

TSE DIGITAL CENTER POLICY PAPER SERIES

MOBILE PAYMENTS AND INTEROPERABILITY: INSIGHTS FROM THE ACADEMIC LITERATURE

*Milo Bianchi, Matthieu Bouvard, Renato Gomes, Andrew Rhodes
and Vatsala Shreeti*

December 20, 2021, No. 4

Mobile Payments and Interoperability: Insights from the Academic Literature*

Milo Bianchi[†] Matthieu Bouvard[‡] Renato Gomes[§]
Andrew Rhodes[¶] Vatsala Shreeti^{||}

December 20, 2021

Abstract

We connect various streams of academic literature to shed light on how the degree of interoperability in mobile payments affects market outcomes and welfare. We organize our discussion around four dimensions of interoperability. First, we consider mobile network interoperability (whether clients of one telecom can access another telecom’s payment services) in connection with the IO literature on tying. Second, we discuss platform level interoperability (the ability to send money off-network) in light of the literature on compatibility. We also build on the behavioral IO literature to suggest how the effects of interoperability may be very heterogeneous across various types of firms and consumers, or even backfire. Third, we consider interoperability in the cash-in-cash-out agent network, in light of the literature on co-investment in network industries, and of more specific studies on ATMs’ interoperability. Fourth, we discuss how the literature in banking and on data ownership can be used to understand interoperability of data. We conclude with some broader remarks on policy implications and on possible directions for future research.

Keywords: Mobile Payments, Interoperability, Financial Inclusion, Competition Policy.

JEL codes: L51; L96; G23; G28; O16.

*This paper is part of the FIT IN Initiative, a research partnership between the Toulouse School of Economics and the Bill & Melinda Gates Foundation. We thank (without implicating) Sabine Bair, Seth Garz and Hakan Özyılmaz for very useful comments.

[†]Toulouse School of Economics, TSM, IUF, University of Toulouse Capitole: milo.bianchi@tse-fr.eu

[‡]Toulouse School of Economics, TSM, University of Toulouse Capitole: matthieu.bouvard@tse-fr.eu

[§]Toulouse School of Economics, CNRS, University of Toulouse Capitole: renato.gomes@tse-fr.eu

[¶]Toulouse School of Economics, University of Toulouse Capitole and CEPR: andrew.rhodes@tse-fr.eu

^{||}Toulouse School of Economics, University of Toulouse Capitole: vatsala.shreeti@tse-fr.eu

1 Introduction

Digital financial services are often viewed as a key tool towards financial inclusion (Demirguc-Kunt, Klapper, Singer and Ansar (2018)). New technologies allow a significant decrease of transaction costs, which may allow reaching consumers who have been traditionally under-served (Goldfarb and Tucker (2019)). Moreover, the use of digital financial services generates data, which may be used to expand the set of services households can have access to.

A key step, often the first key step, in this direction is to move out from a fully cash-based economy and increase the use of digital payments. While giving access to a bank account has been shown to improve financial capabilities (Ashraf, Karlan and Yin (2006), Dupas and Robinson (2013)), a substantial fraction of the world population remains unbanked. At the same time, many of the unbanked have a mobile phone. Accessing digital money through mobile network is then an important way to decrease the reliance of cash, and a natural stepping-stone towards accessing other financial services such as credit and insurance (Demirguc-Kunt et al. (2018)).

When it comes to implementing a mobile payment network, several fundamental questions should be confronted. The first set of questions concerns the balance between competition and cooperation between service providers. As an industry characterized by network externalities, market forces may lead to the emergence of a dominant player. When is it beneficial to promote competition and when is it beneficial to exploit network effects? Moreover, a well-functioning payment network requires significant investments in infrastructures (say, to guarantee network security or to convert cash in/out of digital money), which often have a public good component. How to ensure that market participants have the right incentives to undertake those investments?

A second key dimension concerns the governance of the key assets. For example: What is the role of regulatory authorities, which dimensions should these authorities control, and which dimensions should be left to platforms' self-regulation and to market forces? For example: Should these authorities intervene on prices, and if so how? How should data access and ownership be organized?

A third key dimension, which is of course directly related to the previous ones, concerns the broader welfare impact on the various market participants. The effects of interoperability are likely to be heterogeneous for example on poor vs. non-poor households or on small vs. large businesses. Assessing these welfare implications also requires understanding how access to mobile money may affect the access to a broader set of financial services.

In this paper, we review various streams of the academic literature in order to shed light on some of those questions. In the discussion, we also point out at whether some important questions remain relatively under-explored. We pay particular attention to the role of interoperability of mobile payment networks, which can be broadly defined as “the ability for mass users to perform specific transactions between accounts at different

providers” (Arabehty, Chen, Cook and McKay (2016)).

We view interoperability as a fundamental dimension that affects the way in which the various market participants interact. As such, interoperability intersects the above-mentioned questions in many ways. Is interoperability an effective way to promote entry and competition? How does interoperability affect platforms’ incentives to invest in possibly common infrastructures? Should interoperability be mandated by a public authority or would it emerge from decentralized platforms’ agreements? How does interoperability influence the use of data and the access to other financial services? How do these aspects depend on the specificities, say in terms of concentration or penetration, of the market at hand? The answers to these questions appear of first order importance to assess the ultimate effects of digital financial payments and in particular their welfare implications.

We organize our discussion by identifying four key dimensions of interoperability. We start by considering interoperability at the mobile network level, which refers to the extent to which the subscribers of the phone (basic) service by one telecom can use the payment (ancillary) service of any other telecom. Mobile payments in Africa are often offered jointly with a telecom service and exclusively by one single telecom operator. M-Pesa, for instance, is offered solely to subscribers of Safaricom, MTN Mobile Money solely to subscribers of MTN, Airtel Money to solely subscribers of Airtel, as it also happens in many other markets. As we discuss in Section 2, the lack of interoperability at the mobile network level is allusive of what the literature on competition policy calls tying, whereby a (typically dominant) firm in one market (here, telecom) makes the sale (or price) of its main product conditional upon the purchaser also buying another product (here, mobile payment) from it.

We then consider interoperability at the platform level, which refers to the ability of users of one provider to make transfers to those of another provider. Again, platform level interoperability is more the exception than the rule, although that has recently changed in some markets. Typically, clients of a telecom operator are not only obliged to use its proprietary mobile payment system, but can only send and receive money from other clients from that same telecom operator. As we discuss in Section 3, interoperability at the platform level is parallel to what the literature on competition policy calls compatibility. Interestingly, regulation could well mandate payments to be interoperable at the platform level (possibly under regulated interconnection fees) regardless of whether they are interoperable at the mobile network level. The converse is also true: interoperability at the mobile network level is neither necessary nor sufficient for interoperability at the platform level.

That most countries exhibit non-interoperable payment services at the mobile network level is surprising in light of the traditional Chicago School argument (reviewed below) positing that tying only occurs when it is harmless. Similarly, the Coase Theorem suggests that interoperability at the platform level should occur provided “property rights” are duly defined over the user bases of different providers. Yet, this form of interoperability is not widespread.¹ We highlight how lessons from the competition policy

¹This form of interoperability is de facto absent in some markets, where off-net transfers are only

literature may be used to illuminate these (perhaps with some exaggeration) puzzling market outcomes.

Our main questions are then: Why most markets are not interoperable at the mobile network level? Under what circumstances should this type of interoperability prevail? Why and when should telecom providers oppose interoperability at the platform level? How is the answer to this question affected by the intensity of network effects (inherent to payment services) and to the possibility that users engage in multi-homing? Are there reasons to mandate interoperability at either level?

In Section 4, we discuss what can go wrong with interoperability or, more precisely, we highlight that, when implementing interoperability, several delicate dimensions should be considered so as to make sure that interoperability does not backfire. For concreteness, we mostly frame our discussion in terms of platform level interoperability, but our insights can be naturally extended to other dimensions of interoperability. We build in particular on the literature in behavioral IO, which studies markets populated by consumers with different degrees of experience or understanding of the market.² This literature emphasize how standard effects of pro-competitive measures can be considerably weakened, or even reversed, once one takes into account that a fraction of market participants is “non-standard”, in the sense for example that they face search costs, limited attention, or limited understanding of the various fee structures. This literature is also particularly useful to highlight how market forces may often lead to cross-subsidization between different groups of consumers, thereby speaking directly to how interoperability may affect various market participants in a highly heterogeneous way.

Our third key dimension refers to interoperability at the agent level. Cash-in-cash-out agents are a central component of the mobile network infrastructure. These agents ensure the possibility to transform cash into mobile money and mobile money into cash, which is essential in any economy in which households still largely rely on cash, and bank branches or ATMs are not easily accessible. A well functioning agent network (in terms of liquidity balances, portfolio of services provided, density of agents etc.) is key to promote the usage of mobile payments (Aker, Prina and Welch (2020)), and of one its essential dimension is the extent to which users of one provider can access the network of agents developed and maintained by another provider, that is, agent interoperability.

In Section 5, we discuss how interoperability of mobile-money agents can be analyzed building on the literature on interoperability for ATMs, focusing in particular on the effects of various forms of interchange fees between banks and of surcharges to customers. We also discuss the relation with the literature on on co-investment in network industries, discussing how potential free-riding may severely impede investments in agent networks and how different pricing and co-investment policies may alleviate these concern and improve the network’s coverage. This literature also highlights how the optimality of these policies depends on the degree of market uncertainty (say, about the demand for mobile money) and on the investment cost (depending for example on whether the the investment takes place in rural or urban areas).

available at prohibitively high fees.

²Throughout the article, IO refers to industrial organization.

Our fourth key dimension concerns data interoperability. As mentioned earlier, the use of mobile payment is often viewed as a first step towards a broader access to financial services. Data interoperability is key in this process, as it refers to the extent to which the data generated when dealing with one service provider can be transferred and used by other service providers. This is clearly connected to issues of ownership of data and of privacy regulation.

As we emphasize in Section 6, data interoperability can be understood building on two important streams of literature. First, the standard literature in banking that highlights the complementary between payment and lending, as controlling the former creates an informational advantage in screening and monitoring lenders. Under this view, data interoperability can improve the functioning of credit markets while also potentially bringing important distributional effects both across borrowers and across lenders. Second, the literature on data control rights and on the efficient use of information, discussing how gaining access to the data can promote entry in the payment market and what the optimal level of information sharing is.

Despite this rich picture, this review is far from comprehensive. We have selected what we believe are useful contributions in the existing literature, in particular from IO and finance, to study our key dimensions of interoperability. We refer to Bourreau and Valletti (2015) for an insightful discussion in particular on the effects of various regulatory practices. We also refer to recent surveys which provide a complementary and more practical perspective together with institutional evidence from different countries (see for example Arabehty et al. (2016) and Naji (2020)).

2 Mobile Network Interoperability

This section considers interoperability at the mobile network level; that is, the extent to which the subscribers of the phone service by one telecom can use the payment service of any other telecom. We start by revisiting the traditional Chicago School argument positing that lack of interoperability only occurs when it is harmless. We then review several reasons why this argument may fail in practice.

2.1 The Chicago school argument

In what follows we adapt the standard Chicago-school argument (see, for instance, Posner (1978)) to the markets of telecom and mobile payments.

Consider two telecom providers offering phone services worth v_a and v_b to consumers at zero marginal cost. Each telecom provider also offers payment services worth u_a and u_b , again at zero marginal cost (for simplicity). Suppose firm a is a superior telecom provider, in that $v_a > v_b$, but an inferior payment provider, in that $u_a < u_b$. Moreover, suppose firm a is “overall” better, in that $v_a + u_a > v_b + u_b$.

Let us now consider the following two modes of competition. In the first, firms compete by offering bundles of telecom and payment services (tied competition). For brevity, we will use the word “tying” to refer to the lack of interoperability at the mobile

network level. In the second mode of competition, firms offer telecom and payment services on a stand-alone basis. Again for brevity, we say “untied competition” rather than “competition with services interoperable at the mobile network level”.

Under tied competition, firm a corners the market for the bundle by charging the price $v_a - v_b - (u_b - u_a) > 0$. In turn, under untied competition, firm a conquers the telecom market, charging the price $v_a - v_b > 0$, while firm b conquers the payments market, charging the price $u_b - u_a > 0$. Clearly, untied competition makes both firms better off.

Only when $u_a \geq u_b$, which is contrary to the assumed above, would tying *not* be harmful for firm a . In this case, firm a offers better telecom and payment products, conquering both markets regardless of whether it bundles sales or not (and obtaining the same profit in both cases). In conclusion, tying can only occur when it is harmless.

When does this logic fail?

The argument described above highlights the idea that tying (when it is not redundant) erodes the profits of the firm that is “overall” better. This suggests that tying should not be a concern to competition policy, as it is unprofitable whenever it is consequential for equilibrium outcomes. As it happens with every model, this conclusion might fail under different assumptions. In what follows, we will describe several instances where the Chicago School argument fails, highlighting the sources of departure from this important benchmark.

2.2 Strategic foreclosure

Consider the model of subsection 2.1 with one modification: firm b has yet to enter the payment market, which requires paying a fixed costs of $K > 0$ (e.g., necessary for building a network of agents). Crucially, firm a takes its tying decision before firm b makes the entry choice. Moreover, firm a ’s tying decision is irreversible, in that it cannot untie its product after learning that firm b gave up entering. Typically, technological choices can serve as a commitment device to not undo the tying knot.

Proceeding by backward induction, suppose firm a decides to tie its products. Firm b then anticipates that entering the payment market will be useless, resulting in a loss of $-K$. Firm b ’s optimal response then involves not paying this fixed cost, which implies that firm a will operate alone in the payment market. Firm a can therefore charge $v_a - v_b + u_a$, cornering the bundle market.

If, by contrast, firm a does not tie its products and $K < u_b - u_a$, firm b will enter the payment market. In the continuation game, this firm conquers the payment market, while firm a conquers the telecom market, obtaining a profit of $v_a - v_b$.

Because $u_a > 0$, it then follows that firm a does better by tying its products. This behavior is allusive of multi-product firms that offer “closed systems,” effectively shunning competition for ancillary services (such as payments). Articulated by Bernheim and Whinston (1990), the strategic foreclosure argument described above is reminiscent of the “top dog” strategy described in Fudenberg and Tirole (1984).

The strategic foreclosure argument offers a possible explanation for why a telecom operator may choose technological standards that do not allow other payment services to operate in its system. Indeed, it is possible that entry in payments would have been much more intense (by non-dominant telecoms or fintechs) in the absence of tying.

2.3 Dynamic leverage

Consider again the model of subsection 2.1 under the following modification: firm b has not yet entered either market and, to do so, needs to invest in R&D separately in each market. Suppose, for simplicity, that the investment costs the same amount I and leads to a “success” with the same probability $p \in (0, 1)$ in each market (statistically independently). A success in telecom (resp., payment) guarantees that $v_b = v_a + \Delta$ (resp., $u_b = u_a + \Delta$), where $\Delta > 0$. By contrast, a failure in the telecom (resp., payment) implies that $v_a = v_b + \hat{\Delta}$ (resp., $u_a = u_b + \hat{\Delta}$), where $\hat{\Delta} > \Delta$. Therefore, if at least one failure occurs, firm a is “overall” better, in that $v_a + u_a > v_b + u_b$. No investment leads to failure for sure. As in subsection 2.1, the decision to tie is credible even after investments are realized.

Consider first the case where firm a does not practice tying. In this case, firm b invests in telecom if and only if $p\Delta \geq I$ and invests in payments if and only if $p\Delta \geq I$. Under this condition, which we assume from now on, firm a ’s expected profit is $2(1-p)\hat{\Delta}$.

Consider now the case where firm a ties its products. In this case, firm b needs to be successful in both investments to make a profit (as firm a can corner the bundle market provided one failure occurs). Therefore, firm b refrains from investing if and only if $2p^2\Delta < 2I$. Under this condition, tying will raise firm a ’s expected profit from $2(1-p)\hat{\Delta}$ to $2\hat{\Delta}$. Therefore, tying works to discourage R&D by rivals.³

According to this theory, proposed by Choi and Stefanadis (2001), tying buttresses firm a ’s dominance in both markets, including the tying (telecom) one! Similarly to the strategic foreclosure theory, the dynamic leverage theory has the potential to explain the lack of serious contenders in many markets where a dominant telecom provider ties telecom and payment services.

2.4 Potential competition

Now consider the model of subsection 2.1 under (yet) other modifications: There are two periods, and firm b may enter the payment market in either period, but can enter the telecom market only in the second period. Entry costs K in each of these markets. Upon entry, firm b is more efficient than firm a , that is, $v_b > v_a$ and $u_b > u_a$. One important assumption, not needed until now, is that the payment service is one-way complementary to the telecom service. This is obviously true, as one cannot make a mobile payment without a mobile plan, but one can use the phone while not making payments with it. As before, the decision to tie is credible across periods. There is no discounting. For the sake of this subsection, let us denote the second period equilibrium

³If, by contrast, $p^2\Delta \geq I$, tying is not profitable for firm a .

revenue of firm b , if it enters both markets, by $\pi \equiv v_b + u_b - (v_a + u_a) > 0$. To streamline the exposition, we assume that $v_b - v_a = u_b - u_a$, that is, b 's advantage relative to a is the same in both markets.

Suppose firm a ties its products. Then firm b cannot enter either market in the first period. It would also choose not to enter (any market) in the second period if and only if

$$\pi < 2K, \tag{1}$$

in which case firm a 's profit is $2(v_a + u_a)$. If condition (1) is violated, firm a 's total profit consist of its first-period profit only: $v_a + u_a$.

Suppose firm a does not tie its products. Consider first the putative equilibrium where firm b enters the payment market in the first period, and the telecom market in the second. This sub-game possesses a continuum of equilibria differing in the prices posted in the first period. Namely, any price profile such that firm a charges $v_a + x$ for the telecom service, $\max\{u_a - x, 0\}$ for the payment service, while firm b charges $u_b - x$ for the payment service, with $x \in [u_a, u_b]$, is part of an equilibrium. We select the one where firms a and b equally share the efficiency gain produced by firm b 's entry, in which case $x = \frac{u_a + u_b}{2}$.⁴

Under this selection, the putative equilibrium proposed above is in fact an equilibrium if and only if firm b 's profits across the two period compensate the total entry costs:

$$\frac{u_b - u_a}{2} + \pi > 2K, \tag{2}$$

in which case firm a 's total profit is $v_a + \frac{u_a + u_b}{2}$.⁵

Direct comparison then reveals that tying is not profitable if it does not prevent the entry of firm b in both markets in the second period. However, tying can be profitable whenever firm b never enters any market, as

$$2(v_a + u_a) > v_a + \frac{u_a + u_b}{2} \iff v_a + u_a > \frac{u_b - u_a}{2}.$$

Firm b would enter both markets (one in each period) in the absence of tying, but would enter none with tying, if and only if (1) and (2) simultaneously hold. This illustrates the strategic use of tying to preserve and create market power in evolving industries, as proposed by Carlton and Waldman (1998).

The reasoning above illustrates why dominant telecom operators may close its system, denying access to fringe payment systems (e.g., those offered by fintechs or smaller telecom competitors). The idea is that stand-alone pricing enables early entry in payments, which is needed to amortize the total costs of entry in both markets. Absent this amortization, the smaller firm would refrain from entering at once in both markets, which preserves the incumbent's telecom profits.

⁴Qualitative results are similar under other equilibrium selection criteria.

⁵Note that entering solely on the payment market is worse for firm b than entering both markets.

2.5 Data-enabled learning

Carlton and Waldman (1998) propose another rationale for tying, which is closely related to the one from the previous subsection. Namely, consider again a model with two periods, assuming as before that firm b may enter the payment market in either period, but can enter the telecom market only in the second period. The payment sector exhibits data-enabled learning, in that cornering this market in the first period improves the second-period quality of payments by Δ . There is no entry cost for the payment market, but a cost K for the telecom market. We let the quality of the telecom service be identical across firms, $v \equiv v_a = v_b$, but assume that $u_a < u_b$ and that $\Delta > u_b - u_a$ (i.e., data-enabled learning is stronger than firm b 's initial advantage on payments). Finally, payment is again one-way complementary to telecom.

Suppose firm a ties its products. Then firm b cannot enter either market in the first period. It would also choose not to enter (any market) in the second period, as firm a could easily beat it. Firm a 's profit is then $2(v + u_a) + \Delta$.

If firm a does not tie its products, then the entrant corners the payment market in the first period. As in the previous section, we select equilibria in a way that firm a charges $v + \frac{u_a + u_b}{2}$ for the telecom service and zero for the payment service, while firm b charges $\frac{u_b - u_a}{2}$ for the payment service. In the second period, firm b would enter the telecom market if and only if

$$v + u_b + \Delta - (v + u_a) - K > \frac{u_b - u_a}{2} \iff \frac{u_b - u_a}{2} + \Delta > K.$$

In this case, firm a 's total profit is the first-period profit, which equals $v + \frac{u_a + u_b}{2}$. Tying is therefore profitable if and only if

$$2(v + u_a) + \Delta > v + \frac{u_a + u_b}{2} \iff v + u_a + \Delta > \frac{u_b - u_a}{2}.$$

Similarly to the previous subsection, tying, by eliminating data-enabled learning by the entrant, insulates the incumbent's primary market from potential competition.

2.6 Upgrading

Consider again the model of subsection 2.1 with the following modifications: There are two periods and firm a is the only telecom operator. Both firms offer a payment service, which is one-way complementary to the telecom service. To simplify, we assume consumers assign no value to the telecom service, $v_a = v_b = 0$, and that firm b offers a superior payment service: $u_b > u_a$. In the second period firms have the option to invest $Z > 0$ and upgrade the payment service, which adds $\lambda > Z$ to its original quality. One can think of this upgrade as the addition of extra services, such as insurance and credit features. The marginal cost is $c < \min\{u_a, u_b - u_a\}$ for payments and zero for telecom. A unit purchased in the first period of either good can be used in the second period. Prices have to be non-negative. The consumer can mix and match only if firm a does not tie.

Suppose firm a does separate selling. Then, selecting the most profitable equilibrium for firm a , this firm sells the telecom service by a price of $2u_b - c$, and charges zero for payments. Firm b conquers the payment market charging the price c . In the second period, only firm b upgrades the payment service, charging λ for the upgrade. Firm a 's profit is $2u_b - c$.

Now consider the case of tying. Firm a then charges $2u_a$ for the bundle in the first period, while firm b is excluded. In the second period, firm a upgrades and charges λ for it. Firm a 's profit is then $2u_a - c + \lambda - Z$. Therefore, tying is profitable if and only if

$$2u_a - c + \lambda - Z > 2u_b - c \quad \iff \quad \lambda - Z > 2(u_b - u_a).$$

The crucial idea here is that firms cannot commit in period one to upgrade the payment system in period two. Because the telecom service is durable (a questionable assumption in developing countries, where consumers typically use pre-paid plans and there is no commitment on tariffs), the monopolist telecom cannot practice price squeeze on the future surplus of an upgrade. If this surplus is large, the telecom prefers to exclude the more efficient payment system and appropriate the upgrade surplus herself. This argument is developed in Carlton and Waldman (2012).

2.7 Two-sidedness and the zero lower bound

Finally, let us now incorporate two-sidedness in the payment service. We do so by assuming that payments data can be monetized at a rate $\beta > u_b - u_a$ per transaction. Moreover, we impose a zero lower bound on prices, preventing them from being negative in either market. This restriction captures the idea that consumers could fraud the platform if they were paid to use it - see Choi and Jeon (2021) for further discussion of this assumption. As in subsection 2.1, suppose a is a superior telecom provider, in that $v_a > v_b$, but an inferior payment provider, in that $u_a < u_b$. Moreover, suppose a is "overall" better, in that $v_a + u_a > v_b + u_b$.

Under untied competition, firm a charges $v_a - v_b$ for the telecom service and 0 for the payment service. In turn, firm b charges v_b for telecom and $u_b - u_a$ for payments. Firm a conquers the telecom market, obtaining the profit $v_a - v_b$, while firm b conquers the payment market, obtaining the profit $u_b - u_a + \beta$.

Under tied competition, firm a charges $v_a - v_b + (u_a - u_b)$ for the bundled services, while firm b charges zero. Firm a 's profit is $v_a - v_b + (u_a - u_b) + \beta$ and b is excluded.

Clearly, firm a is better-off under tying! Intuitively, the zero lower bound renders the foreclosed firm b less aggressive in equilibrium. To appreciate this point, note that, under tying, firm b would be happy to set a negative price equal to $-\beta$. Were this possible, one would conclude that tying hurts firm a , as described in subsection 2.1.

The notion that two-sidedness together with the zero lower bound render tying profitable is a major insight of Choi and Jeon (2021).⁶ Crucially, tying is more likely to occur (and decrease welfare) when the tied market is sufficiently two-sided.

⁶See Amelio and Jullien (2012) for an illustration of how tying can relax the zero-lower bound constraint in a monopoly setting.

It is important to note that the argument above differs from the others in that tying is an ex-post profitable strategy for competing *in* the market (i.e., after the entrant firm is operating). In the theories discussed in subsections 2.2 to 2.6, tying worked mainly as a tool to discourage entry, hindering competition *for* the market.

2.8 Conclusion

The theories above debunk the Chicago school argument according to which tying may only occur when it is inconsequential for market outcomes. In particular, the theories of subsections 2.2 to 2.6 give four different arguments for why tying is a potentially effective tool to discourage future competition in either the tying market (telecom) or the tied market (payments), or simultaneously in both of these markets. As discussed above, these theories are compatible with poorly competitive outcomes in both telecom and payment markets.

Yet, none of these theories explains why a telecom operator would like to keep services tied even after the entry of a second (or even third) telecom/mobile payment system, as indeed occurred in many markets.⁷ Only the theory of subsection 2.7 makes the case that tying can be profitable as an *in-the-market* competitive tool.

The models above, however, have little to say about the incentives of competing telecom providers do make their payment services interoperable at the platform level (whether tying occurs or not). This is a feature of the fact that competition is always “all-or-nothing” in these models, preventing the co-existence of more than one service operating. The papers reviewed in the subsequent section relax this restrictive feature.

3 Platform Interoperability

The previous section emphasized the connection between the telecom and payment markets, analyzing how (and why) firms might decide to tie their offers. In this section, we will ignore the “telecom” aspect of mobile payments, rather focusing on how the network effects present on payment services affects firms’ decision to be interoperable at the platform level (or compatible, for short).

To this end, consider two payment providers, a and b , and assume for the moment that these services are not compatible. Consumers differ on their tastes for each provider, described by their position $x \in [0, 1]$ uniformly distributed in the Hotelling segment. Consumers also may differ on their valuation $\beta \geq 0$ for network size. Accordingly, the net utility of a consumer with valuation β located at x when she adopts provider $i \in \{a, b\}$ is

$$U_i(x, \beta, N_i) = \beta N_i - p_i - t_i(x),$$

where N_i represents the total number provider i ’s customers. We choose labels such that $t_a(x) = tx$ and $t_b(x) = t(1 - x)$, where $t \geq 0$, and focus on the case where the market is

⁷These theories can however rationalize this phenomenon if untying is technologically impossible to undo ex-post (i.e., after entry occurred). This seems like a strong assumption in the case of mobile payments.

fully covered (i.e., every consumer joins at least one provider).

If the consumer adopts both providers (multi-homes), her net utility is

$$U_{ab}(x, \beta, N_a, N_b) = \beta - p_a - p_b - t.$$

In turn, if services a and b are compatible, the net utility of a consumer with valuation b located at $x \in [0, 1]$ when she adopts provider $i \in \{a, b\}$ is rather

$$\tilde{U}_i(x, \beta, N_i) = \beta - p_i - t_i(x).$$

Accordingly, compatibility destroys network effects, in that each consumer can reach the same set of other consumers by joining either platform. Obviously, no consumer would multi-home in this case.

We assume that providers, first, simultaneously take compatibility decisions. Services are compatible if and only if both providers decide accordingly, in which case each pays a fixed cost F . Secondly, knowing the outcome of the first stage, firms simultaneously choose prices. We will describe equilibria of the game above assuming consumers can/not multi-home, and considering the cases where the magnitude of network effects (captured by β) fares large/small relative to brand tastes (captured by t). We impose the natural restriction that all prices have to be non-negative. Marginal costs are zero.

3.1 Strong network effects: Market tipping

Consider first the case where network effects are stronger than brand tastes, in that all consumers share the same valuation $\beta > 0$ and $t = 0$ (for simplicity). This captures situations where payment systems constitute homogenous independent products (not inheriting the brand tastes of other elements of the same eco-system), and where payments are often “non-local.” The latter means that agents are likely to transfer money to non-neighbors chosen arbitrarily, so that “global” network effects (i.e., considering the overall population) are important. This is to be opposed to the case where payments are local (within village members, for instance), so that coordination among few individuals is enough to generate maximal value to consumers. Local payments are better captured by assuming that the strength of network effects β is small relative to the intensity of brand preferences, captured by t . This case is treated in the next subsection.

We analyze first the situation where consumers single-home. We select equilibrium assuming beliefs are favorable in the sense of Caillaud and Jullien (2003) and Caillaud and Jullien (2001), meaning that consumers coordinate on a given platform and only migrate to the competitor if doing so is a dominant strategy.

In this environment, choose labels such that provider a is the winning payment service. If services are not compatible, equilibrium prices are then $p_a = \beta$ and $p_b = 0$, following which all consumers join provider a . By contrast, were services compatible, equilibrium prices would be $p_a = p_b = 0$. Obviously, the winning payment provider would oppose compatibility. Because we analyze the extreme case where $t = 0$, imposing compatibility does not raise welfare (although it would were t small but positive). However,

imposing compatibility does shift rents from the winning platform to consumers.

Consider now the situation where consumers can seamlessly multi-home (but lexicographically prefer not doing so), and construct equilibrium beliefs such that a consumer would multi-home only if doing so is a strictly dominant strategy. In the absence of compatibility, equilibrium again involves $p_a = \beta$ and $p_b = 0$, following which all consumers single-home at provider a . Again, compatibility would produce a Bertrand race leading to $p_a = p_b = 0$. So, similarly to the single-homing case, compatibility would be opposed by the dominant provider.

These rather extreme modeling assumptions are chosen to illustrate the idea that, in the absence of product differentiation, under full market coverage, firms will have insufficient incentives to make services compatible, regardless of whether consumers can/not multi-home. This idea is well established in the literature, dating back to Katz and Shapiro (1985).⁸ The analysis is more nuanced when “global” network effects are less important, so that brand tastes are enough to guarantee that the market does not “tip,” in that a single provider attracts all consumers.

3.2 Moderate network effects: Market sharing

We now analyze the market sharing scenario, where more than one payment provider operates. This is the likely outcome of economies where payments occur within small geographic areas, so that local coordination (at the village level, for instance) is enough to generate all network gains. In this case, it is useful to think that market power stems more from brand tastes than from network externalities. Accordingly, we assume in this subsection that $t > \mathbb{E}(\beta)$, where $\mathbb{E}(\beta)$ is the average network valuation across agents. For simplicity, we assume that β takes two values, β^l and β^h , with $\beta^l < \beta^h$, in which case $\mathbb{E}(\beta) = \lambda\beta^h + (1 - \lambda)\beta^l$, where λ is the fraction of high-valuation consumers (distributed iid across agents, and independently of the brand taste x).

We consider first the case where consumer cannot multi-home, having to choose a single payment provider. This assumption approximates reality when, for instance, dual-sim phones are very expensive or simply unavailable to consumers.

3.2.1 Single-homing consumers

Let us denote by N_a^l (resp., N_a^h) the share of low-valuation (resp., high-valuation) consumers joining firm a (and define N_b^l and N_b^h analogously for firm b). Clearly, $N_a = \lambda N_a^h + (1 - \lambda)N_a^l$.

In the absence of compatibility, each consumer picks the provider that maximizes her net utility. It is then straightforward to derive the following linear demand system:

$$N_a^h = \frac{1}{2} + \frac{p_b - p_a}{2t} + \frac{\beta^h(p_b - p_a)}{2t(t - \mathbb{E}(\beta))} \quad \text{and} \quad N_a^l = \frac{1}{2} + \frac{p_b - p_a}{2t} + \frac{\beta^l(p_b - p_a)}{2t(t - \mathbb{E}(\beta))}. \quad (3)$$

⁸For a more recent treatment, see Doganoglu and Wright (2006) and Farrell and Klemperer (2007).

Each provider maximizes profits, as given by $p_i N_i$. In the unique symmetric equilibrium, $p_a = p_b = t - \mathbb{E}(\beta)$, both firms share the market, and each obtains a profit of $\frac{1}{2}(t - \mathbb{E}(\beta))$. Direct computation reveals that the equilibrium welfare is

$$W = \frac{\mathbb{E}(\beta)}{2} - \frac{t}{4}.$$

By contrast, consider the case where payment services are compatible. The absence of network effects brings us back the standard Hotelling demands:

$$N_a^h = \frac{1}{2} + \frac{p_b - p_a}{2t} \quad \text{and} \quad N_a^l = \frac{1}{2} + \frac{p_b - p_a}{2t}.$$

The equilibrium is therefore $p_a = p_b = t$, both firms share the market, and each obtains a profit of $\frac{t}{2} - F$, where, recall, F is the fixed cost that each provider has to pay to render his service compatible. Equilibrium welfare is then

$$\tilde{W} = \mathbb{E}(\beta) - \frac{t}{4} - 2F.$$

Therefore, compatibility is socially optimal if and only if

$$\tilde{W} \geq W \quad \iff \quad F \leq \frac{\mathbb{E}(\beta)}{4}.$$

In turn, compatibility is privately optimal (for payment providers) if and only if

$$\frac{t}{2} - F \geq \frac{1}{2}(t - \mathbb{E}(\beta)) \quad \iff \quad F \leq \frac{\mathbb{E}(\beta)}{2}.$$

It then follows that “symmetric” payment providers (i.e., with similar market shares) have an excessive incentive to be compatible relative to the social optimum. The reason is that network effects render demands, in the absence of compatibility, more elastic, which decreases equilibrium prices.

This result has been established under different forms in the literature. The formulation above follows Doganoglu and Wright (2006). An alternative treatment is offered by Crémer, Rey and Tirole (2000), who assume that firms compete in capacity and prices clear demand (i.e., Cournot competition). In the latter paper, besides reducing demand elasticities, compatibility renders services more attractive relative to the outside option, thus expanding the market. In turn, Chen, Doraszelski and Harrington (2009) investigate a dynamic extension where firms’ past purchases improve the product’s quality over time (a form of data-enabled learning). Random taste shocks may however render the market asymmetric. The authors then show that the dynamic adjustment of prices helps to avoid market tipping, preserving the incentives to offer compatible services. Finally, Malueg and Schwartz (2006) consider an asymmetric market structure where a dominant firm faces a set of small competitors. The authors show that it is a dominant strategy for the small competitors to be interoperable. In equilibrium, the market may

tip towards the smaller firms, as intra-network competition assures consumers of low prices.

3.2.2 Multi-homing consumers

So far, we have seen that market tipping has the effect of rendering compatibility (more often) undesirable for the dominant operator, but socially desirable under moderate brand tastes by consumers. By contrast, market sharing often renders compatibility desirable for the competing operators (as it relaxes price competition), while sometimes being socially detrimental. As we will see next, the latter conclusion is sensitive to the assumption that consumers cannot multi-home.

Consider the equilibrium where high-valuation consumers multi-home, while low-valuation consumers single-home.⁹ Under this configuration, in the absence of compatibility, firm a 's total demand is given by

$$N_a = \frac{1}{2} + \frac{\lambda}{2} + \frac{(1-\lambda)(p_b - p_a)}{2(t - (1-\lambda)\beta^l)}, \quad (4)$$

and similarly for firm b (after changing indexes). Comparing (3) and (4), the total demand faced by firm a increases by $\frac{\lambda}{2}$ and its price sensitivity decreases. Not surprisingly, the equilibrium prices, given by

$$p_a = p_b = \left(\frac{1+\lambda}{1-\lambda} \right) (t - (1-\lambda)\beta^l), \quad (5)$$

are greater than their single-homing counterparts. Profits are also greater, as demands expand due to multi-homing.

Because multi-homing is irrelevant under compatibility, the equilibrium under compatible services is identical to that derived when consumers single-home. Accordingly, firms set prices $p_a = p_b = t$ and enjoy the profit $\frac{t}{2} - F$.

We now assess the firms' incentives for compatibility under multi-homing. It is straightforward to see that, when β^l is small, the equilibrium prices in (5) exceed t . This implies that firms would remain incompatible even if the cost of achieving compatibility was zero! By contrast, the planner may favor compatibility (provided F is positive but low) in order to save duplication in transportation costs and to generate network effects for low-valuation consumers. As can be verified after some tedious algebra, under multi-homing and for β^l low, firms have an insufficient incentive to choose compatibility. Somewhat surprisingly, the ability of consumers to multi-home, by inducing firms to avoid compatibility, may reduce network effects in equilibrium.

⁹Clearly, no equilibrium where all consumer multi-home exists. As argued by Doganoglu and Wright (2006), considering equilibria where a fraction of high-valuation consumers single-home leads to similar qualitative results.

3.3 Conclusion

The model of this section offers two explanations for why telecom/payment providers fail to be interoperable at the platform level (even when it is efficient to do so). The first applies to markets with one dominant firm, such as Kenya. The idea is that the dominant firm’s market power stems from strong network externalities, which would disappear under compatibility. The second explanation applies to markets with multiple “symmetric” firms serving multi-homing consumers. In this case, compatibility shrinks the market (as multi-homing disappears) and makes demand more elastic (as consumers joining decisions are rival), reducing prices, sales and profits. In both cases, there is a case to mandate interoperability at the platform level, which would raise welfare and consumer surplus.

It is worth noting that the theory above compares two extreme outcomes; one with full-fledged platform interoperability (in the sense that all providers have free and unfettered access to each others’ networks and consumers pay the same regardless of their share of on-net or off-net transactions), the other with no interoperability at all. Arguably, one could think of intermediate situations where firms and consumers face, respectively, interchange fees and different prices for on/off net transfers. It is expected, by Coase theorem, that one such intermediate outcome would enable firms to achieve a Pareto improvement, as they would increase efficiency while selling each other access to their respective user bases.

As we discuss in the next section, however, a new set of issues arise in such intermediate cases, in which consumers may have to perform comparisons across providers along various dimensions. These comparisons need not be straightforward, especially for less experienced consumers.

4 Interoperability with Behavioral Consumers

This section uses insights from behavioural IO to understand what should be done to ensure that platform-level interoperability “works well”.^{10,11} For example, if consumers do not understand the fees for transferring money across networks, or if those fees are difficult to compare across wallet providers, then competition between wallet providers will tend to be weak, leading to high fees and low transaction volumes.¹² As another example, if consumers anticipate that interoperability will work poorly, they may prefer to use incumbent mobile money providers that have a large share of transactions, thus inhibiting entry of new players into the market. Market interventions are then required to get the full benefits of interoperability.

¹⁰For a summary of behavioural biases, including evidence from developing countries in the context of broader financial services, see e.g., Garz, Giné, Karlan, Mazer, Sanford and Zinman (2021) .

¹¹If interoperability does not “work well”, this need not imply that consumers on different networks do not transact. They may instead transact via multi-simming (GSMA (2012*a*)) or through cash which leads to travel and coordination costs (Aron (2018)).

¹²Throughout this chapter we take costs of interoperability as fixed. Of course interoperability may impose costs on market providers, acting as another source of high consumer fees. See GSMA (2012*a*).

While we focus our discussion on interoperability at the platform level, the issues highlighted in this section can be naturally applied to other dimensions of interoperability, such as mobile network or agent interoperability, in which consumers need to compare fees and services across providers.

4.1 Fees in practice

Different mobile money providers “frame” their fees in different ways. Broadly speaking there are three different formats (CGAP (2017)). One format is so-called “slab pricing”, where transactions within a given range all have the same fixed fee (but fees differ across slabs, usually in a very regressive way). Another format is percentage-based pricing, where the fee is a fixed percentage of how much money is transacted. Finally, another format is one in which consumers pay no fees at all to transact money.

Conditional on the frame adopted, there is also the question of whether on- and off-net transactions attract different fees. Different countries have adopted different stances regarding this. For example, according to CGAP (2021), there are no restrictions on end-user fees in Australia, and only some restrictions in the Philippines; on the other hand, in Tanzania, it is forbidden to charge different fees for on- and off-net transactions, and in Peru most transactions are free for final consumers.

In turn this translates into quite different pricing models, as illustrated by the following two examples. (In both examples, the fees are correct as of November 2021.) Firstly, in Kenya, Safaricom and Airtel both use slab pricing. One difference is that they use different transaction bands. Another difference is that while Safaricom charges the same fees for on- and off-net transactions, Airtel has a zero fee for on-net transactions but a positive fee for off-net transactions.¹³ Secondly, in Uganda, MTN and Airtel both use slab pricing, with the same transactions bands, and both charge more for off- rather than on-net transactions. The difference in these fees can also be large: for example, for transactions in the top band, MTN charges 44 times and Airtel charges over 60 times more for off-net transactions than on-net transactions.¹⁴

4.2 Search costs

Consumers may, at least initially, be poorly informed about the fees that different mobile money providers charge for transferring money. Search costs—the time and effort needed to learn about those fees—can, in theory, lead to high fees and deter consumers from using mobile money services at all.

In developed country contexts, search costs are believed to be quantitatively important in a broad range of markets, including online.¹⁵ In the context of mobile money,

¹³See here for Safaricom and here for Airtel. As mentioned earlier, fees are highly regressive: on Safaricom, a transfer of 101 Shillings attracts a fee of 6 Shillings, whereas a transfer of 150,000 Shillings attracts a fee of 105 Shillings.

¹⁴See here for MTN and here for Airtel

¹⁵For a recent example of a paper using search costs to rationalize patterns of online purchase behavior, see Jolivet and Turon (2018), and for recent evidence that consumers typically search only a small number of products see Amano, Rhodes and Seiler (2021).

IPA (2021) report that 83% of surveyed mobile money customers in Uganda either did not know the fees charged by their provider, or their guess about the fees was off by more than 10 percent; only 10% of the surveyed customers exactly knew the fees. Moreover, 64% stated that they only learned how much they would pay for a transaction *after* they had carried it out, and 19% complained of unclear or unexpected charges.¹⁶

Economic theory predicts that even small costs of learning about a mobile money provider’s fees could lead to significantly relaxed competition and high fees. To illustrate this, first consider a simple case *without* interoperability. Suppose consumers single-home, and are initially uninformed about each provider’s fees for on-net transactions. According to the well-known Diamond Paradox (Diamond (1971)), if consumers must incur a (even very small) cost to learn about a provider’s fee, all providers optimally charge the monopoly fee, and the consumers who participate in the market only learn about one provider’s fee. The intuition is straightforward: since consumers expect all providers to charge a high fee, and since comparing offers from different providers is costly, consumers search only one provider, which makes that provider a *de facto* monopolist.¹⁷ Moreover, providers cannot attract additional consumers by cutting their fees, because consumers who did not search them do not observe the fee reduction.

The logic of the Diamond Paradox breaks down whenever some consumers are well-informed about fees (perhaps because they enjoy shopping around). Now providers have an incentive to price below the monopoly level, in order to attract the well-informed consumers. Stahl (1989) and Varian (1980) show that this leads to fee dispersion—with some providers charging a high fee to exploit consumers with search costs, and other providers charging a low fee to try and attract the well-informed consumers. As the fraction of well-informed consumers tends to 1 (respectively, 0), equilibrium fees converge to marginal cost (respectively, monopoly). An interesting implication of these models is that fee dispersion is not necessarily a sign of a lack of competition; under the Diamond outcome fees are not dispersed and yet they are very high, and as the fraction of well-informed consumers increases fees may fall on average even though dispersion is (over some range of parameters) increasing.

The theory models described above all assume that the product sold by firms is one-dimensional and hence has a single price. However mobile money—especially after the introduction of interoperability—is inherently multidimensional, given that it provides several different services with several different fees (such as the distinction between on- and off-net services). Nevertheless the main insights from those models carry over to this more complex setting.

Rhodes (2015) considers a setting where consumers need to pay one search cost to learn the fees associated with several goods or services provided by a firm. He shows that fees are high, as in Diamond. However he also shows that adding more goods or services—as would happen *after* interoperability, when providers add off-net transactions—leads

¹⁶IPA (2021) found qualitatively similar results for Kenya. Although consumers were better informed about fees, 72% still reported only learning how much they would pay after carrying out a transaction.

¹⁷Similarly in a context with repeated purchases—as is true with mobile money—there is no incentive to search another provider if the one used in the past did not deviate from the expected (high) fee.

to a reduction in providers' fees. Intuitively this is because adding more goods or services expands the pool of consumers who use mobile money services. Moreover, the "new" consumers must have a relatively low valuation for existing mobile money services, otherwise they would have already been using them; the presence of these "new" lower-valuation consumers makes it optimal to reduce the fees associated with the mobile money services which existed before interoperability. Moreover, when some consumers are well-informed about fees, Shelegia (2012) confirms that, even with multidimensional services, fees should be dispersed, and it is possible that any given provider has low fees on some services but high fees on other services.

We now discuss three possible remedies to high fees caused by search costs and imperfect consumer information: i) increasing competition, ii) regulatory caps on consumer fees, and iii) comparison sites and other ways to facilitate price transparency.

Economic theory suggests that, in the presence of search costs, encouraging entry can have perverse consequences. If all consumers have strictly positive search costs, as in Diamond (1971), then encouraging more providers to enter the market has *no* effect on fees. Meanwhile, if some consumers are well-informed about fees, as in Stahl (1989) and Varian (1980), encouraging entry can even lead to higher fees on average, to the detriment of those consumers who have high search costs. This happens because, facing more competition for the well-informed consumers, providers find it optimal to focus more on exploiting consumers with high search costs via high fees.

Economic theory also provides mixed guidance concerning the efficacy of fee caps. It is straightforward to see that in a setting like Diamond (1971), a cap on fees unambiguously reduces fees and therefore benefits consumers. (However consumers have to be aware of the cap, in order for it to induce more of them to search.) However Armstrong, Vickers and Zhou (2009) show that when some consumers are well-informed, fee caps can have a perverse effect, potentially reducing competition and harming consumers. The intuition is that a fee cap reduces, *ceteris paribus*, the amount of fee dispersion, reducing the incentives for consumers to become well-informed, potentially relaxing competition.¹⁸

Price comparison sites—by collecting information on fees and making it easier for consumers to compare offers—should in theory lead to lower fee and hence benefit consumers. In terms of the models outlined above, a price comparison site can be thought of as a device for increasing the fraction of well informed consumers. (See, e.g., Baye and Morgan (2001).) Price comparison tools do potentially raise some concerns surrounding increased equilibrium obfuscation by providers (e.g., Ellison and Ellison (2009)), degradations in service quality (e.g., Yang (2021)), or higher fees once the commissions charged by the comparison site are taken into account (e.g., Ronayne (2021)). Nevertheless, in the context of mobile money, these issues seem more second-order, especially if the comparison site is run by, for example, a regulator.¹⁹

¹⁸Despite this, using a structural model Galenianos and Gavazza (2021) estimate that a cap on interest rates for US credit cards would reduce market power and benefit consumers.

¹⁹Much in this spirit, Annan (2020) provides evidence from Ghana that when mobile money consumers have better information on what fees they should be charged by cash agents, the fees they pay decrease by about 40%, leading to a substantial increase in volumes traded.

4.3 Complexity and Framing of fees

Even if consumers search for information about fees, or acquire such information from comparison sites, they still need to be able to understand those fees in order to make a well-informed choice. When mobile money products become interoperable, they also become more complex—because (absent any regulation to the contrary) each provider can now choose two sets of fees, one for on-net transactions and another for off-net transactions. This complexity may then confuse consumers. As Grubb (2015) writes, “when prices ... are vectors rather than scalars, consumers have trouble choosing the best price”.

There is substantial evidence from developing country contexts that consumers make poor choices when products are (or are framed as being) complex. For example, Wilson and Waddams Price (2010) examined consumer switching decisions in the UK electricity market, where providers offered two- and in some cases three-part tariffs. Focusing on consumers who stated they had switched provider to reduce their electricity bill, they found—based on consumers’ expected consumption—that at most 20% of consumers had switched to the best available deal in the market, whilst as many as one-third of consumers had actually switched to a *more expensive* supplier.²⁰

Complexity as a barrier to good choice has also been documented in the laboratory. For example, Kalaycı and Potters (2011) run an experiment in which sellers of a product choose how many “dimensions” it has, and buyers have to choose which product to buy. (In the context of mobile money, one might think of there being a single dimension prior to interoperability, and two dimensions once interoperability comes into effect.) They find that when sellers choose more dimensions, buyers are more likely to make suboptimal purchase decisions, and sellers exploit this by charging higher prices. Meanwhile Huck and Wallace (2015) run an experiment where sellers with homogeneous products can present their prices using different frames, and show that relative to simple linear pricing this leads to worse decision making by buyers.

Although we are not aware of empirical evidence on the impact of complexity on mobile money choices, Gine, Martinez Cuellar and Mazer (2017) use laboratory experiments in Mexico and Peru to study complexity of loans and savings products. They examined how consumer behavior changed when, instead of being presented with marketing materials from firms, consumers were presented with simplified statements with “key facts” about the financial products. They found that by framing the products in a more comparable way, the probability of choosing the cheapest loan increased from 42% to 65%, and the elasticity of demand for credit tripled.

Economic theory suggests that firms indeed have incentives to use complexity as a way to exploit consumers. Carlin (2009) considers a model where firms sell a homogeneous product, and choose both what price to charge and how much complexity to use when conveying the price. Some fraction of consumers get confused and choose randomly, while the remaining consumers choose the best price; the fraction of confused

²⁰More strikingly still, Wilson and Waddams Price (2010) report that many of the consumers who switched to a worse deal, actually switched to a supplier who was dominated by their current one for any conceivable level of electricity consumption.

consumers is increasing in the level of complexity chosen by each firm. In equilibrium prices are dispersed—some firms choose a low price and little complexity, while others choose a high price and as much complexity as possible. An increase in the feasible amount of complexity—as might arise once interoperability makes the set of possible fee structures more complex—leads to more confusion and higher prices. Encouraging more competition can have perverse effects because, similar to the search models considered earlier, when competition increases more firms choose complex prices so as to target confused consumers.

Piccione and Spiegler (2012) and Chioveanu and Zhou (2013) then provide a micro-foundation for complexity via the use of frames. In the context of mobile money, one frame would be presenting fees using “slab” pricing, while another would be presenting fees as a percentage of the transaction value. Interoperability increases the potential for complexity—if a provider chooses different fees for on- and off-net transactions, that makes it harder for a consumer to evaluate the provider’s offer, and compare it with other providers’ offers (independent of what frames each provider is using).

Although Piccione and Spiegler (2012) and Chioveanu and Zhou (2013) make different assumptions on how different frame choices interact, they both find that sellers of a homogeneous good will generally randomize over different frames, leading to some consumers getting confused, which is then exploited via dispersed (and supra-competitive) prices. One important insight from the second paper is that increased competition—for example, by encouraging more mobile money providers to enter—may harm consumers, by encouraging firms to make their frames more complex to soften competition.²¹ One important insight from the first paper is that policy which makes fees somewhat easier to compare—for example requiring all providers to use slab pricing—may backfire and harm consumers when providers have other ways to make fees complex—for example, by choosing different on- and off-net fees.

Turning to policy, all the theory models outlined above (implicitly) assume a cap on fees, otherwise in their (admittedly static) environments firms would have an incentive to charge very high fees in order to exploit confused consumers. One lesson is that, absent any regulation on how fees should be presented to consumers, encouraging more competition between providers may backfire. Another lesson is that partial restrictions on how fees are presented to consumers may again backfire, or at best be only partially effective, as providers seek greater complexity in other ways. Broadly speaking, these models suggest that in order for consumers to make the best choices possible under interoperability, regulators should mandate that fees be easily comparable, possibly requiring on- and off-net fees to be the same.

²¹As Gabaix, Laibson, Li, Li, Resnick and de Vries (2016) point out, a more reduced-form way to model complexity is to assume that consumers make small errors when evaluating different products, leading them to act as in a discrete choice model. Depending on the distribution of the errors, greater competition can again increase prices; for regular noise distributions prices fall with competition, but the authors show that prices may fall only very slowly.

4.4 Add-on pricing

Another way to analyze the impact of interoperability on consumer fees is through the lens of add-on pricing theory. Two seminal papers in this literature are Gabaix and Laibson (2006) and Ellison (2005). In Gabaix and Laibson (2006) firms provide a base good and an add-on good. These two goods are naturally bundled together, such that consumers buy from (at most) one firm. Firms choose a price for each good upfront. The price of the base good is salient, and accounted for by all consumers. However firms can, if they wish, “shroud” their add-on price, meaning that consumers cannot observe it when choosing which firm to buy from. Some consumers are myopic: if each firm shrouds then these consumers neglect add-on prices when deciding which firm to buy from. However, if one or more firm unshrouds, this educates (some) myopic consumers, causing them to account for add-on prices. Other consumers are sophisticated: they account for add-ons, and where appropriate form an expectation about their price. Sophisticated consumers can also exert effort in order to avoid paying the add-on charges if they are (expected to be) excessive.

Gabaix and Laibson (2006) show that if there are enough myopic consumers, all firms choose to shroud. Add-on prices end up being very high, such that sophisticated consumers engage in socially costly activities to avoid paying them. Nevertheless overall profits need not be high—because competition for myopic consumers, and the high add-on prices they will pay later on, can lead to low (and even below-cost) prices for the base good. Intuitively, no firm has an incentive to unshroud, because in doing so it would simply induce the educated myopic consumers to buy from the firm with the lowest base good price and take actions to avoid ever paying the add-on price—in particular, unless the unshrouding fee drastically changes its own prices, it will not be able to capitalize on these newly educated consumers.

In mapping this model to mobile money, one can think of a provider’s telephony and basic wallet services as its base good, and transactions (both on- and off-net) as its add-on good. Sophisticated consumers could multi-sim or use cash whenever (especially off-net) transaction fees are too high. There are some reasons to believe that interoperability could *increase* incentives for providers to shroud. Firstly, since interoperability creates a new stream of fees (for off-net transactions) this increases the potential revenue from add-on charges, leading to yet fiercer competition for the telephony/wallet services; this in turn makes it even less attractive for a firm to unshroud and win new demand. Secondly, the model just presented assumes that all firms are symmetric. In practice, for historical reasons, some telephony providers will have larger market shares than others. Absent interoperability, one might expect that newer and smaller operators have greater incentives to unshroud as a way to get a foothold in the market. However, with interoperability, a smaller operator may now be able to collect lots of add-on fees from its customers who transfer money to the larger operators, thereby reducing its incentives to unshroud.

Alan, Cemalcilar, Karlan and Zinman (2018) provide some empirical evidence of the relevance of the unshrouding effect described above. Specifically, in collaboration with a large Turkish bank, they found that promotional text messages offering a 50% discount

off the usual APR for going overdrawn *reduced* overdraft usage by around 1.2 percentage points (from a baseline of around 30%). On the other hand, a message about overdrafts which did not draw attention to their cost increased uptake by around 0.9 percentage points.

An interesting feature of Gabaix and Laibson (2006) is that firms may end up earning low profits because they largely compete away the rents from add-ons through a low base price. A Chicago school critique would then be that high transaction fees—whether due to add-on pricing, or search costs, or framing—are ultimately passed through to consumers, suggesting they are not harmful.²² However there are at least three possible issues with this critique. Firstly, as we describe shortly, Ellison (2005) argues that high transaction fees may not be passed through very much. Secondly, even if pass-through does occur, shrouding shifts rents towards sophisticated consumers and away from naive consumers—who may be poorer. Thirdly, since sophisticated consumers react to high transaction fees by avoiding them, this is socially inefficient. Relatedly, even naive consumers may transact less when they eventually learn that fees are high.

Ellison (2005) casts doubt on the idea that high add-on fees will be passed through to consumers in the form of a lower base good price. To see why, suppose that sophisticated consumers are more price-sensitive on the base good than are myopic consumers. (This is reasonable if, for example, sophisticated consumers are less influenced by providers' attempts to appear more differentiated.) If a firm reduces its base good price it will attract more consumers, but relatively many of those consumers will be sophisticated. Since sophisticated consumers do not buy add-ons, they are less profitable, and hence the incentive to undercut on the base good is reduced. In equilibrium this leads to only partial passthrough of high add-on fees.

Finally, Gabaix and Laibson (2006) suggest several different possible regulatory interventions in markets with add-on pricing. In the context of mobile money, three are of particular interest. Firstly, mobile money providers could be compelled to disclose fees. However as we saw earlier, if providers can frame fees as they wish, this may not help resolve the problem of high fees. Secondly, consumers might receive a warning about fees when signing up to a wallet. However, as we also saw earlier, with search costs, a mere warning may not induce consumers to search and force providers to price more competitively. Thirdly, then, some form of caps may be required.

4.5 Conclusion

In order for interoperability to work well, fees need to be set in a clear and transparent manner, and at levels which make it attractive for consumers to transfer money both on- and off-net. In this section we reviewed some reasons why fees may end up being high. One reason is imperfect consumer information and search costs, which give firms monopoly power over at least a fraction of the market. Another reason is complexity due

²²In the case of mobile money, GSMA (2012b) argues that even without the fees wallet services can enhance profitability of the base good (mobile telephony), by reducing churn in the mobile telephony market, and by reducing distribution costs of airtime since consumers can buy airtime directly using their wallet.

to the way that providers frame or present their fees, making it difficult for consumers to compare offers, and ultimately leading to relaxed competition on fees. A final reason is shrouded pricing, whereby consumers do not fully account for fees when choosing a mobile money provider, and providers have little incentive to “educate” consumers about their fees. We also argued that interoperability—by making products more complex, and by expanding the scope of possible add-on fees—could potentially further weaken competition over fees. We then surveyed possible policy interventions, such as fee caps, comparison sites, and attempts to increase competition via entry. A common theme was that, depending on the particular cause of a lack of competition, well-intentioned policy interventions could potentially backfire. For example, encouraging entry (without imposing any other regulation) could lead to higher fees, because providers focus more on capturing consumers with high search costs, or seek to offset the increased competitive pressure by making their fees more complex and hence less comparable.

5 Agent Interoperability

One of the key features of digital payments and finance in developing countries is the prevalence of agent-based banking and mobile money. In most developing countries, cash remains the most preferred means of transaction and traditionally, individuals rely on ATMs and bank branches for obtaining cash. In rural areas with low population density, it is often economically unviable for banks to have many branches or ATMs. Even consumers that use digital financial services like mobile money might need to rely on physical infrastructure to deposit and withdraw cash. The agent-based model of mobile money and payment platforms addresses precisely this gap in physical infrastructure.

Typically, mobile money agents facilitate cash-in/cash-out services in areas underserved by traditional banking. Apart from cash-in/cash-out, mobile money agents provide services like person to person transfers, recharges for mobile airtime and bill payments. The mobile money provider’s main source of revenue is the fee charged to the consumer for every transaction, with the agent receiving a part of this fee. The agent may receive training from the mobile money provider but typically bears the cost of operating herself.

There can be many different forms of the relationship between agents and providers, leading to different types and extent of interoperability. Depending on the country, agents can either be exclusive to one mobile money provider or represent several providers. For example, an interoperable agent might provide cash-in services to customers of other providers with her agent account, but not cash out. An agent could also represent several companies but only by having multiple agent accounts (non-exclusivity). There could also be a full-interoperability regime where an agent can provide all its services to the customers of another service provider.

Interoperability of mobile agents is a contested policy issue. While investment in agent networks is the main source of product differentiation for mobile money providers in many markets, having interoperable agents might further digital financial inclusion. Even when there is agreement about having interoperable agents, it is not obvious when

it should happen. Interoperability of agents can have different consequences for welfare depending on whether it is introduced in a mature market or in a nascent one.

Nevertheless, there is sufficient evidence on the benefits of a large agent network (which may or may not be as a result of mandating agent interoperability). Suri and Jack (2016) find that the key driver of adoption of mobile money services in Kenya are agent networks. In Niger, Aker et al. (2020) finds that despite a relatively high demand for mobile money services, adoption remains limited because agent networks are sparse.

In this section, we describe economic frameworks to think about interoperability of mobile money agents using first, the experience of interoperability for ATMs, and then the theory of access pricing and co-investment in network industries.

5.1 Interoperability of ATMs

In New York, ATMs were introduced by Citibank in the 1970s (CGAP (2019)). Even though it would have been cheaper to introduce ATMs in consortium with other banks, Citibank chose to differentiate itself through the ATM technology and expand its market. It only became interoperable with ATMs of other banks ten years later after it started losing market share to a consortium of six banks with interoperable ATMs (CGAP (2019)). Thus, the decision of the market leader (Citi) to allow interoperability came as a response to a competitive market.

The experience of interoperability for ATMs has been different in different jurisdictions. In most parts of the world, ATMs are now interoperable with a pre-negotiated interchange fee between banks and/or a surcharge to use another bank's ATM. It is possible to draw parallels from this system to the case of interoperability of mobile money agents. We can think of agents as ATMs and mobile money service providers as banks. To have interoperable agents, mobile money providers would need to agree on the terms to do so. This could be done by setting an interchange fee as in the case of ATMs and debit/credit cards. If a consumer chooses to transact with a "foreign" agent (Donze and Dubec (2009)) affiliated to a different provider (foreign provider) than her own, then the customer's provider (home provider) would pay an interchange fee to the foreign provider. This fee would compensate the cost faced by the foreign provider for providing services to a foreign customer. The home provider may choose to charge a part of this fee to its customer (akin to foreign fees for using another bank's ATM). The foreign provider may also charge the foreign customer for the transaction (commonly called surcharges in the case of ATMs). The key question would then be how providers set these fees and if there is a need to regulate them. These fees have a direct impact on consumer welfare by affecting consumer search for the most economical ATM (or agent, in our case). They can also have an indirect effect on consumer welfare by affecting the number of ATMs deployed, the price of transactions themselves, and the level of competition between service providers.

The existing economics literature has tried to address several contentious policy issues in ATM pricing. Though most of the examples of these policy issues come from the context of ATM deployment in a mature market like the US, the theoretical insights can inform the discussion of interoperability of mobile money agents in the developing world.

The first policy issue, broadly defined, concerns the setting of the interchange fees. Introducing interchange fees can be an incentive for providers to share their network (or have interoperability), however, the ultimate welfare outcomes depend on how providers set these fees. Empirical evidence for ATMs from the UK, Australia and the US has shown that interchange fees were set by banks in a way that led to mark-ups of close to 100%, meaning that these fees were much higher than the cost of providing ATM transactions. Several theoretical models rationalize this empirical finding by showing that providers cooperate to set the interchange fee and this amounts to tacit collusion (McAndrews (2003)). A high interchange fee softens competition for end consumers but increases competition for processing foreign transactions (Donze and Dubec (2006)). This makes it possible for the interchange fee to be used as a device for collusion by providers. This is widely recognized by regulators in the ATM and electronic card industries, and in these industries it is not uncommon to have oversight on how interchange fees are set.

Even more contentious than the setting of the interchange fee has been the debate around ATM surcharges (the fees charged by the ATM owner bank to a foreign customer). Proponents of surcharging argued that it was necessary to sustain ATMs in remote, and otherwise unprofitable locations, thus expanding access to cash withdrawal services (McAndrews (2003)). On the other hand, two main concerns against the practice of surcharging were highlighted. First, surcharging a foreign customer by the ATM owner added to the foreign fee already paid by the customer to its home bank, making the fee structure complicated for consumers (McAndrews (2003)). Second, consumers would be inclined to switch to banks with a larger network of ATMs to avoid paying these fees, thus reducing competition in the market (McAndrews (2003)). Some concerns were also raised that surcharging amounted to consumers paying a fee to access their own money (Hannan, Kiser, Prager and McAndrews (2003)). The issue of surcharging became prominent in public debates in the US in 1996 when state governments and courts made illegal any rules imposed by service providers to disallow surcharging.

Following this ruling, ATM deployment dramatically increased in the US, as surcharging was a new incentive for banks to provide this service. Theoretical literature rationalizes this empirical finding. Donze and Dubec (2009) provide a model of two horizontally differentiated banks that first choose the interchange fee jointly and then provide ATMs, competing for customers non-cooperatively. A key parameter of their model is the size of the shopping space (or the dispersion parameter) - in a larger shopping space, consumers value the deployment of an additional ATM more than in a smaller space. After setting up the model, they compare welfare outcome under three regimes: i) consumers can access all ATMs (home and foreign) free of cost, ii) consumers pay a foreign fee to their home bank for using a foreign ATM, iii) consumers pay a foreign fee to their home bank and a surcharge to the foreign bank when using a foreign ATM.

The first set of results from this work shed light on the size of the ATM network under each of the three regimes. Under the first regime, the consumer views all ATMs as identical and the only incentive for banks to deploy more ATMs is to generate revenue through the interchange fee. Under the second regime, since consumers have to pay a foreign fee for using a foreign ATM, they are attracted to banks that have a larger

network of ATMs. Banks thus have an additional incentive to deploy more ATMs as they attract new customers through a larger network of ATMs. The incentive to deploy ATMs is even bigger under the third regime, as surcharging increases a foreign banks revenue from a foreign customer. Thus, regime three leads to the most number of ATMs deployed.

On pricing and setting the three fees, the paper presents an interesting set of results. Surprisingly, the paper finds that consumer welfare is higher in regime two (foreign fees) than in regime one (free ATM usage for home and foreign ATMs). The economic intuition driving the result is the following: under regime one, since all ATMs are identical for consumers, they make many foreign transactions for no additional fee which generates a large gross surplus. The banks take this into account in pricing the end product (deposits) and is able to extract the surplus by setting high account fees. Under regime two, consumers make fewer foreign transactions and ultimately pay lesser to the bank.

Even more striking is the result on regulation of surcharges. The authors find that when surcharging is banned, the profits of the bank in equilibrium depend on the interchange fee and banks have the incentive to collude to set this fee. Under regime three, however, the authors show that the interchange fee cannot be used to collude by banks. Further, they evaluate consumer welfare while allowing banks to jointly set the interchange fee. They find that consumers prefer regime three to regime two when the shopping space is large and the value of having more ATMs is high. When the shopping space is relatively smaller, consumers do not benefit from paying the additional surcharge as accessing ATMs is easier than in a bigger shopping space. This result is consistent with empirical evidence from the US. In 1996, the US changed from regime two to regime three and Knittel and Stango (2008) find that in areas with high travel cost (where the shopping space is large) consumer welfare improved after the regime change. The opposite was true in areas of low travel cost.

The empirical literature on ATM pricing apart from Knittel and Stango (2008) is limited, but illustrative. Hannan et al. (2003) find that the probability of surcharging increases with the bank's market share of ATMs and decreases with increasing density of the ATM network. They also find evidence that banks use surcharges to attract new customers to their own networks, but do not find any evidence that big banks use surcharges to attract existing customers of smaller, local banks.

To conclude, discussions on interoperability of mobile money agent can be informed by the experience of interoperability for ATMs, and particularly by the literature on ATM pricing. While this literature still has many open ends, some important theoretical findings presented here are consistent with empirical evidence on ATM pricing.

5.2 Agent Interoperability in Network Industry Framework

Another way to think about interoperability of mobile money agents is to use the framework of network industries. In network industries like telecommunications, providers face high investment cost to expand access. In our case, training agents and expanding the agent network is a considerable investment for mobile money providers (though considerably less than the case of telecommunications infrastructure). For a provider

that has invested in a large network of agents, introducing agent interoperability would mean that a part of the gains from its investment are recovered by competing providers. If mobile money providers expect that agent interoperability might be mandated in the future, they would have fewer incentives to invest in the agent network. On the other hand, once one provider has made investments in the agent network, other providers may free-ride on this investment in the case where agent interoperability is mandated.

Bourreau, Cambini and Hoernig (2018) and Bourreau, Cambini, Hoernig and Vogel-sang (2021) discuss a similar problem taking the general case of network industries. In any network industry (telecom, mobile money etc), an under-investment problem may arise because new infrastructure is costly and the gains from the infrastructure will likely be shared by other firms and their consumers. This under-investment problem is made worse when there is uncertainty about the level of demand for the firm's service.

Typically, in industries like telecommunications, to expand coverage, regulation mandates incumbents to have access obligations. For example, in the Europe, during the initial roll-out of broadband infrastructure, incumbent operators were mandated to share their infrastructure with other providers under these access obligations. The other providers pay an access fee to the incumbent, which may or may not be regulated. Though mandatory access regulations have prevented incumbents from exerting monopoly power, they might reduce the incentives to create new investments (Bourreau et al. (2018)).

A recent regulatory alternative to access obligations (in order to expand coverage of services) is allowing, encouraging or mandating "co-investment" in infrastructure (Bourreau et al. (2018)). Co-investment, instead mandating access to competing providers, works as a way to share the cost of infrastructure by different service providers. Bourreau et al. (2018) provide a theoretical comparison of the access pricing regime and the co-investment regime in terms of spurring coverage, and in terms of total welfare.

The authors consider the case of a network industry where providers face uncertain demand and might have incentives to co-invest. In their framework, they consider a country with several areas. The areas have the same distribution of uncertain demand but differ in the sunk cost required for deploying new infrastructure. There are two firms, one of which is the incumbent in the market and the other a potential entrant. The incumbent chooses the areas to invest in and the entrant chooses whether to co-invest or to pay an access fee. After co-investment or paying an access fee, the entrant can access the shared infrastructure in the same way as the incumbent, and at no additional cost. In our case, this translates to the following: A mobile money provider decides to invest in expanding the agent networks in certain areas. Once this investment takes place, another provider that wants to enter the market can co-invest in the agent network by taking up half of the cost of the investment incurred by the incumbent firm. It then has access to these (interoperable) agents at no extra cost.

The authors compare three different scenarios. In the first "pure access" scenario, there is no possibility of co-investment. The entrant can ask for access to the incumbent's infrastructure by paying an access fee (linear, and fixed by the incumbent). The second scenario that they consider is of "pure co-investment" where the entrant can share the

cost on investment in areas that the incumbent has planned to invest in. Crucially, in this scenario, the entrant cannot ask for access to the incumbent's infrastructure as an alternative to co-investment. The final scenario that the authors consider is a combination of the first two: the entrant is allowed to decide whether to pay an access fee to the incumbent or to co-invest.

The authors find that compared to the first scenario of pure access, the pure co-investment regime fosters competition in areas where the network is shared, and also leads to greater coverage. Although in the pure co-investment regime, the authors find that there will be areas where the incumbent can set monopoly prices, most of these areas would not have been covered at all under the pure access regime. Under the mixed regime, the authors find that both total coverage and co-investment coverage is lower than in the pure co-investment scenario. The intuition behind this result is that having the option of paying an access fee reduces the entrant's incentives to take up co-investment. The authors conclude that in more realistic scenarios with a high degree of uncertainty of demand, the pure co-investment regime is the most preferred regulatory regime. Finally, they also provide a comparison of the pure co-investment regime with other variations of the access-pricing regime and find that the pure co-investment leads to greater welfare in every scenario. This result is attributed to the higher intensity of competition and greater incentives to invest under the co-investment regime.

The next crucial question to consider is the timing of the co-investment. Regulators may allow firms to "co-invest" in infrastructure deployment ex-post, meaning that entrants can wait until any uncertainty of demand is resolved and then decide to co-invest. The importance of demand uncertainty and the timing of co-investment is explored in Bourreau et al. (2021). The set up of the theoretical model is largely similar to Bourreau et al. (2018). After the investment stage, the two firms compete with each other and set the prices of their services. In areas where the entrant and the incumbent both co-invest, they each make duopoly profit. In areas where only the incumbent operates, it makes monopoly profits or less. In all cases that the authors consider, the information about the level of demand is only revealed after investment takes place.

The first case that Bourreau et al. (2021) considers (the benchmark case) is where the entrant makes its co-investment decision before the deployment of the infrastructure and without information about the level of demand in the area. This means that both the entrant and the incumbent face the same demand uncertainty. In this case, the authors find that in equilibrium, firms choose to co-invest only in those areas of the market where the sunk cost of investment is low. In this equilibrium, there might also be areas where only the incumbent invests and gets monopoly profits (under the condition that duopoly profits are lower).

The authors contrast this case of ex-ante co-investment with the case where the entrant has the option to wait for the demand to be realized and then co-invest. In practice, this case is more likely to occur than the first one. In this case, the authors highlight two distortions that arise: market structure distortion and coverage distortion. First, the probability of entry in co-investment areas decreases if the entrant knows the level of demand before the co-investment decision. At the same time, the probability

of entry in monopoly areas increases. This shift in probabilities is called the market structure distortion. Second, in this case, the incumbent’s incentive to invest are reduced and this translates to lower coverage of the end service compared to the benchmark case. This is called the coverage distortion by the authors. Coverage distortion is especially likely to arise when there is a high degree of demand uncertainty and decreases social welfare compared to the benchmark case.

These results can be applied to the mobile money industry when thinking about agent interoperability. If co-investment is proposed as a way to ensure agent interoperability, the timing of an entrant’s co-investment is crucial. In case the entrant can observe the demand for mobile money before co-investing in areas served by the incumbent, the incumbent has fewer incentives to invest in the first place. This leads to lower coverage of the service, and lower social welfare.

Bourreau et al. (2021) suggest two remedies that partially address these distortions. The first one is a co-investment option that the entrant purchases from the incumbent ex-ante (that is before the demand is realized) and it can be exercised ex-post. The authors find that this first remedy cannot correct the market structure distortion since it does not affect the entrant’s co-investment decisions in areas already covered by the incumbent. However, the co-investment option is able to reduce the coverage distortion by increasing the incumbent’s incentives to invest.

The second remedy suggested by the authors is a risk premium paid by the entrant to the incumbent ex-post. The authors find that a positive risk premium increases total coverage but reduces the entrant’s willingness to co-invest. However, this does not address the market structure distortion. Only a negative risk premium (i.e. a payment/subsidy from the incumbent to the entrant, which might be unlikely in practice) can reduce the market structure distortion.

For the coverage distortion, the authors suggest a combination of the two remedies - a co-investment option for areas that have intermediate costs of investment and a risk premium for areas that more costly or hard to service. They show that this combination of remedies can increase total welfare.

5.3 Conclusion

For discussions about interoperability of mobile money agents, pricing is a very important dimension to consider. In this section, we present two analytical frameworks to analyze interoperability of mobile money agents. For the first, we draw lessons from the experience of interoperability of ATMs and focus on three types of fees - interchange fees, foreign fees and surcharges. The ATM literature finds that the interchange fee can be used as a collusive device, and may need regulatory oversight. On the other hand, surcharges might be useful in preventing providers from using interchange fees as a collusive device. The number of ATMs deployed is the highest when there are both surcharges and foreign fees. Notably, these results depend on density of the ATM network (or the size of the shopping space).

Next, we focus on understanding providers’ incentives to invest in agent networks when agent interoperability might be possible. We view agent interoperability as a

parallel of the problem of investment incentives in network industries. We provide a review of the latest findings on co-investment, a new strategy recommended by the European Union to facilitate access to the next generation of digital technologies. Recent literature suggests that co-investment may be better regulatory strategy in terms of social welfare and investment incentives than the traditional regulatory tools based on access pricing. Further, the timing of co-investment is crucial, especially in the presence of a high degree of demand uncertainty. In case entrants can wait until the demand uncertainty is resolved before co-investing, a market structure distortion and a coverage distortion is likely to arise. In this case, to expand the coverage of mobile money services through a larger (and interoperable) network of agents, a combination of a co-investment option (in areas with intermediate costs of investment) and a risk premium (in areas with high cost of investment) can be used.

6 Data Interoperability

Digital payments, unlike cash payments, only exist to the extent they are recorded. Therefore one side product of digital payments is data on the consumption patterns and financial decisions of users. This information is typically relevant to evaluate risks that are key in the provision of financial products such as credit or insurance. This suggests that digital payments may not only expand access to basic (yet important) financial services such as storage and transfer of money but can also be the gateway to more sophisticated products where risk evaluation is key. We understand data interoperability here as the extent to which the information generated with one payment provider can be transferred to a third party.

6.1 The value of payment data

Digital payment generates “hard information” in the sense that it is verifiable, easy to store and to a large extent quantitative. A first implication is that payment data can provide a cost-effective resource to assess the individual risk of a user: the collection of payment data and its processing and transformation into risk metrics can be standardized and automated. This can eliminate some of the interactions between loan officers and borrowers, and more generally the need for physical proximity between borrowers and lenders through a costly network of branches. A stream of empirical literature in banking uses that physical proximity as a proxy for information frictions: when hard information is not available, banks need local loan officers with sustained relationships with potential borrowers in order to evaluate credit-worthiness. Petersen and Rajan (2002) attribute the steady growth of the physical distance between lenders and small business borrowers in the US to an increased reliance on hard information and information technologies. The wider availability of payment data and the implied productivity gains could be particularly meaningful in the context of developing countries: Mian (2006) shows that foreign banks that open subsidiaries in Pakistan are willing to lend at arm’s length when hard information is available, but are reluctant to do so when the lending decision

requires unstructured “soft information” collected through repeated relationship.

Another dimension of information being hard is that it can be used and interpreted independently from the person or entity that collected it unlike, for instance, the subjective assessment of a loan officer (Liberti and Petersen (2019)). This feature is important for data interoperability because it makes payment information portable: it can be exploited by any third party provided that party can access it. In that sense, data interoperability can be distinct from other forms of payment interoperability: interoperable payment providers inherently share data, but payment data can be portable across firms with no existing connection. In practice, digital infrastructures such as Application Programming Interfaces (APIs) can facilitate the interconnection between payment providers and third parties and allow for faster and cheaper automated processing of data, but the mere fact that users can provide banking statements is a form of data sharing.

Ghosh, Vallee and Zeng (2021) show how the transfer of payment information can have a downstream effect on the provision of credit. In the theoretical part of the paper, they highlight two channels through which the additional signal provided by payment data affects lending outcomes. First, additional information improves the borrower’s ability to screen lenders. While this may result in a more efficient allocation of credit, enhanced screening also has differential effects across loan applicants: better-than-average lenders (given the extra signal) are more likely to receive funding while worse-than average lenders are less likely to receive funding. This result echoes the more general finding that improving lenders’ screening ability through information technologies has distributional effects. For instance, Fuster, Goldsmith-Pinkham, Ramadorai and Walther (Forthcoming) show that the use of machine learning algorithms in the US mortgage market is less likely to benefit Black and Hispanic minorities than the White majority group. Second, additional information reduces the uncertainty lenders face when originating a loan. If lenders are risk-averse (more generally, subject to financial constraints), lower uncertainty reduces the required return for lenders, hence the cost of funding across the board for borrowers. Next, Ghosh et al. (2021) use data from a FinTech lender in India to confirm the prediction of the model. They show that borrowers’ ability to provide digital payment data to the lender is associated with better lending outcomes for small businesses: an interquartile increase in the share of cashless payment increases the likelihood of obtaining a loan by 2.7%, i.e., more than 10% of the baseline likelihood of getting approved, and a 42 basis point decrease in interest rate. This relationship is stronger for internet transfers that provide richer information about counterparts than for mobile payment. The paper exploits an exogenous variation to the adoption of digital payment (the 2016 Indian demonetization) to show that the relationship between the use digital payment and credit outcomes is likely causal.

6.2 Who should have control over payment data?

The Indian case studied in Ghosh et al. (2021) highlights how data interoperability is distinct from other forms of payment interoperability that require pooling resources and investments across players (e.g., payment networks, agent networks) or coordination on

compatible technical solutions. On the contrary, data interoperability merely requires that data be portable from one firm to another firm. This suggests that obstacles to data interoperability may not lie so much in the cost of building common infrastructures as in the incentives to share data. This naturally raises the question of who should have control over data, and hence the ability to share it.

Jones and Tonetti (2020) provide a general framework that supports consumers' ownership of data. In their model, the distinctive feature of data relative to other factors of production is that it is non-rival: data is costlessly replicable and can therefore be used infinitely many times with no opportunity cost. This suggests there are significant economic gains to data being used broadly. The use of payment data for credit allocation discussed earlier is consistent with meaningful efficiency gains from transferring data from one service to a different service. Jones and Tonetti (2020) show there can be quantitatively significant under-sharing of data if firms control it and are concerned that access to data makes them more likely to be displaced by competitors. By contrast, consumer ownership implements a higher degree of data sharing and comes closer to the social optimum. Key to this result is the assumption that consumers cannot commit not to share data in the future, i.e., they cannot enter into an exclusive data sharing agreement with one particular firm.

Some of the regulatory frameworks and public infrastructures around payment data favor consumer's control over payment data, consistent with Jones and Tonetti (2020). For instance, PSD2 in Europe creates a mandate for banks to give third parties such as tech firms access to payment history through APIs, upon consent by the consumer. This effectively implements data interoperability but also a deeper form of integration as these third parties could use APIs to initiate payments through the banking system infrastructure. Note that tech firms are not subject to a reciprocal data sharing mandate with banks (Carstens, Claessens, Restoy and Shin (2021)). In India, the central bank created a legal framework for regulated entities ("account aggregators") that can collect and consolidate information from multiple financial institutions as well tax authorities and transmit it, with the customer consent to a third party (e.g., banks, insurance companies). Making data available through accounts aggregators is optional for financial institutions but firms that do not provide access are also precluded from using the information provided by accounts aggregators (D'Silva, Filková, Packer and Tiwari (2019)). In both cases, one rationale for giving users control over personal data is to facilitate access to other financial services, such as credit. This is different from Jones and Tonetti (2020) or Dosis and Sand-Zantman (2019) where users' control over personal data allows them to monetize it.

6.3 Payment data production

Data are different from other factors of production in that it is to a large extent a by-product of production. This has several consequences. First it can create increasing returns to scale economy-wide whereby higher output and consumption generates more data which improves productivity, hence leads to higher output, more data etc. In Jones and Tonetti (2020), this occurs within the firm, which takes advantage of data generated

by its own production. In addition, because data is non-rival, returns to scale can be magnified at the economy level if the structure of the market for data incentivizes data sharing across firms.²³ Second, firms may alter their production decisions to generate more data. In Farboodi and Veldkamp (2021), firms offer products at lower prices (even possibly at zero-prices) because they value access to the data generated by these sales. This general insight suggests that incentives to invest in payment infrastructures and the pricing of payment services should be sensitive to the payment provider’s ability to exploit payment data for other services and to its ability to retain control over this data and monetize it.

Parlour, Rajan and Zhu (2020) provide some insights into the relationship between the pricing of payment services, the production of payment data and consumers’ ability to transfer data. They model a payment market where a legacy payment provider, a monopolistic bank, faces the entry of a new payment provider, a FinTech firm. First, entry has a competitive effect: lower prices for payment allow new users to access digital payments, and induces some users to switch from the bank to the FinTech firm. Users who remain with the bank (those who have a strong “taste” preference for the bank over the FinTech firm) may however face higher prices for payment. In addition, because payment data is used by the bank to assess credit risk, the quality of the bank’s loans declines as it loses access to part of the payment market. Introducing data portability has an upside and a downside for the bank. On the one hand access to payment data produced by the FinTech firm improves the bank’s ability to screen borrowers in the credit market. On the other hand, consumers are less inclined to use the bank’s payment services to generate information and improve their access to credit. Overall, Parlour et al. (2020) argue that the development of digital payment providers outside the banking system and the interoperability of data between these new players and banks can have significant impacts not only on the payment activity of banks, but also on their credit activity and ultimately on their stability. Another takeaway is that the interoperability of payment data is likely to affect the pricing of payment services.

6.4 Payment data sharing and credit markets

As Ghosh et al. (2021) and Parlour et al. (2020) suggest, payment data can be particularly effective in credit markets. The question that the interoperability of data then raises is the impact of making this information available across lenders. The extant literature in banking has studied the impact of information sharing mechanisms such as credit bureaus. Information sharing has the potential to alleviate the adverse selection problem that lenders face in credit markets and can endogenously arise (absent regulation) when competition between banks is not too high (Pagano and Jappelli (1993)). Jappelli and Pagano (2002) and Djankov, McLiesh and Shleifer (2007) provide cross-country evidence that data sharing leads to higher aggregate bank lending and lower

²³Farboodi and Veldkamp (2021) note that when data is used to improve the precision of predictions, as would be the case with payment data for credit or insurance applications, there is a natural limit to returns to scale since forecast errors cannot be reduced beyond zero.

credit risk. Liberman, Neilson, Opazo and Zimmerman (2018) try to assess the equilibrium effects of restricting information sharing. They use a change in the regulation of credit bureaus in Chile in 2012 that removed information about past default for a part of the population, and estimate both distributional and aggregate effects. They show a decrease in borrowing by 6.4% for borrowers negatively impacted by the reform (those who did not experience past defaults) and an increase by 11.8% for those positively impacted. However, because more borrowers are negatively impacted than positively impacted, the aggregate effect is a contraction in credit supply of 3.5%, consistent with the idea that adverse selection has increased in the credit market.

Data sharing is also likely to affect competition in credit markets. A first potential effect is to weaken banks' bargaining power based on the information accumulated over time on their pool of borrowers. Rajan (1992) shows this type of relationship lending allows banks to leverage their informational advantage to extract surplus from borrowers. The portability of data could then mitigate this information hold-up problem. The interaction between adverse selection and competition leads to more nuanced predictions. A stream of papers have studied the impact of screening abilities on credit when lenders competes with imperfect information about borrowers' quality (e.g. Hauswald and Marquez (2003)). He, Huang and Zhou (2020) build on this literature by introducing the possibility for consumers to transfer data from one lender to its competitor. In their models, Fintech firms do not produce as much information as banks but can better exploit it to evaluate credit risk. That improved precision can have ambiguous effects on borrowers' welfare (and even on total welfare). On the one hand, higher precision for all lenders leads to a more efficient allocation of credit by lenders. On the other hand, higher precision for some lenders increases adverse selection for the other lenders. The latter may then compete less aggressively, which allows the former to raise prices (interest rates). In that context, the systematic transfer of bank payment information to fintech lenders may eventually make borrowers worse off by weakening competition. This happens even if the transfer of payment information to a potential lender is optional. Indeed, denying access to personal payment information is then a signal of poor credit risk in equilibrium. In that case, good credit risks are forced to share information and bad credit risks suffer either from the stigma of not sharing or from the improved precision of the FinTech's screening when they share information.

The idea that allowing for payment data sharing leads to equilibria where keeping this information private carries an extra cost because it is interpreted as a bad signal is also present in Parlour et al. (2020) and Ghosh et al. (2021). This points to potential downsides of giving consumers control over payment data. First, data sharing can be associated with privacy costs for consumers. Second, lenders' improved ability to screen loan applicants need not lead to more efficient credit allocation when these borrowers face financial constraints. In that case, additional information can eliminate one channel through which stronger borrowers cross-subsidized weaker ones (similar to the Hirshleifer effect, see Hirshleifer (1971)).

6.5 Conclusion

Payment data interoperability raises questions at the intersection of two streams in economics and finance. The first one has its roots in a long-standing literature on banking that studies the origin and consequences of banks' informational advantage in screening and monitoring lenders. This literature highlights that the use of payment information for lending decisions creates a complementarity between the two activities. Payment data sharing through interoperability has then the potential to improve the efficiency of credit markets by favoring entry and competition, by reducing asymmetric information both across lenders and between lenders and borrowers, and by lowering uncertainty for lenders. This literature also cautions that higher data sharing can have distributional effects that adversely impacts certain groups of borrowers despite positive aggregate effects, and that differential abilities to process data can have a countervailing effects on the reduction of information asymmetries across lenders. The second stream focuses on the allocation of data control rights and its impact on the efficient use of information. Because payment data is a highly valuable input, access to it may in itself justify entry into the payment market and pricing below marginal cost. The benefit of access is higher if the payment provider can retain control over the data its customers generate. Firms' control over data may however lead to suboptimal information sharing ex post (inefficiently low data interoperability), even when markets for data exist. This concern underlies policy frameworks where customers have control over payment data.

7 Concluding remarks

We have reviewed various streams of academic literature to shed light on how the degree of interoperability in mobile payments affects market outcomes and welfare. First, building on the IO literature on tying, we have provided several arguments on why mobile network interoperability may have significant impacts on market outcomes. In particular, lack of interoperability may affect competition both in the telecom and in the payment markets, both discouraging entry and softening in-the-market competition.

Second, in relation to the IO literature on compatibility, we have shown that, absent regulation, interoperability at the platform level may fail to emerge, even if welfare-improving, when a dominant player enjoys strong network externalities or when consumers can join several platforms. We have also discussed some insights from the behavioral IO literature, showing that when consumers face search costs or when fees' structures are complex or hard to observe, the intended pro-competitive effects of interoperability may be considerably weakened or even backfire.

Third, we have shown how interoperability at the agent level can be analyzed through the lens of the literature on co-investment in network industries and of studies on ATMs' interoperability, showing when in these markets we are likely to observe consumers being served by agents of different providers, and what the implications in terms of pricing and welfare are.

Finally, building on the banking literature and on studies on data ownership, we

have highlighted how data interoperability creates a complementarity between payments and other financial services, which may increase overall efficiency but also have adverse impacts on some types of firms and consumers. We have also highlighted how access to consumers' data may in itself promote entry in the payment market, while at the same time leading to inefficiently low data interoperability ex-post.

Many of the dimensions we have highlighted would certainly benefit from further research. One key and probably under-explored aspect concerns the pricing of interoperable systems. Indeed, in practice, tariff systems for interoperable mobile-payment systems are a recent phenomenon, sometimes limited to certain kinds of transactions (e.g., P2P but not P2M), and still far from universal. An interesting topic for future research is to study the optimal design of such a system (from both the firms' and the regulator's perspective). One novel element of this problem is that consumer multi-homing imposes a natural upper bound on the difference between on-net and off-net transfers that telecom/payment providers can charge each other. Another novel element is that consumers' ability to compare tariffs across providers may be imperfect, thereby opening the door to possibly more complex anti-competitive practices.

Another, we believe, very interesting dimension for future research concerns the connection between data interoperability and other forms of payment interoperability. Key aspects of interoperability such as allowing payments across networks, sharing distributors and agents, or bundling telecom and payment networks, have implications for access and control over payment data. In settings where capturing the value of that data is central for firms' decisions, the question of data interoperability should be a key driver of payment interoperability at a broader level.

References

- Aker, J. C., Prina, S. and Welch, C. J. (2020), Migration, money transfers, and mobile money: Evidence from niger, *in* 'AEA Papers and Proceedings', Vol. 110, pp. 589–93.
- Alan, S., Cemalcilar, M., Karlan, D. and Zinman, J. (2018), 'Unshrouding: Evidence from bank overdrafts in turkey', *The Journal of Finance* **73**(2), 481–522.
- Amano, T., Rhodes, A. and Seiler, S. (2021), Flexible demand estimation with search data, Technical report, Harvard Business School.
- Amelio, A. and Jullien, B. (2012), 'Tying and freebies in two-sided markets', *International Journal of Industrial Organization* **30**(5), 436–446.
- Annan, F. (2020), Misconduct and reputation under imperfect information, Technical report.
- Arabehty, P. G., Chen, G., Cook, W. and McKay, C. (2016), 'Digital finance interoperability & financial inclusion', *CGAP report* .
- Armstrong, M., Vickers, J. and Zhou, J. (2009), 'Consumer protection and the incentive to become informed', *Journal of the European Economic Association* **7**(2-3), 399–410.

- Aron, J. (2018), ‘Mobile money and the economy: A review of the evidence’, *The World Bank Research Observer* **33**(2), 135–188.
- Ashraf, N., Karlan, D. and Yin, W. (2006), ‘Tying odysseus to the mast: Evidence from a commitment savings product in the philippines’, *Quarterly Journal of Economics* **121**(2), 635–672.
- Baye, M. R. and Morgan, J. (2001), ‘Information gatekeepers on the internet and the competitiveness of homogeneous product markets’, *American Economic Review* **91**(3), 454–474.
- Bernheim, B. D. and Whinston, M. D. (1990), ‘Multimarket contact and collusive behavior’, *The RAND Journal of Economics* pp. 1–26.
- Bourreau, M., Cambini, C. and Hoernig, S. (2018), ‘Cooperative investment, access, and uncertainty’, *International Journal of Industrial Organization* **56**, 78–106.
- Bourreau, M., Cambini, C., Hoernig, S. and Vogelsang, I. (2021), ‘Co-investment, uncertainty, and opportunism: ex-ante and ex-post remedies.’, *Information Economics and Policy* p. 100913.
- Bourreau, M. and Valletti, T. (2015), ‘Enabling digital financial inclusion through improvements in competition and interoperability: What works and what doesn’t’, *CGD Policy Paper* **65**, 1–30.
- Caillaud, B. and Jullien, B. (2001), ‘Competing cybermediaries’, *European Economic Review* **45**(4-6), 797–808.
- Caillaud, B. and Jullien, B. (2003), ‘Chicken & egg: Competition among intermediation service providers’, *RAND journal of Economics* pp. 309–328.
- Carlin, B. I. (2009), ‘Strategic price complexity in retail financial markets’, *Journal of Financial Economics* **91**(3), 278–287.
- Carlton, D. W. and Waldman, M. (1998), ‘The strategic use of tying to preserve and create market power in evolving industries’.
- Carlton, D. W. and Waldman, M. (2012), ‘Upgrades, switching costs and the leverage theory of tying’, *The Economic Journal* **122**(561), 675–706.
- Carstens, A., Claessens, S., Restoy, F. and Shin, H. S. (2021), ‘Regulating big techs in finance’, *BIS Bulletin No 45* .
- CGAP (2017), How do mobile money fee structures impact the poor?, Technical report, Blog post by Kyle Holloway, Rebecca Rouse, and William Cook.
- CGAP (2019), Interoperability: Why and how providers should pursue it, Technical report.

- CGAP (2021), Building faster better: A guide to inclusive instant payment systems, Technical report.
- Chen, J., Doraszelski, U. and Harrington, Jr, J. E. (2009), ‘Avoiding market dominance: Product compatibility in markets with network effects’, *The RAND Journal of Economics* **40**(3), 455–485.
- Chioveanu, I. and Zhou, J. (2013), ‘Price competition with consumer confusion’, *Management Science* **59**(11), 2450–2469.
- Choi, J. P. and Jeon, D.-S. (2021), ‘A leverage theory of tying in two-sided markets with non-negative price constraints’, *American Economic Journal: Microeconomics* **13**(1), 283–337.
- Choi, J. P. and Stefanadis, C. (2001), ‘Tying, investment, and the dynamic leverage theory’, *RAND Journal of Economics* pp. 52–71.
- Crémer, J., Rey, P. and Tirole, J. (2000), ‘Connectivity in the commercial internet’, *The Journal of Industrial Economics* **48**(4), 433–472.
- Demirguc-Kunt, A., Klapper, L., Singer, D. and Ansar, S. (2018), *The Global Findex Database 2017: Measuring financial inclusion and the fintech revolution*, World Bank Publications.
- Diamond, P. (1971), ‘A model of price adjustment’, *Journal of Economic Theory* **3**(2), 156–168.
- Djankov, S., McLiesh, C. and Shleifer, A. (2007), ‘Private credit in 129 countries’, *Journal of Financial Economics* **84**(2), 299–329.
- Doganoglu, T. and Wright, J. (2006), ‘Multihoming and compatibility’, *International Journal of Industrial Organization* **24**(1), 45–67.
- Donze, J. and Dubec, I. (2006), ‘The role of interchange fees in atm networks’, *International Journal of Industrial Organization* **24**(1), 29–43.
- Donze, J. and Dubec, I. (2009), ‘Paying for atm usage: good for consumers, bad for banks?’, *The Journal of Industrial Economics* **57**(3), 583–612.
- Dosis, A. and Sand-Zantman, W. (2019), ‘The ownership of data’, *Available at SSRN 3420680* .
- D’Silva, D., Filková, Z., Packer, F. and Tiwari, S. (2019), ‘The design of digital financial infrastructure: lessons from india’, *BIS Paper No 106* .
- Dupas, P. and Robinson, J. (2013), ‘Savings constraints and microenterprise development: Evidence from a field experiment in kenya’, *American Economic Journal: Applied Economics* **5**(1), 163–92.

- Ellison, G. (2005), ‘A model of add-on pricing’, *The Quarterly Journal of Economics* **120**(2), 585–637.
- Ellison, G. and Ellison, S. F. (2009), ‘Search, obfuscation, and price elasticities on the internet’, *Econometrica* **77**(2), 427–452.
- Farboodi, M. and Veldkamp, L. (2021), ‘A growth model of the data economy’, *NBER Working Paper No 28427*.
- Farrell, J. and Klemperer, P. (2007), ‘Coordination and lock-in: Competition with switching costs and network effects’, *Handbook of industrial organization* **3**, 1967–2072.
- Fudenberg, D. and Tirole, J. (1984), ‘The fat-cat effect, the puppy-dog ploy, and the lean and hungry look’, *The American Economic Review* **74**(2), 361–366.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. and Walther, A. (Forthcoming), ‘Predictably unequal? the effects of machine learning on credit markets’, *Journal of Finance*.
- Gabaix, X. and Laibson, D. (2006), ‘Shrouded attributes, consumer myopia, and information suppression in competitive markets’, *The Quarterly Journal of Economics* **121**(2), 505–540.
- Gabaix, X., Laibson, D., Li, D., Li, H., Resnick, S. and de Vries, C. G. (2016), ‘The impact of competition on prices with numerous firms’, *Journal of Economic Theory* **165**, 1–24.
- Galenianos, M. and Gavazza, A. (2021), Regulatory interventions in consumer financial markets: The case of credit cards, Technical report.
- Garz, S., Giné, X., Karlan, D., Mazer, R., Sanford, C. and Zinman, J. (2021), ‘Consumer protection for financial inclusion in low- and middle-income countries: Bridging regulator and academic perspectives’, *Annual Review of Financial Economics* **13**(1), null.
- Ghosh, P., Vallee, B. and Zeng, Y. (2021), ‘Fintech lending and cashless payments’, *Available at SSRN 3766250*.
- Gine, X., Martinez Cuellar, C. and Mazer, R. K. (2017), Information disclosure and demand elasticity of financial products : Evidence from a multi-country study, Technical report, World Bank, Washington.
- Goldfarb, A. and Tucker, C. (2019), ‘Digital economics’, *Journal of Economic Literature* **57**(1), 3–43.
- Grubb, M. D. (2015), ‘Failing to choose the best price: Theory, evidence, and policy’, *Rev Ind Organ* **47**, 303–340.
- GSMA (2012a), The case for interoperability: Assessing the value that the interconnection of mobile money services would create for customers and operators, Technical report, Neil Davidson and Paul Leishman.

- GSMA (2012*b*), Is there really any money in mobile money?, Technical report, Paul Leishman.
- Hannan, T. H., Kiser, E. K., Prager, R. A. and McAndrews, J. J. (2003), ‘To surcharge or not to surcharge: An empirical investigation of atm pricing’, *Review of Economics and Statistics* **85**(4), 990–1002.
- Hauswald, R. and Marquez, R. (2003), ‘Information technology and financial services competition’, *Review of Financial Studies* **16**(3), 921–948.
- He, Z., Huang, J. and Zhou, J. (2020), ‘Open banking: credit market competition when borrowers own the data’, *NBER Working Paper No 28118*.
- Hirshleifer, J. (1971), ‘The private and social value of information and the reward to inventive activity’, *American Economic Review* **61**(4), 561–574.
- Huck, S. and Wallace, B. (2015), The impact of price frames on consumer decision making: Experimental evidence, Technical report, UCL.
- IPA (2021), Ipa consumer protection research initiative: Rfp overview, Technical report, <https://www.povertyaction.org/sites/default/files/presentation/IPA-Consumer-Protection-RFP-Info-Session-Feb-10-11-2021-Presentation-Final.pdf>.
- Jappelli, T. and Pagano, M. (2002), ‘Information sharing, lending and defaults: Cross-country evidence’, *Journal of Banking & Finance* **26**(10), 2017–2045.
- Jolivet, G. and Turon, H. (2018), ‘Consumer Search Costs and Preferences on the Internet’, *The Review of Economic Studies* **86**(3), 1258–1300.
- Jones, C. I. and Tonetti, C. (2020), ‘Nonrivalry and the economics of data’, *American Economic Review* **110**(9), 2819–58.
- Kalaycı, K. and Potters, J. (2011), ‘Buyer confusion and market prices’, *International Journal of Industrial Organization* **29**(1), 14–22. Special Issue: Experiments in Industrial Organization.
- Katz, M. L. and Shapiro, C. (1985), ‘Network externalities, competition, and compatibility’, *The American economic review* **75**(3), 424–440.
- Knittel, C. R. and Stango, V. (2008), ‘Incompatibility, product attributes and consumer welfare: evidence from ATMs’, *The BE Journal of Economic Analysis & Policy* **8**(1).
- Lieberman, A., Neilson, C., Opazo, L. and Zimmerman, S. (2018), ‘The equilibrium effects of information deletion: Evidence from consumer credit markets’, *NBER Working Paper No 25097*.
- Liberti, J. M. and Petersen, M. A. (2019), ‘Information: Hard and soft’, *Review of Corporate Finance Studies* **8**(1), 1–41.

- Malueg, D. A. and Schwartz, M. (2006), ‘Compatibility incentives of a large network facing multiple rivals’, *The Journal of Industrial Economics* **54**(4), 527–567.
- McAndrews, J. J. (2003), ‘Automated teller machine network pricing – a review of the literature’, *Review of Network Economics* **2**(2), 146–158.
- Mian, A. (2006), ‘Distance constraints: The limits of foreign lending in poor economies’, *Journal of Finance* **61**(3), 1465–1505.
- Naji, L. (2020), ‘Tracking the journey towards mobile money interoperability’, *GSMA report*.
- Pagano, M. and Jappelli, T. (1993), ‘Information sharing in credit markets’, *Journal of Finance* **48**(5), 1693–1718.
- Parlour, C. A., Rajan, U. and Zhu, H. (2020), ‘When fintech competes for payment flows’, *Available at SSRN 3544981*.
- Petersen, M. A. and Rajan, R. G. (2002), ‘Does distance still matter? the information revolution in small business lending’, *Journal of Finance* **57**(6), 2533–2570.
- Piccione, M. and Spiegler, R. (2012), ‘Price Competition Under Limited Comparability’, *The Quarterly Journal of Economics* **127**(1), 97–135.
- Posner, R. A. (1978), ‘The Chicago school of antitrust analysis’, *U. Pa. L. Rev.* **127**, 925.
- Rajan, R. G. (1992), ‘Insiders and outsiders: The choice between informed and arm’s-length debt’, *Journal of finance* **47**(4), 1367–1400.
- Rhodes, A. (2015), ‘Multiproduct retailing’, *The Review of Economic Studies* **82**(1), 360–390.
- Ronayne, D. (2021), ‘Price comparison websites’, *International Economic Review* **62**(3), 1081–1110.
- Shelegia, S. (2012), ‘Multiproduct pricing in oligopoly’, *International Journal of Industrial Organization* **30**(2), 231–242.
- Stahl, D. (1989), ‘Oligopolistic pricing with sequential consumer search’, *American Economic Review* **79**(4), 700–712.
- Suri, T. and Jack, W. (2016), ‘The long-run poverty and gender impacts of mobile money’, *Science* **354**(6317), 1288–1292.
- Varian, H. (1980), ‘A model of sales’, *American Economic Review* **70**(4), 651–659.
- Wilson, C. M. and Waddams Price, C. (2010), ‘Do consumers switch to the best supplier?’, *Oxford Economic Papers* **62**(4), 647–668.
- Yang, Y. (2021), Price transparency in online markets, Technical report, Toulouse School of Economics.