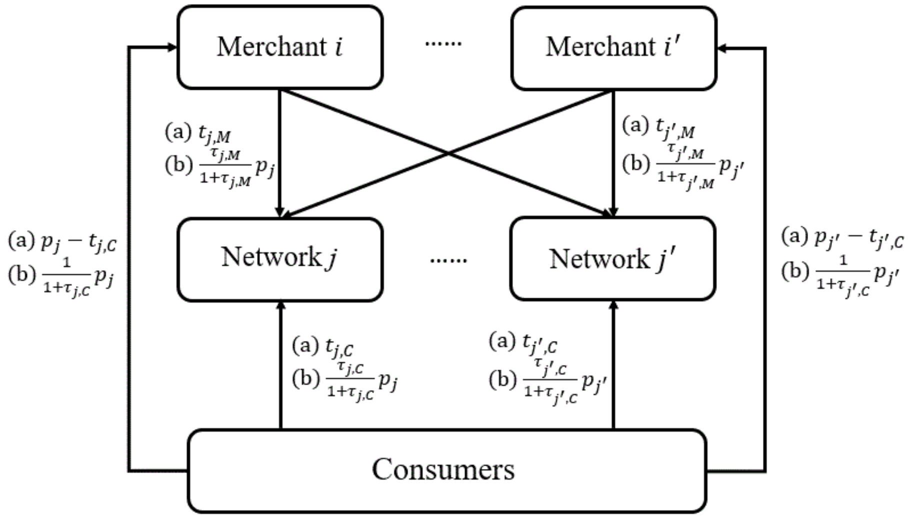

Figure 6: Payments flows when merchants multi-home: (a) unit fees; (b) proportional fees

such concerns. One way to connect the two settings is to imagine that starting from single-homing, every n merchants, one from each network, merge into a single merchant. Since network services are differentiated, a merged merchant's businesses from different networks are not homogeneous, therefore it will keep all n businesses.[34]

4.1 Unit fees

Merchant $i \in \{1, 2, \cdots, m\}$ sells on all n networks, so its profit under unit fees is $\pi_i^U = \sum_{j=1}^n [p_j - t_j - \kappa_M] q_{ij}$. Merchant FOC, $\frac{\partial \pi_i^U}{\partial q_{ij}} = 0$, determines the quantity sold on network $j \in \{1, 2, \cdots, n\}$. Summing up the FOCs over i, we have the average revenue for network j:

$$D_j^U = p_j + \frac{1}{m}\frac{\partial p_j}{\partial Q_j} Q_j + \frac{1}{m}\sum_{-j} \frac{\partial p_{-j}}{\partial Q_j} Q_{-j} - \kappa_M, \tag{25}$$

where $Q_j = \sum_{i=1}^m q_{ij}$ is the total output sold on network j, and p_{-j} is the price on a network other than j. The system of merchant FOCs (there are n of them similar to (25)) will collectively determine $Q_j\left(t_j, \mathbf{t}_{-j}\right)$, $j \in \{1, .., n\}$, where $\mathbf{t}_{-j}$ is the vector of other networks' unit fees.

Network j's profit is $\Pi_j^U\left(t_j\right) = (t_j - \kappa_N) Q_j\left(t_j, \mathbf{t}_{-j}\right)$. Then network FOCs for all the n networks will collectively determine a unique equilibrium in terms of the n unit fees. In equilibrium, we have $Q_j^U = Q_{-j}^U$, and the virtual output, $Q^U = kQ_j^U$. Details of the equilibrium characterization are provided in Appendix 7.7.

[34]In the limiting case of $\alpha = 1$, network services are homogeneous. Every merchant will carry out all transactions on the network with the lowest fee. In equilibrium, all networks charge the same fee level (in fact their marginal cost), and therefore every merchant continues to split its transactions equally among all networks.

4.2 Proportional fees

Under proportional fees, the profit of merchant i is $\pi_i^P = \sum_{j=1}^{n} \left[(1-\tau_j)\, p_j - \kappa_M\right] q_{ij}$. Summing up merchant FOCs over i to obtain network j's average revenue:

$$D_j^P = p_j + \frac{1}{m}(1-\tau_j)\frac{\partial p_j}{\partial Q_j} Q_j + \frac{1}{m}\sum_{-j}(1-\tau_{-j})\frac{\partial p_{-j}}{\partial Q_j} Q_{-j} - \kappa_M. \tag{26}$$

A system of n such equations will collectively determine $Q_j(\tau_j, \boldsymbol{\tau}_{-j})$, $j \in \{1, .., n\}$, where $\boldsymbol{\tau}_{-j}$ is the vector of other networks' proportional fees. Network j's profit is $\Pi_j^P(\tau_j) = (\tau_j p_j - \kappa_N) Q_j(\tau_j, \boldsymbol{\tau}_{-j})$. Then network FOCs from all n networks will collectively determine a unique equilibrium. Again, we focus on the condition that characterizes the equilibrium virtual output, $Q^P = Q_j^P + \alpha \sum_{-j} Q_{-j}^P = kQ_j^P$ as equilibrium outputs of every networks are identical. Details can again be found in Appendix 7.7.

4.3 The discount effect

On competing networks with multi-homing merchants, the discount effect is:

$$D_j^P = D_j^U + \underbrace{\underbrace{\left[-\frac{1}{m}\tau_j \frac{\partial p_j}{\partial Q_j} Q_j\right]}_{\text{Direct Effect } (+)} + \underbrace{\left[-\frac{1}{m}\sum_{-j}\tau_{-j}\frac{\partial p_{-j}}{\partial Q_j} Q_{-j}\right]}_{\text{Indirect Effect } (+)}}_{\text{Discount Effect } (+)}, \tag{27}$$

where τ_j is network j's proportional rate. In addition to a direct discount effect, there is an indirect effect, which captures the additional impacts due to multi-homing.

As before, here we are fixing the outputs of network j and all other networks, and look at how j's average revenue changes when all networks simultaneously switch from unit fees to proportional fees. Let us consider a particular merchant, i (in fact all merchants transact on j). When network j switches to proportional fees, merchant i tends to sell more through network j. This is the (direct) discount effect established earlier in the single-homing setting where i sells only on j. With multi-homing, i also sells its product on another network, say j'. Because p_j tends to drop due to the direct discount effect, and because i's products on j and j' are imperfect substitutes, $p_{j'}$ also tends to drop even though $Q_{j'}$ is fixed. Under proportional fees (but not under unit fees), the lower $p_{j'}$ on network j' implies that i will pay a lower fee to j', which is an additional benefit to i. As a result, multi-homing by merchant i allows network j to take away an even larger amount of fees (on top of those feasible with single-homing) from i without reducing i's output.

4.4 Consumer surplus and social welfare

Network j's additional profit under proportional fees can be written as:

$$\Delta\Pi_j = \Pi_j^P - \Pi_j^U = -\frac{1}{m}\left[\tau_j p_j' Q_j^2 + \alpha\sum_{-j}\tau_{-j} p_{-j}' Q_{-j} Q_j\right],$$

which is a function of τ_j. We can show (see Appendix 7.8):

$$
\begin{aligned}
& sign\left\{\frac{\partial \Pi_j^P/\partial \tau_j}{\partial Q_j/\partial \tau_j} - \frac{\partial \Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \qquad (28) \\
= \;& sign\left\{\frac{\partial \Delta \Pi_j/\partial \tau_j}{\partial Q_j/\partial \tau_j} + (n-1)\frac{\partial \Pi_j^U}{\partial Q_{-j}}\left(\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j} - \frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j}\right)\right\} \\
= \;& sign\left\{\begin{array}{c}
\underbrace{\frac{Q}{\partial Q_j/\partial \tau_j}}_{\text{Concession Effect } (-)} + \underbrace{\tau(\rho+2)\frac{\partial Q/\partial \tau_j}{\partial Q_j/\partial \tau_j}}_{\text{(Direct) Earnings Effect } (+)} + \underbrace{2\tau(k-1)\left(1-\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j}\right)}_{\text{Business-Stealing Effect } (+)} \\
\underbrace{+\tau(k-1)\left[(\rho+2)\left(\alpha+(k-\alpha)\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j}\right)+(k-2\alpha)\left(1-\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j}\right)\right]}_{\text{(Indirect) Earnings Effect } (+)} \\
+\underbrace{\left\{-(k-1)\left[(k+1-\alpha)\rho+(m+1)k\right]\left(\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j} - \frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j}\right)\right\}}_{\text{Allocation Effect } (+)}
\end{array}\right\},
\end{aligned}
$$

where $\tau = \tau_j$ is the symmetric proportional fee ratio. As before, $\frac{\partial \Pi_j^P/\partial \tau_j}{\partial Q_j/\partial \tau_j} - \frac{\partial \Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}$ captures each network's additional incentive to expand output due to proportional fees, again for any given Q.

As can be seen from the decomposition, when merchants multi-home, proportional fees affect the equilibrium output through five forces. The first three, namely the (direct) earnings effect, concession effect, and business-stealing effect, are the same as when merchants single-home. Multi-homing gives rise to two additional effects. Recall the indirect discount effect. When a network lowers the fee to its merchants, these merchants will reap some benefits from their products sold on other networks, and these benefits can then be turned into additional profits for the network that is lowering the fee. Referred to as an *indirect earnings effect*, this consideration gives every network an additional incentive to expand.

The last effect comes from multi-homing merchants' incentive and feasibility to allocate its transactions across networks. When merchants single-home, the intensity of network interaction is the same between the two fee schemes, i.e., $\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j} = \frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j}$. This is because networks do not compete directly; they compete indirectly through merchants' output choices. A network's output responds to changes in another network's output in the same way regardless of the source of the change, i.e., whether it is caused by a change in unit fees or proportional fees on the source network. When merchants multi-home, however, a network responds more strongly to another network under proportional fees than under unit fees, i.e., $\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j} < \frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j} < 0$. This is because networks compete on two fronts. In addition to the indirect, quantity competition through merchants' output levels, there is also a direct, price competition between networks given that each merchant also reallocates its total transactions in favor of a network that charges a lower fee.[35] This latter effect is stronger

[35] When a network lowers it fee level, the total output on this network expands, to which other networks'

under proportional fees than unit fees. Therefore, proportional fees encourage each network to further lower its fee level in order to expand. This is referred to an *allocation effect*.

For the net effect, we can show that $Q^P > Q^U$ if and only if (evaluated at Q^U):

$$\begin{aligned} \kappa_N &> -\Phi \times \{[\varepsilon(\rho+1)+1] + \frac{\varepsilon(k-1)(\rho+m+1)(k-\alpha+1)}{(1-\alpha)[(1-\alpha)\rho+m+1]}\} \\ &\equiv \widetilde{\kappa}_N, \end{aligned} \tag{29}$$

where $\Phi = \frac{p(Q^U)(1-\alpha)^2(\rho+m+1)[(1-\alpha)\rho+m+1][(1-\alpha)\rho+k(m+1)]}{\varepsilon\{(\rho+m+1)k(k-\alpha)-\alpha(1-\alpha)\rho\}\{(1-\alpha)[(1-\alpha)\rho+m+1][m\varepsilon(\rho+2)+m-1]+(k-1)[km\varepsilon+(1-\alpha)(2m\varepsilon-1)](\rho+m+1)\}} > 0$. When $\alpha = 0$ or $n = 1$, the setting degenerates into monopoly network, in which case (29) degenerates into (11).

The second term in $\widetilde{\kappa}_N$ is positive if and only if

$$\varepsilon(\rho+1)+1 > -\frac{\varepsilon(k-1)(k+1-\alpha)(\rho+m+1)}{(1-\alpha)[(1-\alpha)\rho+m+1]}. \tag{30}$$

Again, if $\kappa_N = 0$, condition (30) is sufficient and necessary for $Q^P > Q^U$; if $\kappa_N > 0$, it is sufficient but not necessary. Similar to the single-homing case, here when the demand is sub-convex or constant-elastic, $\kappa_N \geq 0 > \widetilde{\kappa}_N$, and proportional fees always improve social welfare. When the demand is super-convex, the result is also qualitatively similar: welfare improves if network competition is sufficiently strong.

The difference between (29) and (23) captures the effects of multi-homing. Although the endogenous variables (Q, ε and ρ) complicate the comparison in the general setting, we can show that multi-homing is more likely than single-homing to generate welfare improvement through proportional fees if networks have zero cost (i.e. $\kappa_N = 0$). Figure 7 shows such examples in the space of $\{\varepsilon, \rho\}$ with $\kappa_N = 0$.

Proposition 5 *On competing networks,*

(i) when merchants multi-home, proportional fees' welfare consequences are qualitatively similar to those when merchants single-home (i.e., as stated in Proposition 3).

(ii) When networks have zero cost (i.e. $\kappa_N = 0$*), multi-homing makes welfare improvement more likely as compared with single-homing.*

Fixing a fee scheme, moving from single-homing to multi-homing may raise or lower the equilibrium output.[36] Our focus here, however, is how multi-homing changes the *output*

outputs will respond; this is the same as in single-homing. In addition, each merchant will reallocate its transaction toward the focal network because of its lower fee relative to other networks.

[36] As explained earlier, multi-homing can be regarded as a merger among single-homing merchants, and a merged merchant naturally internalizes the negative externalities between its outputs on different networks. Fixing the network fee levels, such a coordination effect tends to reduce equilibrium output. However, each merchant will also reallocate its transactions in favor of the network that charges the lowest fee. Such an allocation effect (as established earlier) will intensify network competition, resulting in lower network fees and hence tends to raise the equilibrium output. Either effect may dominate (so the overall competition may be strengthened or weakened by multi-homing), in both unit fees and proportional fees.

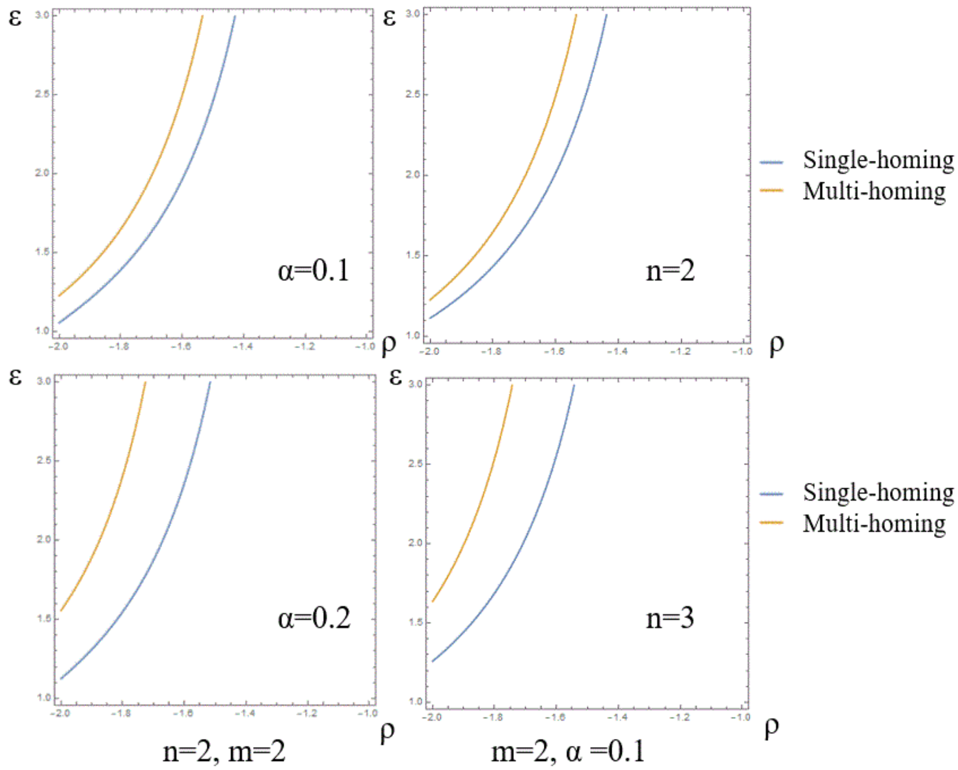

Figure 7: Single-homing versus multi-homing

comparison between unit and proportional fees, and this additional incentive is best captured by the two new effects in the decomposition in (28). Although the equilibrium Q^U will differ between single- and multi-homing, (22) and (28) are still comparable because both are for an arbitrary Q, not just the equilibrium Q. We have explained that the two new effects under multi-homing, namely the indirect earnings effect and the allocation effect, both encourage each network to further expand its output. That is why multi-homing is more conducive to proportional fees' benign welfare consequence than single-homing.

A consistent message arising from the analysis so far is that intensified network competition is conducive to welfare improvement by proportional fees. Fixing single-homing, network competition is intensified if networks are closer substitutes (i.e., α is larger) or grow in number (i.e., n is larger). Moving from single-homing to multi-homing, network competition is also intensified because a network competes not only by expanding its affiliated merchants' output level, but also by inducing each multi-homed merchant to re-allocate its total transactions across networks.

4.5 Profits

For any given Q, the joint profit of all merchants and networks is the same as in single-homing, but each party's profit will change. Under unit fees, the total profits of all merchants and networks are, respectively:

$$\begin{aligned} \Pi_M^U(Q^U) &= -\frac{n}{mk}p'(Q^U)(Q^U)^2; \\ \Pi_N^U(Q^U) &= [p(Q^U) - \kappa_N - \kappa_M]Q^U + \frac{n}{mk}p'(Q^U)(Q^U)^2. \end{aligned}$$

Under proportional fees, they are

$$\begin{aligned} \Pi_M^P(Q^P) &= -\left[1 - \tau(Q^P)\right]\frac{n}{mk}p'(Q^P)(Q^P)^2; \\ \Pi_N^P(Q^P) &= [p(Q^P) - \kappa_N - \kappa_M]Q^P + \left[1 - \tau(Q^P)\right]\frac{n}{mk}p'(Q^P)(Q^P)^2. \end{aligned}$$

Compared to single-homing, for any given Q, merchants' total profit is larger under multi-homing due to their ability to allocate transactions across networks, which in turn reduces networks' profit, as the two parties' joint profit remains unchanged.

Despite the differences, the profit functions have the same properties as in single-homing: (1) $\Pi_M^P(Q)$ and $\Pi_M^U(Q)$ increase with Q, while $\Pi_N^P(Q)$ and $\Pi_N^U(Q)$ are concave in Q; (2) $\Pi_M^P(Q) < \Pi_M^U(Q)$ and $\Pi_N^P(Q) > \Pi_N^U(Q)$; (3) profit comparison pins down to how proportional fees change the equilibrium output and how much the networks are differentiated. As a result, proportional fees' impacts on these two parties' profits are again qualitatively similar between multi-homing and single-homing. Appendix 7.9 gives examples of several scenarios.

A special case is $\alpha = 1$, i.e., network services are perfect substitutes. If merchants single-home, consumers gain, while the supply chain loses (unless the number of merchants and networks are both extremely small). By contrast, if merchants multi-home, every merchant will concentrate all its transactions on the network with the lowest fee. Under either fee scheme, such Bertrand competition drives the equilibrium fees to networks' marginal cost. Networks earn zero profits, but the discount effect still exists. As a result, proportional fees raise the equilibrium output and hence hurt merchants (because merchants already over-produce under unit fees, and the expanded output under proportional fees moves the equilibrium output further away from the joint profit maximizing level). In fact, it can be easily verified that (30) always holds when $\alpha = 1$.

Proposition 6 *On competing networks,*

(i) when merchants multi-home, proportional fees' impacts on networks and merchants' profits are qualitatively similar to those when they single-home (i.e., as stated in Proposition 4).

(ii) When merchants multi-home and networks are perfectly substitutable ($\alpha = 1$), consumers gain from proportional fees; merchants lose; networks are indifferent (their profits are zero in both fee schemes).

5 Extensions

The discount effect plays a central role in our analysis because all the other effects are built upon it. In what follows, we will further demonstrate the robustness of discount effect by extending the model to Bertrand competition, and two-sided markets with costly entry. In addition, we will investigate the situation in which merchants and consumers are levied different types of fees, where price neutrality breaks down and the advantage of proportional fees disappears. For simplicity, we will focus on a monopoly network, but it is clear that the analysis and conclusion can be easily extended to networks competition.[37]

5.1 Bertrand competition

In the main models, we have assumed that merchants compete in output quantity. Now consider the case in which merchants carry out Bertrand competition with differentiated products. The setting is the same as in Section 2 except that all merchants face symmetric final demand $Q_i(p_i, \mathbf{p}_{-i})$ with $\frac{\partial Q_i}{\partial p_i} < 0$. Unique equilibrium is guaranteed by usual assumptions about Bertrand competition, which are omitted for brevity. Again, price neutrality is valid, so we denote $t = t_C + t_M$ and $\tau = \frac{\tau_C + \tau_M}{1+\tau_C}$.

Under unit fees, the profit of merchant i is $\pi_i^U = [p_i - t - \kappa_M]Q_i(p_i, \mathbf{p}_{-i})$. The merchant FOC is:

$$p_i - t - \kappa_M + \frac{Q_i}{\frac{\partial Q_i}{\partial p_i}} = 0,$$

which in turn generates the network's average revenue from merchant i and its customers:

$$t = p_i - \kappa_M + \frac{Q_i}{\frac{\partial Q_i}{\partial p_i}}.$$

Under proportional fees, the profit of merchant i is $\pi_i^P = [(1-\tau)p_i - \kappa_M] Q_i(p_i, \mathbf{p}_{-i})$. The merchant FOC is:

$$(1-\tau)\left[Q_i + p_i \frac{\partial Q_i}{\partial p_i}\right] = \kappa_M \frac{\partial Q_i}{\partial p_i}.$$

As a result, the average revenue is:

$$\tau p_i = p_i - \kappa_M + \frac{Q_i \kappa_M}{Q_i + p_i \frac{\partial Q_i}{\partial p_i}}.$$

[37]If consumers or merchants have outside options such as non-card transaction, the qualitative results continue to hold. This is because proportional fees expand a network's profitability frontier, making it always *feasible* to increase both consumer surplus and the supply chain's joint profit. If some parties have to earn some minimum payoffs, the feasibility still exists. The presence of outside options will therefore only change the surplus allocation among various parties (and possibly the equilibrium output), but not the profitability of proportional fees for a monopoly network. In fact, social welfare is more likely to rise.

If merchants' costs differ, the equilibrium total output will remain the same as long as their average marginal cost does not change. In that case, proportional fees will generate an additional welfare gain by encouraging efficient merchants to produce more, and less efficient merchants to produce less. This can be seen from $q_j^P - q_i^P = -\frac{\kappa_{M,i} - \kappa_{M,j}}{(1-\tau)p'(Q)} > -\frac{\kappa_{M,i} - \kappa_{M,j}}{p'(Q)} = q_j^U - q_i^U$ for $\kappa_{M,i} > \kappa_{M,j}$.

The discount effect is

$$\tau p_i - t = \underbrace{-\tau \frac{Q_i}{\frac{\partial Q_i}{\partial p_i}}}_{\text{Discount Effect } (+)}.$$

Therefore, discount effect still shows up in Bertrand competition as long as merchants have market power, i.e., $\frac{\partial Q_i}{\partial p_i} < \infty$.

5.2 Two-sided markets

The payment card industry is commonly regarded as a two-sided market. There are several definitions in the literature about what constitutes a two-sided market (Weyl, 2010). According to Rochet and Tirole (2003, 2006), the key characterization of a two-sided market is price non-neutrality. They also argue that price neutrality holds for payment cards as long as three mild conditions are satisfied, so the industry should be regarded as a one-sided market despite apparent two-sidedness (Rochet and Tirole, 2006, p.648). Indeed, price neutrality holds in our model. That is why the setting can be interpreted (or framed) as an vertical market structure.

Another definition for two-sided market is the existence of cross network externality, i.e., the utility derived by users on one side increases with the network size of the other side (Armstrong 2006, and Rysman 2009, among others). To investigate how this might affect our results, we generate cross network externalities endogenously through merchant free entry. Again we focus on monopoly network. There are many identical merchants, each of which must pay an entry cost, $H > 0$, in order to operate. There are n_C identical consumers, each of whom has the demand $p(Q)$. For any fee scheme, after the network announces its particular choice of fee level, all potential merchants simultaneously decide whether or not join the network (by paying H). Therefore, the number of merchants on the network, m, is endogenized.

Under unit fees, given m, the merchant FOC is identical to that in Section 2, (1). In equilibrium, $q_i = \frac{Q}{m}$, and merchant i's profit is

$$\pi_i^U = -n_C p'(Q) \left(\frac{Q}{m}\right)^2 - H.$$

The free entry condition requires $\pi_i^U = 0$, which means that the endogenous number of merchants, $m^U(Q)$, must satisfy:

$$-n_C p'(Q) \left(\frac{Q}{m^U(Q)}\right)^2 = H.$$

Here m^U is treated as a function of Q. The network's average revenue from every individual consumer is

$$D^U(Q) = p(Q) - \kappa_M - \frac{m^U(Q)}{Q} H.$$

Under proportional fees, the same procedure is used. Merchant i's profit is

$$\pi_i^P = -n_C (1-\tau) p'(Q) \left(\frac{Q}{m}\right)^2 - H.$$

The endogenous number of merchants, $m^P(Q)$, must therefore satisfy:

$$-n_C (1-\tau) p'(Q) \left(\frac{Q}{m^P(Q)}\right)^2 = H.$$

The network's average revenue from every individual consumer is

$$D^P(Q) = p(Q) - \kappa_M - \frac{m^P(Q)}{Q} H$$

From the expressions of π_i^U and π_i^P, it is clear that the more consumers, the larger is each merchant's profit, and thus a larger number of merchants. That is, consumers have positive cross network externality to merchants. On the other hand, the larger number of merchants, the larger is equilibrium output. Since consumers surplus, $S = \int_0^Q p(x)dx - p(Q)Q$, increases with the level of output ($\frac{\partial S}{\partial Q} = -p'(Q)Q > 0$), the more merchants, the larger is consumer surplus. Therefore, merchants also exert positive cross-side network externality to consumers.

For any given Q, since $-n_C [1-\tau(Q)] p'(Q) \left(\frac{Q}{m^P}\right)^2 = -n_C p'(Q) \left(\frac{Q}{m^U}\right)^2 = H$, we have

$$m^P = \sqrt{1-\tau(Q)} m^U < m^U.$$

Then

$$\begin{aligned}
D^P(Q) - D^U(Q) &= \underbrace{\frac{H}{Q}\left(m^U - m^P\right)}_{\text{Net effect } (+)} \qquad (>0 \text{ because } m^U > m^P) \\
&= [1-\tau(Q)]p'(Q)\frac{Q}{m^P} - p'(Q)\frac{Q}{m^U} \\
&= \underbrace{-\tau(Q)p'(Q)\frac{Q}{m^P}}_{\text{Discount Effect } (+)} + \underbrace{p'(Q)Q\left(\frac{1}{m^P} - \frac{1}{m^U}\right)}_{\text{Entry/Exit effect } (-)}.
\end{aligned}$$

Therefore, merchant free entry would reduce the advantage of proportional fees (for any given Q), but will never completely eliminate it. If the number of merchants, m, is the same between the two fee schemes, $D^P(Q) > D^U(Q)$ for any given Q, which is the discount effect established in the main models. When m is endogenous, given Q, proportional fees will reduce the merchants' markup, which leads to a smaller m. The reduced competition would partially restore merchant markup. This is the entry/exit effect. As established above, the discount effect always dominates the entry/exit effect and, hence, $D^P(Q, m^P(Q)) > D^U(Q, m^U(Q))$. An immediate corollary is that, even with free-entry merchants, proportional fees continue to benefit a monopoly network.

5.3 Mixed fees

In the main analysis a network is assumed to charge the same type of fees to both merchants and consumers. This leads to price neutrality: although the network seems to have two instruments, what really matters is a composite fee, $t_M + t_C$ in the case of unit fee, and $\frac{\tau_M+\tau_C}{1+\tau_C}$ in the case of proportional fees. What will happen if the two sides of the market are charged different types of fees (hereby referred to as mixed fees)?

Suppose a monopoly network charges a unit fee, t_M, to merchants, and a proportional fee, τ_C, to consumers. Merchant FOC is:

$$\frac{1}{m}p'(Q)Q + p(Q) = (\kappa_M + t_M)(1 + \tau_C).$$

For any given Q, then, there is an inverse relationship between t_M and τ_C. The network's profit is

$$[p(Q) - \kappa_M - \kappa_N]\,Q + \frac{1}{1+\tau_C}\frac{1}{m}p'(Q)Q^2.$$

After substituting one of the fees, the other fee remains an independent choice for given Q. In other words, price neutrality no longer holds. It can be shown that the network's profit increases with the proportional fee, or equivalently decreases with the unit fee. The optimal choice is $t_M \to -\kappa_M$ and $\tau_C \to +\infty$. Notice that we do not require $\tau_C \leq 1$. If the infinite proportional fee looks a little strange, we only need to point out that the true proportion is $\frac{\tau_C}{1+\tau_C}$, which approaches 1, a much more reasonable outcome. In equilibrium, the network achieves monopoly outcome (i.e., Q^*, which maximizes the joint industry profit), and leaves zero profit for the merchants.

If the network charges a unit fee, t_C, to consumers and a proportional fee, τ_M, to merchants, the result is the same. In particular, merchant FOC is:

$$\frac{1}{m}p'(Q)Q + p(Q) = \frac{\kappa_M}{1-\tau_M} + t_C.$$

And the network's profit is

$$[p(Q) - \kappa_M - \kappa_N]\,Q + (1-\tau_M)\,\frac{1}{m}p'(Q)Q^2.$$

The optimal choice is $\tau_M \to 1$ and $t_C = -\frac{\kappa_M}{1-\tau_M}$. Again, the network is able to obtain the monopoly outcome.

Proposition 7 *If a monopoly network uses mixed fees,*

(i) it will subsidize one side (to which it charges the unit fee) and extract maximal surplus from the other side (to which it charges the proportional fee).

(ii) The network earns the monopoly profit, while merchants earn zero profit.

(iii) Social welfare and consumer surplus are larger under mixed fees than under pure fees (i.e., when both sides pay unit fees or both sides pay proportional fees).

Mixed fee is a combination of proportional and unit fee schemes, but the equilibrium is not a convex combination of each pure component's outcome. Basically, mixed fees give the network two truly independent instruments. The network will use a negative unit fee to subsidize one side and maximize the size of the pie, and use the remaining tool of proportional fee to extract all the surplus from the other side. In the main models, subsidy to one side does not work because, due to price neutrality, what really matters is the composite fee. Since the network's average revenue comes from the composite fee, it cannot be negative in equilibrium.

As a result, starting from pure fees, a one-sided switch (either from unit fee to proportional fee, or from proportional fee to unit fee) will result in mixed fees and therefore will always improve both welfare and network profits.

6 Antitrust Cases and Policy Implications

6.1 Antitrust cases involving proportional or unit fees

As mentioned in Introduction, merchants have long complained about interchange fees, in terms of both the high level and the way it is calculated, for both debit cards and credit cards.[38] Partly in response to the complaints about debit cards, the Federal Reserve Board imposed a cap in 2011 that effectively halved the interchange fees for most domestic debit card transactions. The legislation, however, brought mixed and unintended consequences, with banks cancelling free current accounts and encouraging cardholders to use more costly credit cards to recover the loss (Manuszak and Wozniak 2017). Both the merchants and card-issuing banks have opposed this regulation, and the Financial CHOICE Act of 2017 would repeal the so-called Durbin Amendment (Getter 2017).

Prior to the introduction of Apple's iPad in 2010, Amazon was the dominant player in the ebook market with its Kindle, which used the wholesale model. Apple entered the competition by signing agency contracts with five major publishers, who in turn forced Amazon to switch to the agency model. Apple was eventually convicted of price-fixing, and its appeals to the circuit and supreme courts both failed. In its ruling in 2013, DOJ forced Apple to modify or terminate agency agreements.[39] This case is complicated because it involves complements (i.e., ebook reader such as Amazon's Kindle and Apple's iPad or iPhone), substitutes (the sale of an ebook cannibalized the same title's hardcover version, which is more lucrative for

[38] For credit cards, merchants and trade associations alleged that Visa, MasterCard, and major credit card issuers had engaged in price fixing and other allegedly anti-competitive practices. In December 2013, a U.S. District Court Judge approved a settlement of $7.25 billion, which lowered interchange fees in return for a provision that bar future lawsuits over the same issue. Many merchants disputed the provision. In June 2016, the United States Court of Appeals for the Second Circuit overturned the settlement.

[39] See United States v. Apple, Inc., et al., Civil Action No. 1: 12-CV-2826, https://www.justice.gov/atr/case-document/plaintiff-united-states-final-judgment-and-plaintiff-states-order-entering.

the publisher), as well as complementary arrangements such as the Most Favored Nation (MFN) clause that Apple imposed on publishers along with the agency model. The antitrust case focused on whether Apple and the five publishers conspired to raise ebook prices rather than the agency model per se. In fact, Supreme Court explicitly asserted that "both courts emphasized that their decisions cast no doubt on the 'broader legality' of the contract terms at issue, such as the 'agency model and MFNs'" (Supreme Court 2015). The agency model is involved in many other cases mentioned in Introduction because price-fixing is easier when retail prices are set by retailers, especially coupled with MFN.

In a more recent case in November 2018, the Supreme Court of the United States is considering whether consumers have the right to sue Apple Inc. for anticompetitive conduct. At issue is Apple's practice of charging 30 per cent commission to any software sold in Apple's exclusive Apple Store.

6.2 Policy implications

For payment cards, there have been extensive studies about the socially optimal fee levels;[40] that is why we have focused instead on the fee type. In terms of policy implementation, enforcing a particular fee type is much easier than targeting a socially optimal fee level, as the latter is more demanding in information collection and subject to strategic manipulation.

The possibility for proportional fees to reduce welfare has obvious policy implications. For example, based on their theoretical finding that proportional fees always hurt merchants and always benefit consumers, Shy and Wang (2011) suggest that the antitrust authority should not heed the complaints by merchants. We show that it is perfectly possible for proportional fees to hurt both consumers and merchants, so the merits of proportional fees must be evaluated more carefully. For another example, recent court rulings in technology licensing usually favor the per-unit regime, especially in cases involving standard-setting organizations (SSOs) such as the Institute of Electrical and Electronics Engineers (IEEE). This worries Llobet and Padilla (2016), who argue that ad valorem royalty welfare dominates unit royalty. Llobet and Padilla's conclusion is established for a monopoly inventor facing sub-convex demands. Our research points out that unit royalty may well be socially desirable if the demand is super-convex.

Because proportional fees improve welfare when network competition is sufficiently strong, the easiest way to promote welfare is to encourage competition (among networks and/or among merchants as well as merchant multi-homing) rather than forcing the fee to drop, which usually involve costly litigations.[41] Competition can be encouraged by regulations that reduce

[40] See, for example, Baxter (1983), Bedre-Defolie and Calvano (2013), Rochet and Tirole (2002), Wright (2004; 2013), and Wang (2010) among others.

[41] Our analysis demonstrates that the private and social incentives diverge in different ways depending on the competition among networks. When competition is weak, the welfare consequence is ambiguous, but a monopoly network always benefits. When competition is strong, welfare always improves but all networks may be hurt. For monopoly, therefore, the regulatory concern is an over-utilization of proportional fees, as a

switching cost. For example, Facebook users should be allowed to take their social connections with them when migrating to another network. Ditto for competing entertainment networks such as Netflix, Disney, AT&T-Time Warner, Comcast, YouTube, Amazon and Apple.[42]

If network competition is infeasible (in, for example, licensing of a monopoly technology, product or service standards such as 5G mobile-communications technology, or a monopoly platform with strong network effect such as Amazon.com or Google), then the antitrust authority should consider carefully all the market conditions including demand, cost, and merchant competition. For examples, although the costs of network and merchants both affect the equilibrium output in the same way, they have opposite impacts on proportional fees' welfare consequences.[43]

For monopoly network, since decreased merchant profitability is necessary but insufficient to indicate a decreased consumer surplus, if merchants do not complain, then the authority needs not bother. Moreover, proportional fees lower social welfare if and only if it lowers the joint profit of merchants and the network. If profit transfer is feasible between the two parties and yet merchants continue to complain, then it is a strong indication that proportional fees may be hurting social welfare, and antitrust concerns should be raised.

Finally we would like to point out a peculiar feature of proportional fees, that the merchants' equilibrium profit is non-monotonic in their own costs. In particular, when κ_M is sufficiently small, the profit increases with the cost.[44] When a firm's profit increases with its own cost, the firm's incentive to reduce the cost through innovation will be discouraged. Since this perversive effect exists only for proportional fees, an important implication is that if merchant innovation is of some concern, proportional fees may have an additional, adverse effect on social welfare.

social planner may have preferred unit fees in some situations. For competing networks, however, an under-utilization of proportional fees needs not worry policy makers, as each network's individual incentive can be strong even if every network earns lower profits in equilibrium. That is, competition among networks may cause an across-the-board adoption of proportional fees to the benefits of consumers and social welfare.

[42] "Who will win the media wars?" *The Economist*, Nov.14, 2019, https://www.economist.com/leaders/2019/11/14/who-will-win-the-media-wars.

[43] For constant elastic demand, welfare improves when the network cost is positive, and remains unchanged if the cost is zero. For super-convex demands, if the network's cost is zero, then welfare always decreases. If the network's cost is positive but the merchants' cost is zero, welfare always increases. The importance of network cost suggests that many existing conclusions established under the innocuous assumption of zero cost need to be re-evaluated.

[44] A formal, analytical proof is available upon request. The property is also noted by Llobet and Padilla (2016, p.55) in a numerical calculation.

7 Appendix

7.1 Monopoly network: welfare

$$
\begin{aligned}
& sign\left\{\frac{\partial \Pi^P}{\partial Q}-\frac{\partial \Pi^U}{\partial Q}\right\} \\
= \; & sign\left\{\frac{\partial (\Pi^P-\Pi^U)}{\partial Q}\right\} \\
= \; & sign\left\{\frac{\partial\left[-\tau\frac{1}{m}p'Q^2\right]}{\partial Q}\right\}=-sign\left\{\tau\left(p''Q+2p'\right)-\frac{\partial\tau}{\partial Q}p'Q\right\} \\
= \; & sign\left\{\tau\left(\rho+2\right)-\frac{\partial\tau}{\partial Q}\right\} \\
= \; & sign\left\{\underbrace{\tau(\rho+2)}_{\text{Earnings Effect }(+)}+\underbrace{\left[-\left(1-\tau\right)\frac{\rho+m+1}{m\varepsilon-1}\right]}_{\text{Concession Effect }(-)}\right\},
\end{aligned}
$$

where $\frac{\partial\tau}{\partial Q}=\left(1-\tau\right)\frac{\rho+m+1}{m\varepsilon-1}$.

7.2 Monopoly network: comparative statics

$$
MR^P(Q)-MR^U(Q)=\left\{\underbrace{\tau(Q)\left(\rho+2\right)}_{\text{Earnings effect }(+)}+\underbrace{\left[-\left(1-\tau(Q)\right)\frac{\rho+m+1}{m\varepsilon-1}\right]}_{\text{Concession Effect }(-)}\right\}\times\left[-p'(Q)\frac{Q}{m}\right],
$$

which indicates $MR^P(Q)>MR^U(Q)$ is equivalent to

$$
\kappa_M<p(Q)\frac{\left(\rho+2\right)\left(m\varepsilon-1\right)^2}{m\varepsilon\left\{m\left[\varepsilon\left(\rho+2\right)+1\right]-1\right\}}\equiv\widehat{\kappa}_M. \tag{31}
$$

This is for any arbitrary Q. If $Q=Q^U$, then condition (31) is equivalent to (11). Since κ_M is exogenous, for any given Q and $p(Q)$ (hence ε and ρ are fixed), (31) is more likely to hold when κ_M is smaller.

Note that, $\frac{\partial\frac{(\rho+2)(m\varepsilon-1)^2}{m\varepsilon\{m[\varepsilon(\rho+2)+1]-1\}}}{\partial m}=\frac{(\rho+2)(m\varepsilon-1)\varepsilon\{m[\varepsilon(2\rho+3)+2]-1\}}{\{m\varepsilon\{m[\varepsilon(\rho+2)+1]-1\}\}^2}$. For any given Q and $p(Q)$, $\frac{(\rho+2)(m\varepsilon-1)^2}{m\varepsilon\{m[\varepsilon(\rho+2)+1]-1\}}$ increases in m when $\rho\in\left(-\frac{m(3\varepsilon+2)-1}{2m\varepsilon},-\frac{\varepsilon+1}{\varepsilon}\right)$, and decreases in m when $\rho\in\left(-2,-\frac{m(3\varepsilon+2)-1}{2m\varepsilon}\right)$.

Finally, since both MRs are downward sloping, a larger κ_N will lead to a smaller Q and consequently a larger $p(Q)$. If ρ is constant, $\frac{(\rho+2)(m\varepsilon-1)^2}{m\varepsilon\{m[\varepsilon(\rho+2)+1]-1\}}$ decreases in ε and consequently decreases in Q (as ε increases in Q for super-convex demands). Therefore, when ρ is constant, (31) is more likely to hold when κ_N is larger.

7.3 Single-homing: equilibrium characterization

7.3.1 Unit fees

By taking derivative of (13) with respect to t_j and t_{-j}, we have

$$\begin{aligned} \left[\rho+(m+1)k\right]\frac{\partial Q_j}{\partial t_j}+(k-1)\left(\rho+mk\right)\frac{\partial Q_{-j}}{\partial t_j} &= \frac{mk}{p'}; \\ \alpha\left(\rho+mk\right)\frac{\partial Q_j}{\partial t_j}+\left\{(k-\alpha)\rho+k\left[1+m\left(k-\alpha\right)\right]\right\}\frac{\partial Q_{-j}}{\partial t_j} &= 0, \end{aligned}$$

where $\frac{\partial Q_{-j}}{\partial t_{-j}}=\frac{\partial Q_j}{\partial t_j}$ and $\frac{\partial Q_{-j}}{\partial t_j}=\frac{\partial Q_j}{\partial t_{-j}}$. Solving this equations system, we have

$$\begin{aligned} \frac{\partial Q_j}{\partial t_j} &= -\frac{Q}{p}\frac{m\varepsilon\left[(k-\alpha)\rho+k\left(mk-m\alpha+1\right)\right]}{\left(\rho+mk+1\right)\left[\left(1-a\right)\left(\rho+mk\right)+k\right]}; \\ \frac{\partial Q_{-j}}{\partial t_j} &= \frac{Q}{p}\frac{\alpha m\varepsilon\left(\rho+mk\right)}{\left(\rho+mk+1\right)\left[\left(1-a\right)\left(\rho+mk\right)+k\right]}. \end{aligned}$$

Plug $\frac{\partial Q_j}{\partial t_j}$ into $\frac{\partial \Pi_j^U}{\partial t_j}=Q_j+(t_j-\kappa_N)\frac{\partial Q_j}{\partial t_j}=0$ or, equivalently, $(t_j-\kappa_N)+\frac{Q_j}{\partial Q_j/\partial t_j}=0$. Finally, substitute $t_j=\frac{1}{mk}p'Q+p-\kappa_M$ and $Q_j=\frac{Q}{k}$ to arrive at the equilibrium characterization:

$$\kappa_M+\kappa_N=p'Q+p-p\frac{\left(\rho+mk+1\right)\left[\left(1-a\right)\left(\rho+mk\right)+k\right]}{km\varepsilon\left[(k-\alpha)\rho+k\left(mk-m\alpha+1\right)\right]}, \tag{32}$$

which will determine a unique equilibrium total output, Q^U. Note that p, p', p'' are all functions of Q. To summarize, under unit fees,

$$\begin{aligned} \frac{\partial Q_j}{\partial t_j} &= -\frac{Q}{p}\frac{m\varepsilon\left[(k-\alpha)\rho+k\left(mk-m\alpha+1\right)\right]}{\left(\rho+mk+1\right)\left[\left(1-a\right)\left(\rho+mk\right)+k\right]}; \\ \frac{\partial Q_{-j}}{\partial t_j} &= \frac{Q}{p}\frac{\alpha m\varepsilon\left(\rho+mk\right)}{\left(\rho+mk+1\right)\left[\left(1-a\right)\left(\rho+mk\right)+k\right]}; \\ t_j &= \frac{1}{mk}p'Q+p-\kappa_M. \end{aligned}$$

To ensure proper behavior of network choices, we need $\frac{\partial Q_j}{\partial t_j}<0$ and $\frac{\partial Q_{-j}}{\partial t_j}>0$, which is guaranteed if $\rho+mk>0$.

7.3.2 Proportional fee

By taking derivative of (17) with respect to τ_j and τ_{-j}, we have

$$\begin{aligned} \left[\rho+(m+1)k\right]\frac{\partial Q_j}{\partial t_j}+(k-1)\left(\rho+mk\right)\frac{\partial Q_{-j}}{\partial t_j} &= \frac{mk}{p'}\frac{\kappa_M}{(1-\tau)^2}; \\ \alpha\left(\rho+mk\right)\frac{\partial Q_j}{\partial t_j}+\left\{(k-\alpha)\rho+k\left[1+m\left(k-\alpha\right)\right]\right\}\frac{\partial Q_{-j}}{\partial t_j} &= 0, \end{aligned}$$

where $\frac{\partial Q_{-j}}{\partial \tau_{-j}} = \frac{\partial Q_j}{\partial \tau_j}$ and $\frac{\partial Q_{-j}}{\partial \tau_j} = \frac{\partial Q_j}{\partial \tau_{-j}}$. Solving this equations system, we have

$$\begin{aligned}
\frac{\partial Q_j}{\partial \tau_j} &= -\frac{Q\kappa_M}{p(1-\tau)^2}\frac{m\left[(k-\alpha)\rho + k(mk - m\alpha + 1)\right]}{(\rho + mk + 1)\left[(1-a)(\rho + mk) + k\right]}; \\
\frac{\partial Q_{-j}}{\partial \tau_j} &= \frac{Q\kappa_M}{p(1-\tau)^2}\frac{\alpha m(\rho + mk)}{(\rho + mk + 1)\left[(1-a)(\rho + mk) + k\right]}.
\end{aligned}$$

Plug $\frac{\partial Q_j}{\partial \tau_j}$ and $\frac{\partial Q_{-j}}{\partial \tau_j}$ into $\frac{\partial \Pi_j^P}{\partial \tau_j} = \left\{p_j + \tau_j p_j' \left[\frac{\partial Q_j}{\partial \tau_j} + (k-1)\frac{\partial Q_{-j}}{\partial \tau_j}\right]\right\}\frac{Q}{k} + (\tau_j p_j - \kappa_N)\frac{\partial Q_j}{\partial \tau_j} = 0$. Finally, substitute $\tau_j = 1 - \frac{mkp}{p'Q + mkp}\kappa_M$ and $Q_j = \frac{Q}{k}$ to arrive at the equilibrium characterization:

$$\begin{aligned}
\kappa_M + \kappa_N &= p - \frac{\kappa_M}{km\varepsilon - 1} \\
&\quad - p\frac{(\rho + mk + 1)\left[(1-a)(\rho + mk) + 1\right]}{\varepsilon\left[(k-\alpha)\rho + k(mk - m\alpha + 1)\right]} \\
&\quad - \kappa_M\frac{mk\left\{(1-\alpha)(\rho + mk)(1 + \varepsilon + \varepsilon\rho) + k\varepsilon\left[\rho + m(k-1) + 1\right] + 1\right\}}{(km\varepsilon - 1)^2\left[(k-\alpha)\rho + k(mk - m\alpha + 1)\right]},
\end{aligned} \tag{33}$$

which will determine the unique equilibrium output under proportional fees, Q^P. To summarize, under proportional fees,

$$\begin{aligned}
\frac{\partial Q_j}{\partial \tau_j} &= -\frac{Q\kappa_M}{p(1-\tau)^2}\frac{m\left[(k-\alpha)\rho + k(mk - m\alpha + 1)\right]}{(\rho + mk + 1)\left[(1-a)(\rho + mk) + k\right]}; \\
\frac{\partial Q_{-j}}{\partial \tau_j} &= \frac{Q\kappa_M}{p(1-\tau)^2}\frac{\alpha m(\rho + mk)}{(\rho + mk + 1)\left[(1-a)(\rho + mk) + k\right]}; \\
\tau_j &= 1 - \frac{mkp}{p'Q + mkp}\kappa_M.
\end{aligned}$$

To ensure proper behavior of network choices, we need $\frac{\partial Q_j}{\partial \tau_j} < 0$ and $\frac{\partial Q_{-j}}{\partial \tau_j} > 0$, which again is guaranteed if $\rho + mk > 0$.

7.4 Single-homing: welfare

On competing networks with single-homing merchants, we can write the profit function $\Pi_j^P = \Pi_j^U + \Delta\Pi_j$. These profits ($\Pi_j^P$, Π_j^U and $\Delta\Pi_j$) are functions of Q_j and Q_{-j}, which are in turn functions of $\mathbf{t}$ or $\boldsymbol{\tau}$.

The decomposition of changes in equilibrium outputs:

$$
\begin{aligned}
& sign\left\{\frac{\partial \Pi_j^P/\partial \tau_j}{\partial Q_j/\partial \tau_j}-\frac{\partial \Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \\
= & sign\left\{\frac{\partial\left(\Pi_j^U+\Delta\Pi_j\right)/\partial \tau_j}{\partial Q_j/\partial \tau_j}-\frac{\partial \Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \\
= & sign\left\{\frac{\partial \Delta\Pi_j/\partial \tau_j}{\partial Q_j/\partial \tau_j}+\frac{\partial \Pi_j^U/\partial \tau_j}{\partial Q_j/\partial \tau_j}-\frac{\partial \Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \\
= & sign\left\{\frac{\partial \Delta\Pi_j/\partial \tau_j}{\partial Q_j/\partial \tau_j}+\frac{(n-1)\frac{\partial \Pi_j^U}{\partial Q_{-j}}\frac{\partial Q_{-j}}{\partial \tau_j}+\frac{\partial \Pi_j^U}{\partial Q_j}\frac{\partial Q_j}{\partial \tau_j}}{\partial Q_j/\partial \tau_j}-\frac{(n-1)\frac{\partial \Pi_j^U}{\partial Q_{-j}}\frac{\partial Q_{-j}}{\partial t_j}+\frac{\partial \Pi_j^U}{\partial Q_j}\frac{\partial Q_j}{\partial t_j}}{\partial Q_j/\partial t_j}\right\} \\
= & sign\left\{\frac{\partial \Delta\Pi_j/\partial \tau_j}{\partial Q_j/\partial \tau_j}+(n-1)\frac{\partial \Pi_j^U}{\partial Q_{-j}}\left(\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j}-\frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j}\right)\right\} \\
= & sign\left\{\frac{\partial \Delta\Pi_j/\partial \tau_j}{\partial Q_j/\partial \tau_j}\right\} \text{ (Due to } \frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j}=\frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j}) \\
= & sign\left\{\frac{\partial\left[-\tau_j\frac{1}{m}p_j'Q_j^2\right]/\partial \tau_j}{\partial Q_j/\partial \tau_j}\right\} \\
= & sign\left\{\underbrace{\tau(\rho+2)\frac{\partial Q/\partial \tau_j}{\partial Q_j/\partial \tau_j}}_{\text{Earnings effect }(+)}+\underbrace{\frac{Q}{\partial Q_j/\partial \tau_j}}_{\text{Concession Effect }(-)}+\underbrace{2\tau(k-1)\left(1-\frac{\partial Q_{-j}/\partial \tau_j}{\partial Q_j/\partial \tau_j}\right)}_{\text{Business-stealing Effect }(+)}\right\}.
\end{aligned}
$$

By comparing (32) with (33), we can verify that $Q^P > Q^U$ if and only if (23) holds. Condition (23) is expressed in terms of κ_N. Alternatively, it can also be expressed in terms of κ_M:

$$\kappa_M < p(Q^U)\frac{(km\varepsilon-1)^2}{km\varepsilon}\frac{(\rho+2)[(1-\alpha)\rho+k((1-\alpha)m+1)]+(k-1)(\rho+2mk)}{\{mk[\varepsilon(\rho+2)+1]-1\}[(1-\alpha)(\rho+mk)+1]+(k-1)(2km\varepsilon-1)(\rho+mk+1)}.$$

7.5 Single-homing: Constant-elastic demand

For general demand, comparative statics are very complicated. To show the effects of each parameters, we offer an example with a constant-elastic demand, $p_j=\left(Q_j+\alpha\sum_{-j}Q_{-j}\right)^{-1/\varepsilon}$ with $\varepsilon\in(1,+\infty)$. Then the equilibrium can be directly solved.

Taking the second-order conditions (SOCs) into consideration, Figure 8 is plotted at $\varepsilon=1.5$, $\alpha=0.1$, $m=2$, $n=2$, $\kappa_M=\kappa_N=0.1$. If a parameter is chosen as the independent variable, its value is omitted. The following observations can be verified from the graphs. First, since the demand is constant-elastic, proportional fees always raise the equilibrium output. Second, networks, merchants, and consumers may all be better off, as seen from the upper four panels. Third, a larger output under proportional fees benefit merchants but hurt networks. This result is obtained from the fact that merchants' incremental profits are positively correlated with the incremental outputs (except the middle left figure where the change of output is due to the change of merchants' number), while networks' incremental profits are negatively correlated with the incremental outputs. Fourth, the supply chain's joint profit decreases when network differentiation is smaller, the number of networks or merchants is larger, or demand elasticity is smaller. Fifth, when merchants' or networks' cost is larger, networks tend to prefer proportional fees, while merchants tend to prefer unit fees. Sixth, $\Pi_M^P - \Pi_M^U > 0$ when α, ε, m, or n is large, or κ_N is small; $\Pi_N^P - \Pi_N^U > 0$ under the opposite conditions.

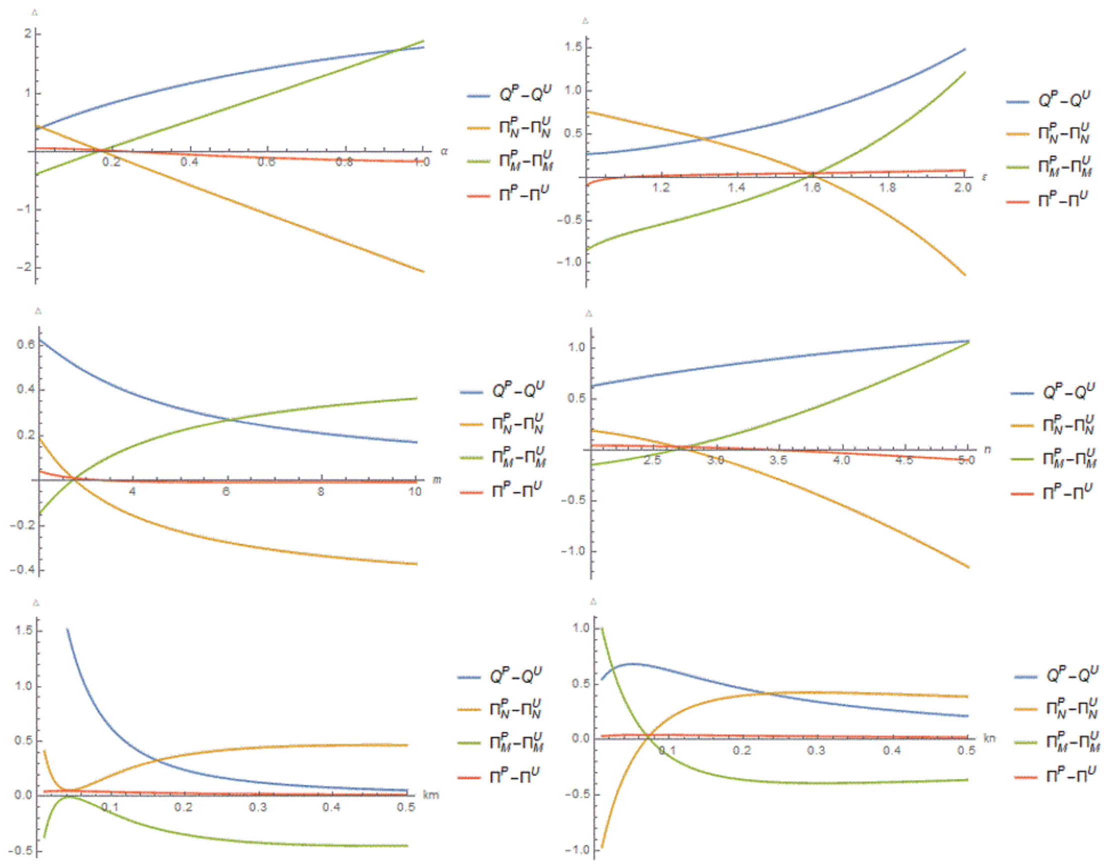

Figure 8: Single-homing and constant elastic demands

7.6 Single-homing: Joint profit and individual incentives

Suppose $\alpha = 1$. Consider linear demand $p = 1 - Q$. The first example demonstrates that proportional fees may raise the joint profit when $n = 2$, $m = 1$ (and when merchants' cost is large):

Table 2: $n = 2$, $m = 1$

	Unit fees				Proportional fees			
Profit	Network	Merchant	Joint	Price	Network	Merchant	Joint	Price
$\kappa_M = 0.2$; $\kappa_N = 0.2$	0.0267	0.0178	0.0444	0.733	0.0324	0.0118	0.0442	0.660
$\kappa_M = 0.4$; $\kappa_N = 0.2$	0.0119	0.0079	0.0198	0.822	0.0134	0.0065	0.0199	0.789

The next two examples demonstrate network's individual incentive to adopt proportional fees. Let $n = m = 2$. Then,

Table 3: Small κ_M

	Network's profit		Merchants' profit	
$\kappa_M = 0.01$; $\kappa_N = 0.2$	Unit fees	Proportional fees	Unit fees	Proportional fees
Unit fees	0.047; 0.047	0.018; 0.082	0.028; 0.028	0.01; 0.005
Proportional fees	0.082; 0.018	0.042; 0.042	0.005; 0.01	0.003; 0.003

Table 4: Large κ_M

	Network's profit		Merchants' profit	
$\kappa_M = 0.2$; $\kappa_N = 0.2$	Unit fees	Proportional fees	Unit fees	Proportional fees
Unit fees	0.027; 0.027	0.019; 0.037	0.016; 0.016	0.0117; 0.0123
Proportional fees	0.037; 0.019	0.028; 0.028	0.0123; 0.0117	0.009; 0.009

In both examples, proportional fee is a dominant strategy. It is Prisoners' Dilemma in Table 3 (when κ_M is small), but not in Table 4 (when κ_M is large). Note that merchants' profit increases with their own cost when both networks charge proportional fees ($0.009 > 0.003$), and merchants are worse off when networks switch from unit fees to proportional fees ($0.003 < 0.028$, and $0.009 < 0.016$).

7.7 Multi-homing: equilibrium characterization

7.7.1 Unit fees

From merchants' FOCs and combine them together, we have

$$
\begin{aligned}
t_j &= p_j + \frac{1}{m} p_j' Q_j + \frac{\alpha}{m} \sum_{-j} p_{-j}' Q_{-j} - \kappa_M; \\
\sum_{j=1}^{n} t_j &= \sum_{j=1}^{n} p_j + \frac{k}{m} \sum_{j=1}^{n} p_j' Q_j - n\kappa_M.
\end{aligned}
$$

By taking derivative with respect to t_j on both sides of the equations, we obtain an equation system, from which we can calculate $\frac{\partial Q_j}{\partial t_j}$ and $\frac{\partial Q}{\partial t_j} = \frac{\partial Q_j}{\partial t_j} + (k-1)\frac{\partial Q_{-j}}{\partial t_j}$. Then, we have

$$\begin{aligned}
\frac{\partial Q_j}{\partial t_j} &= \frac{m\{k(k-\alpha)(\rho+m+1)-\alpha(1-\alpha)\rho\}}{p'(1-\alpha)(\rho+m+1)k[(1-\alpha)\rho+k(m+1)]};\\
\frac{\partial Q_{-j}}{\partial t_j} &= -\frac{\alpha m\{k(\rho+m+1)+(1-\alpha)\rho\}}{p'(1-\alpha)(\rho+m+1)k[(1-\alpha)\rho+k(m+1)]};\\
t &= p+\tfrac{1}{m}p'Q-\kappa_M.
\end{aligned}$$

To ensure the equilibrium is well-behaved, we need $\frac{\partial Q_j}{\partial t_j} < 0$ and $\frac{\partial Q_{-j}}{\partial t_j} > 0$, which is guaranteed if $k(\rho+m+1)+(1-\alpha)\rho > 0$.

Substituting them into networks' FOC, we have

$$\kappa_M+\kappa_N = p+\frac{p'Q}{m}+\frac{p'Q(1-\alpha)(\rho+m+1)[(1-\alpha)(\rho+m+1)+\alpha n(m+1)]}{m\{k(k-\alpha)(\rho+m+1)-\alpha(1-\alpha)\rho\}}.$$

7.7.2 Proportional fees

From merchants' FOCs and combine them together, we have

$$\begin{aligned}
\kappa_M &= (1-\tau_j)p_j+\frac{1}{m}(1-\tau_j)p_j'Q_j+\frac{\alpha}{m}\sum_{-j}(1-\tau_{-j})p_{-j}'Q_{-j};\\
n\kappa_M &= \sum_{j=1}^{n}(1-\tau_j)p_j+\frac{k}{m}\sum_{j=1}^{n}(1-\tau_j)p_j'Q_j.
\end{aligned}$$

By taking derivative with respect to τ_j on both sides of the equations, we obtain an equation system, from which we can solve $\frac{\partial Q_j}{\partial \tau_j}$ and $\frac{\partial Q}{\partial \tau_j} = \frac{\partial Q_j}{\partial \tau_j} + (k-1)\frac{\partial Q_{-j}}{\partial \tau_j}$. Then, we have

$$\begin{aligned}
\frac{\partial Q_j}{\partial \tau_j} &= \frac{p'Q^2}{\kappa_M}\frac{(m\varepsilon-1)\{(\rho+m+1)k[m\varepsilon(k-\alpha)-(1-\alpha)]-\alpha\rho(1-\alpha)(m\varepsilon-1)\}}{m(1-\alpha)(\rho+m+1)k[(1-\alpha)\rho+k(m+1)]};\\
\frac{\partial Q_{-j}}{\partial \tau_j} &= -\frac{p'Q^2}{\kappa_M}\frac{\alpha(m\varepsilon-1)\{k(\rho+m+1)+(1-\alpha)(m\varepsilon-1)\rho\}}{m(1-\alpha)(\rho+m+1)k[(1-\alpha)\rho+k(m+1)]};\\
\tau &= 1-\frac{m\kappa_M}{mp+p'Q}.
\end{aligned}$$

To ensure the equilibrium is well-behaved, we need $\frac{\partial Q_j}{\partial \tau_j} < 0$ and $\frac{\partial Q_{-j}}{\partial \tau_j} > 0$, which is guaranteed if $k(\rho+m+1)+(1-\alpha)(m\varepsilon-1)\rho > 0$.

Substituting them into networks' FOC, we have

$$\begin{aligned}
\kappa_M+\kappa_N &= p-\frac{\kappa_M}{m\varepsilon-1}\\
&\quad -\frac{\kappa_M}{m\varepsilon-1}\frac{(1-\alpha)m\{(\rho+m+1)[\varepsilon(\rho+k)+1]-\alpha\rho[\varepsilon(\rho+1)+1]\}}{\{(\rho+m+1)k[m\varepsilon(k-\alpha)-(1-\alpha)]-\alpha(1-\alpha)(m\varepsilon-1)\rho\}}\\
&\quad +\frac{p'Q(1-\alpha)\{(\rho+m+1)(km\varepsilon-1)-\alpha\rho(m\varepsilon-1)\}}{\{(\rho+m+1)k[m\varepsilon(k-\alpha)-(1-\alpha)]-\alpha(1-\alpha)(m\varepsilon-1)\rho\}}.
\end{aligned}$$

7.8 Multi-homing: welfare

We can again write the profit function $\Pi_j^P = \Pi_j^U + \Delta\Pi_j$, with Π_j^P, Π_j^U and $\Delta\Pi_j$ being functions of Q_j and Q_{-j}, which in turn are functions of $\mathbf{t}$ or $\boldsymbol{\tau}$.

The decomposition of changes in equilibrium outputs:

$$
\begin{aligned}
& sign\left\{\frac{\partial\Pi_j^P/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial\Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \\
= \; & sign\left\{\frac{\partial\left(\Pi_j^U + \Delta\Pi_j\right)/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial\Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \\
= \; & sign\left\{\frac{\partial\Delta\Pi_j/\partial\tau_j}{\partial Q_j/\partial\tau_j} + \frac{\partial\Pi_j^U/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial\Pi_j^U/\partial t_j}{\partial Q_j/\partial t_j}\right\} \\
= \; & sign\left\{\frac{\partial\Delta\Pi_j/\partial\tau_j}{\partial Q_j/\partial\tau_j} + \frac{(n-1)\frac{\partial\Pi_j^U}{\partial Q_{-j}}\frac{\partial Q_{-j}}{\partial\tau_j} + \frac{\partial\Pi_j^U}{\partial Q_j}\frac{\partial Q_j}{\partial\tau_j}}{\partial Q_j/\partial\tau_j} - \frac{(n-1)\frac{\partial\Pi_j^U}{\partial Q_{-j}}\frac{\partial Q_{-j}}{\partial t_j} + \frac{\partial\Pi_j^U}{\partial Q_j}\frac{\partial Q_j}{\partial t_j}}{\partial Q_j/\partial t_j}\right\} \\
= \; & sign\left\{\frac{\partial\Delta\Pi_j/\partial\tau_j}{\partial Q_j/\partial\tau_j} + (n-1)\frac{\partial\Pi_j^U}{\partial Q_{-j}}\left(\frac{\partial Q/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial Q/\partial t_j}{\partial Q_j/\partial t_j}\right)\right\} \\
= \; & sign\left\{\frac{\partial\left[-\frac{1}{m}\tau_j p_j' Q_j^2 - \frac{\alpha}{m}\sum_{-j}\tau_{-j}p_{-j}'Q_{-j}Q_j\right]/\partial\tau_j}{\partial Q_j/\partial\tau_j} + (n-1)\frac{\partial\Pi_j^U}{\partial Q_{-j}}\left(\frac{\partial Q/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial Q/\partial t_j}{\partial Q_j/\partial t_j}\right)\right\} \\
= \; & sign\left\{\begin{array}{c}
\underbrace{\frac{Q}{\partial Q_j/\partial\tau_j}}_{\text{Concession Effect }(-)} + \underbrace{\tau(\rho+2)\frac{\partial Q/\partial\tau_j}{\partial Q_j/\partial\tau_j}}_{\text{(Direct) Earnings Effect }(+)} + \underbrace{2\tau(k-1)\left(1 - \frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j}\right)}_{\text{Business Stealing Effect }(+)} \\
\underbrace{+\tau(k-1)\left[(\rho+2)\left(\alpha + (k-\alpha)\frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j}\right) + (k-2\alpha)\left(1 - \frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j}\right)\right]}_{\text{(Indirect) Earnings Effect }(+)} \\
+\underbrace{\left\{-(k-1)\left[(k+1-\alpha)\rho + (m+1)k\right]\left(\frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j}\right)\right\}}_{\text{Allocation Effect }(+)}
\end{array}\right\}.
\end{aligned}
$$

Note that $\partial Q_j/\partial\tau_j < 0$, $\partial Q_{-j}/\partial\tau_j > 0$, $\partial Q/\partial\tau_j < 0$. It is easy to verify that direct earnings effect and business stealing effect are positive, and concession effect is negative. For the indirect earning effect, $\alpha + (k-\alpha)\frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j} = \frac{\partial Q_{-j}^*/\partial\tau_j}{\partial Q_j/\partial\tau_j}$ where $Q_{-j}^* = Q_{-j} + \alpha\sum_{i\neq -j}Q_i$ is the virtual output in network $-j$. Since the demand is downward sloping, we have $\partial Q_{-j}^*/\partial\tau_j < 0$. Therefore, $\alpha + (k-\alpha)\frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j} > 0$. Together with $1 - \frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j} > 0$, $(k-2\alpha) > 0$, and $(\rho+2) > 0$, this proves that indirect earning effect is positive.

For allocation effect, we have $\frac{\partial\Pi_j^U}{\partial Q_{-j}} = -\left[(k+1-\alpha)\rho + (m+1)k\right] < 0$ and $\frac{\partial Q_{-j}/\partial\tau_j}{\partial Q_j/\partial\tau_j} - \frac{\partial Q_{-j}/\partial t_j}{\partial Q_j/\partial t_j} = -\frac{\alpha(1-\alpha)k(\rho+m+1)[(1-\alpha)\rho+k(m+1)]}{(m\varepsilon-1)\{(\rho+m+1)k[m\varepsilon(k-\alpha)-(1-\alpha)]-\alpha\rho(1-\alpha)(m\varepsilon-1)\}\{k(k-\alpha)(\rho+m+1)-\alpha(1-\alpha)\rho\}} < 0$.

7.9 Multi-homing: Constant-elastic demand

Let $p_j(Q) = Q^{-\frac{1}{\varepsilon}} = \left(Q_j + \alpha \sum_{-j} Q_{-j}\right)^{-\frac{1}{\varepsilon}}$ with $\varepsilon > 1$. Then the equilibrium can be solved directly. Taking SOCs into consideration, Figure 9 shows comparative statics results for $\varepsilon = 1.5$, $\alpha = 0.5$, $m = 1$, $n = 4$, $\kappa_M = \kappa_N = 1$. As before, if a parameter is chosen as the independent variable, its value is omitted. The following observations can be verified from the graphs. First, since the demand is constant-elastic, the equilibrium output always increases. Second, when $\alpha = 1$, networks are indifferent between the two fee schemes, while merchants are strictly worse off (the upper left panel). Third, the supply chain's joint profit decreases when network differentiation is small, the number of networks or merchants is large, merchants' cost is small, or networks' cost is large. If that is the case, networks and merchants cannot both gain. Fourth, consumers and networks always gain from proportional fees (this is consistent with Johnson (2017) as the demand here is constant elastic). However, the merchants gain only if network differentiation is large (together with small number of networks and merchants, small demand elasticity, and large merchant/network cost ratio).

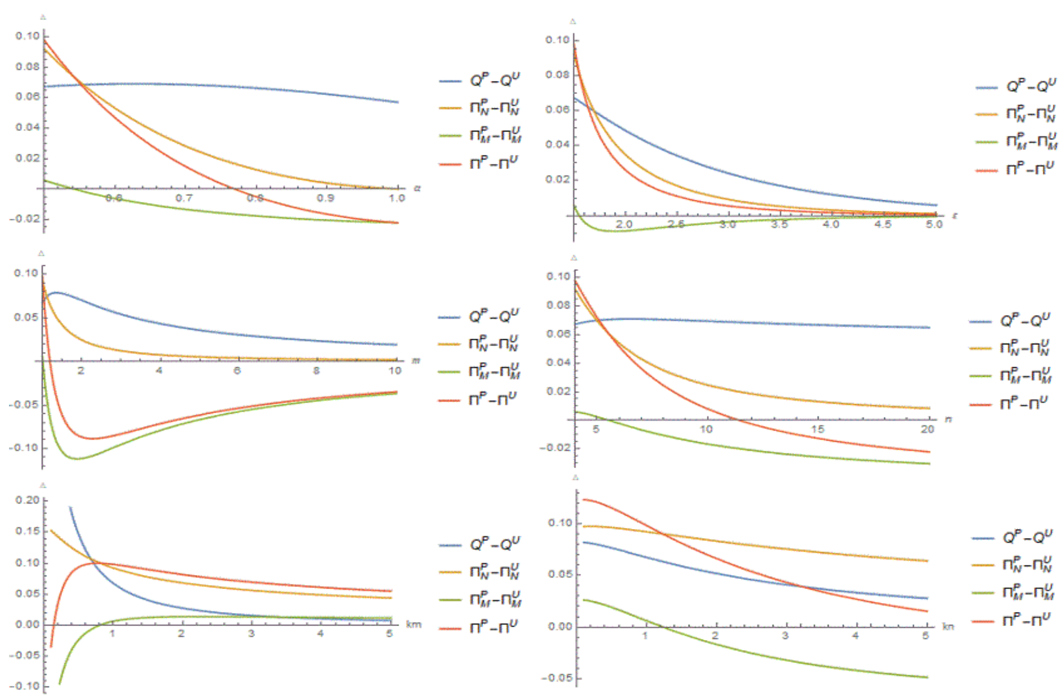

Figure 9: Multi-homing and constant elastic demands

References

[1] Amir, R., & Lambson, V. E. 2000. "On the effects of entry in Cournot markets." *The Review of Economic Studies*, 67(2), 235-254.

[2] Anderson, Simon P., Andre de Palma, and Brent Kreider. 2001. "The efficiency of indirect taxes under imperfect competition." *Journal of Public Economics* 81(2): 231–251.

[3] Armstrong, Mark. 2006. "Competition in two-sided markets." *RAND Journal of Economics* 37 (3): 668–91.

[4] Badinger, H. 2007. "Has the EU's Single Market Programme Fostered Competition? Testing for a Decrease in Mark-up Ratios in EU Industries." Oxford *Bulletin of Economics and statistics*, 69(4), 497-519.

[5] Baxter, William F. 1983. "Bank Interchange of Transactional Paper: Legal and Economic Perspectives." *Journal of Law and Economics* 26 (3): 541–88.

[6] Bedre-Defolie, Ozlem and Emilio Calvano. 2013 "Pricing Payment Cards." *American Economic Journal: Microeconomics* 5(3): 206–231.

[7] Besen, Stanley M., and Joseph Farrell. 1994. "Choosing How to Compete: Strategies and Tactics in Standardization." *Journal of Economic Perspectives* 8(2): 117–31.

[8] Bishop, Robert L. 1968. "The effects of specific and ad valorem taxes." *The Quarterly Journal of Economics* 82(2): 198–218.

[9] Chen, Y., & Riordan, M. H. 2007. "Price and variety in the spokes model." *The Economic Journal*, 117(522), 897-921.

[10] Delipalla, Sofia, and Michael Keen. 1992. "The comparison between ad valorem and specific taxation under imperfect competition." *Journal of Public Economics*, 49(3): 351–367.

[11] Farrell, Joseph, and Garth Saloner. 1986. "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation." *American Economic Review* 76(5): 940–55.

[12] Federal Reserve: The Federal Reserve Payments Study: 2017 Annual Supplement.

[13] Gaudin, Germain, and Alexander White. 2014a. "On the Antitrust Economics of the Electronic Books Industry." mimeo.

[14] Gaudin, Germain, and Alexander White. 2014b. "Unit vs. Ad Valorem Taxes under Revenue Maximization." mimeo.

[15] Getter, Darryl E. 2017. "Regulation of Debit Interchange Fees." *Congress Research Service Report 7-5700*.

[16] Gu, Dingwei, Zhiyong Yao, Wen Zhou, and Rangrang Bai. 2018. "How Can Downstream Firms Benefit from Upstream Collusion?" mimeo.

[17] Gu, Dingwei, Zhiyong Yao, Wen Zhou, and Rangrang Bai. 2019. "When is Upstream Collusion Profitable?." *RAND Journal of Economics*. 50(2): 326–341.

[18] Hesse, Renate. 2015. "Business review letter from the acting assistant attorney general of the Antitrust Division, US Department of Justice, to Michael A. Lindsay, Institute of Electrical and Electronics Engineers." http://www.justice.gov/atr/public/busreview/311470.pdf.

[19] Katz, Michael L., and Carl Shapiro. 1985. "Network Externalities, Competition, and Compatibility." *American Economic Review* 75(3): 424–40.

[20] Katz, Michael L., and Carl Shapiro. 1994. "System Competition and Network Effect." *Journal of Economic Perspectives* 8(2): 93–115.

[21] Johnson, Justin P. 2017. "The Agency Model and MFN Clauses." *Review of Economic Studies* 84: 1151–1185.

[22] Llobet, Gerard, and Jorge Padilla. 2016, "The Optimal Scope of the Royalty Base in Patent Licensing." *Journal of Law and Economics* 59: 45–73

[23] Manuszak, Mark D., and Krzysztof Wozniak. 2017. "The Impact of Price Controls in Two-Sided Markets: Evidence from US Debit Card Interchange Fee Regulation." FEDS Working Paper No. 2017-074. Available at SSRN: https://ssrn.com/abstract=2999628 or http://dx.doi.org/10.17016/FEDS.2017.074

[24] Mrázová, Monika, and J. Peter Neary. 2013. "Selection effects with heterogeneous firms." *Journal of the European Economic Association*. 1–41.

[25] Mrázová, Monika, and J. Peter Neary. 2017. "Not So Demanding: Demand Structure and Firm Behavior." *American Economic Review* 107(12): 3835–3874.

[26] Oates, Wallace E. 1985. "Searching for Leviathan: an empirical analysis." *American Economic Review* 75, 748–757.

[27] Rauscher, Michael, 1998. "Leviathan and competition among jurisdictions: the case of benefit taxation." *Journal of Urban Economics* 44, 59–67.

[28] Rochet, Jean-Charles, and Jean Tirole. 2002. "Cooperation among competitors: some economics of payment card associations." *RAND Journal of Economics* 33 (4): 549–70.

[29] Rochet, Jean-Charles, and Jean Tirole. 2003. "Platform Competition In Two-Sided Markets." *Journal of the European Economic Association* 1 (4): 990–1029.

[30] Rochet, Jean-Charles, and Jean Tirole. 2006. "Two-sided markets: a progress report." *RAND Journal of Economics* 37 (3): 645–67.

[31] Rysman, Marc. 2009. "The Economics of Two-Sided Markets." *Journal of Economic Perspectives* 23 (3): 125–43.

[32] Skeath S E , Trandel G A. 1994. "A Pareto comparison of ad valorem and unit taxes in noncompetitive environments." *Journal of Public Economics* 53 (1): 53-71.

[33] Suits, D. B., Musgrave, R. A. 1953. "Ad valorem and unit taxes compared." *The Quarterly Journal of Economics* 67(4): 598–604.

[34] Shy, Oz, and Zhu Wang. 2011. "Why Do Payment Card Networks Charge Proportional Fees?" *American Economic Review* 101(4): 1575–1590.

[35] U.S. Supreme Court. 2015. "Apple Inc. Petitioner versus United States et al." No. 15-565.

[36] Ward, M. B., Shimshack, J. P., Perloff, J. M., & Harris, J. M. 2002. "Effects of the private-label invasion in food industries." *American Journal of Agricultural Economics*, 84(4), 961-973.

[37] Wang, Zhu. 2010. "Market structure and payment card pricing: What drives the interchange?" *International Journal of Industrial Organization* 28 (1): 86–98.

[38] Wang, Zhu, and Julian Wright. 2017, "Ad valorem platform fees, indirect taxes, and efficient price discrimination." *RAND Journal of Economics* 48(2): 467–484.

[39] Weyl, E. Glen. 2010. "A Price Theory of Multi-Sided Platforms." *American Economic Review* 100 (4):1642–72.

[40] Wilson, John D. 1986. "A theory of interregional tax competition." *Journal of Urban Economics* 19, 296–315.

[41] Wilson, John D. 1999. "Theories of tax competition." *National Tax Journal* 52, 269–304.

[42] Wilson, John D., and David E. Wildasin. 2004. "Capital tax competition: bane or boon." *Journal of Public Economics* 88: 1065–1091.

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

[44] Wright, Julian. 2013. "Why payment cards fees are biased against retailers." *RAND Journal of Economics* 43 (4): 761–80.

[45] Zhelobodko, E., Kokovin, S., Parenti, M., & Thisse, J. F. 2012. "Monopolistic competition: Beyond the constant elasticity of substitution." *Econometrica*, 80(6), 2765-2784.

[46] Zodrow, George R. and Peter Mieszkowski. 1986. "Pigou, Tiebout, property taxation, and the underprovision of local public goods." *Journal of Urban Economics* 19, 356–370.