Shifting branchless banking regulation from enabling to fostering competition

Ignacio Mas, Senior Research Fellow, Saïd Business School, University of Oxford[1]
29 July 2014

Abstract

Branchless banking solutions in most countries tend to be dominated by a few large players, and exhibit low levels of innovation. The paper argues that there is a need to evolve the regulatory framework for branchless banking from one that enables participation by banks and telcos to one that fosters competition by a broader range of players. Regulations on e-money issuers, retail agents and account opening need to be recast so as to reduce the cost of entry and give much more scope for service and business model innovation. In addition, there is a growing need for policymakers to ensure there is a level playing field across all players, and that mobile operators do not exploit their dominance in the mobile communications market to gain advantage in the new market for mobile financial services.

The promise of branchless banking

The development of branchless banking is still incipient in a large number of developing countries, but there is a select group of countries that has been experimenting with such solutions for a decade or longer. The success of services such as M-PESA in Kenya and Caixa Economica Federal´s and Bradesco´s agents in Brazil have long inspired commercial entities and policymakers alike to think that there are alternative, lower-cost delivery models which have the promise of bringing formal financial services to millions of previously unbanked people.

At the same time, an increasing body of literature shows how varied and nuanced people´s financial habits and needs are. People´s money management practices are riddled with psychological discipline devices, wrapped up in personal relationships, and charged with social meaning. It will not be easy to displace the rich variety of informal financial practices with a few simple products that can be effectively marketed and communicated through lower-touch branchless channels.[2]

There is therefore going to be much scope for experimenting with and refining the service propositions from branchless banking platforms, so that they offer a clear reason to switch (benefits relative to current alternatives) while minimizing switching costs (the amount of customer education they require for people to become comfortable in using them).

---

[1] Funding for the author's fellowship at the Saïd Business School and for this paper comes from the Bill & Melinda Gates Foundation. The author thanks both organizations for their support.

[2] For a description of the product development challenges involved with replicating informal financial practices digitally, see I Mas, "Digitizing the Kaleidoscope of Informal Financial Practices." Journal of Business Management & Social Sciences Research Vol. 3, No.5, May 2014.

Organizations with more affinity and trust with their customers are likely to be more successful in this process than providers who merely have network or financial muscle.

The limitations of branchless banking systems today

We have seen how, when they work well, branchless banking systems can indeed reach millions of people. But beyond the headline numbers on customers reached, the record of branchless banking systems is still mixed. It is fair to say that branchless banking remains an unproven approach to effective financial inclusion at a global scale. We can draw some stylized facts from the international experience:

- *Branchless banking systems have only tended to work at scale*. There does not appear to be an easy, gradual *incremental path* for providers wishing to deploy branchless banking solutions. There seems to be a chasm between the large numbers of institutions that have run sub-scale pilots and the much smaller set that have succeeded in establishing commercially sustainable branchless banking operations.[3] As a result, there are very few examples of smaller entities –whether banks, mobile operators, microfinance institutions, or other third parties— successfully incorporating branchless banking solutions in a sustainable, impactful way.[4]

- *Branchless banking is still dominated by mobile operators*. Few banks in the world seem to have made sizable bets to develop agent networks, and most of those who have built agent networks have tended to see them as an add-on for specific services (e.g. bill payment) or for specific segments (e.g. poorer people) rather than as an extension of their core business.[5] Non-financial companies with a retail or distribution background have been reticent to jump into the space. Therefore, the space has been left largely to mobile operators, who have an easier time conceiving of a transactional, high-volume, low-touch approach.

- *Customers tend to use branchless banking systems relatively infrequently, and only for a limited range of applications*. The median active user is likely to make a transaction only once or twice a month. Customers tend to use branchless banking systems principally for remote person-to-person payments, bill payment, and mobile airtime purchases.[6] The mix of these particular services vary country by country (actually: more like continent by continent), but it is not common to see branchless banking

[3] Zoona in Zambia is probably a counterexample: a small organization growing a purely mobile-based money system incrementally by exploiting specific niche opportunities.

[4] Perhaps the most notable exception is bKash in Bangladesh. Though it is backed by BRAC Bank, which is part of one of the most influential organizations in the country, it operates largely as an independent entity.

[5] Banks that have deployed more extensive agent networks tend to be concentrated in Latin America, where the agent model originated because of high bill payment volumes that were cluttering bank branches, and in Kenya probably due to the pressure from the success of the telco-led M-PESA service.

[6] For useful statistics on mobile money usage based on a survey of a large number of mobile money providers, see Claire Pénicaud, "State of the Industry: Results from the 2012 Global Mobile Money Adoption Survey," GSMA, Mobile Money for the Unbanked, February 2013.

being a "stepping stone" or "gateway" into the use of a fuller range of financial services.[7]

- *Branchless banking is not fundamentally reducing people´s reliance on cash*. Most mobile money transactions start and end in cash. We may refer to it as a mobile or electronic transaction, but most customers would understand it as a cash-to-cash money transfer, akin to what Western Union has always done. The payment may be electronified, and as a result the distance that cash needs to move is much reduced. But the underlying money is not electronified, since the value is largely held in cash before and after the transaction. Branchless banking systems have generally failed to position the store-of-value function of customer accounts among the previously un- or under-banked, and the result is that the majority of accounts are actually or practically empty.

- *Branchless banking systems tend to exhibit relatively low levels of service innovation*. Branchless banking –and in particular mobile money— systems are about exposing financial service platform functionalities directly to the customer by digital means. But this has not brought on the kind of constant innovation that has been the hallmark of internet business models. Of course, the need to work on basic phones has hampered the ability to innovate, but the fact remains that most branchless banking providers have brought on new services or optimized their user interfaces not more frequently than annually, if at all.[8]

All of these factors are inter-related, like distinct symptoms of a broader malaise. The seemingly inescapable reality of conversion from and to cash with each transaction not only raises the cost of each transaction substantially but it also presents a brutal business challenge of having to ensure sufficient density of liquid agents in each locality served. Higher transaction costs make the system less compelling for lower customer-value-adding transactions, such as savings or face-to-face merchant payments, which on the other hand, offer the highest potential pool of transactions. In the face of low usage levels per customer and the inherent network effects of payment businesses, the economics can only work for those able to aggregate the largest number of customers, and in particular mobile operators with a mass market transactional business model. Other big players such as banks may not see a positive business case, or if they do, may fear that the new branchless banking activity may cannibalize their core business or be margin dilutive. As a result, few players in each market enter the business, and when they do they tend to underinvest in IT platforms, staffing and marketing spend. With such shoestring resources, they become easily overwhelmed by day-to-day operational issues and do not devote much attention to the service roadmap. With lack of effective competition, innovation falters.

The case for more competition and innovation

The branchless banking models we see today follow fundamentally an *extensive* logic: they are focused on getting lots of customers to do one or two transactions per month on

[7] Where mobile money has flourished, it is far more common to see the opposite: fully-banked people adopting mobile money as "liquidity extension" to their banking service.

[8] For a review of product innovations in mobile money, see M Almazán and I Mas, "Product Innovations on Mobile Money Platforms," forthcoming in the European Journal of Business Management.

average, made up largely of higher-value ($15 and upwards) transactions, occurring in remote (i.e. not face-to-face) settings that are not so well suited to cash. The math only works for the largest players.[9]

Over-reliance on the largest players, especially those whose interest in financial services is only ancillary, to "solve" the financial inclusion challenge at scale has several potential pitfalls. First, it limits the range of players who can contest the market to just a handful in any given market. Since not all may be interested in developing this market, the outcome may be just two, only one, or even no serious players. This may not produce enough competition to drive efficiency gains leading to sustained price reductions, or to foster the development of new service innovative on a regular basis. Second, these large players may feel emboldened to seek to control the market and entrench their position by curtailing interoperability with other players. Therefore, their solutions will lack universality, to the detriment of their own customers. Third, the largest players may not have the kind of granular segmentation capability and affinity with the target market of the unbanked that would be required to effectively market to them. The largest players, by definition, tend to focus their marketing on the more affluent segments within the mass market.

While these large players can play a valuable role in driving solutions to scale, there is a need to have smaller, more nimble players contesting the market. Lacking scale, they will need to search for *intensive* models, based on a smaller number of customers doing far more regular transactions and unleashing extra customer value. They too have a valuable role to play: providing a countervailing power to the larger players, developing more targeted offers based on a deeper understanding of particular customer segments, and introducing a regular flow of innovations.

Moreover, having more specialized players provides its own path to scalability, as each player becomes strong in providing particular elements of service and they leverage each other to create strong complementarities.

In the end, financial exclusion occurs because there are no or few financial institutions wishing to serve the poor or sustaining enough learning and experimentation about how best to serve them. The recipe for financial inclusion ought to be, therefore, increasing competition. And in particular, increasing competition among different service and delivery models. That sort of competition is not operating today in most markets: the few large players contesting a given market are likely to be following largely the same *intensive* model we have outlined above.

Why haven´t we seen successful intensive models emerged? Why haven´t we seen more diversity of models and more experimentation by smaller players whose success depends on it?

So what is an innovator-entrepreneur to do?

Imagine that you are an entrepreneur with a clear idea for a cool new mobile-based money management application that you think will make people want to hold their liquidity (not just transact) in electronic form. Based on that, you expect to accumulate more saved

[9] See I Mas, "Mobile Money Maths," Mobile Money Asia blog, 19 June 2013, available here.

balances, trigger more electronic payments as people dispose of their higher electronic money balances, and capture more relevant information about customers´ financial behaviors from their electronic balances and payment habits which can help you position more credit responsibly. This is an intensive-use model, with a triple path to revenues (low-cost funding, transactional revenue, credit).

But now consider the steps you need to take even before you start developing your app:

1. You need to get some kind of license as a financial institution, since you will be collecting and holding the public´s money. You may be forced to get one with onerous terms that are not proportional to the activity you propose to undertake, or, if you get a more limited one, it may come with restrictions that limit your vision.
2. If your target customers are previously unbanked, you will need to collect, process and store KYC information from them, *before* you and they determine how valuable the relationship is.
3. You need to develop your own national network of cash in/out points, so that your customers can opt into your service.
4. You are likely to have to negotiate a good deal with the one or more of the big mobile operators, whom you will in fact likely have to compete with in your financial service, so you can get access to an acceptable mobile channel at an affordable price.

Each of these requirements is a major barrier to entry in its own right. Collectively, they constitute a daunting list that ought to dissuade all but the most determined, well-funded startups.

Note that each of these requirements has a regulatory dimension: the first three have to do with financial integrity and safety, and the fourth with competition policy. The conclusion therefore seems to be that the current regulatory environment is stacked against entrepreneurs, because it presents large barriers to entry and does not guarantee a level playing field for the smaller, independent players.

Towards second-generation branchless banking regulations

This need not be so: there are steps that can be taken by regulators to curtail barriers to entry while still entirely protecting the integrity and safety of branchless banking system. What is required is to evolve the regulatory framework for branchless banking from one that *enables* participation to one that *fosters* competition.

First-generation branchless regulations have typically centered on the following three key elements:

1. *Agent regulations for cash in/out*. Authorizing licensed financial institutions to engage retail shops as agents, under contract, with the financial institution assuming full liability for the actions of its agents.
2. *Progressive KYC*. Introducing tiered Know Your Customer (KYC) regimes, permitting account opening with less customer information and documentation capture, based on the nature of risks involved for particular classes of customers.

3. *E-money licensing*. Creating an e-money license distinct to banking, so as to permit a new class of players to develop business models focused on electronic payments rather than on financial intermediation.

In the rest of the paper, we will argue that these three elements need to be reinterpreted so as to reduce the cost of entry and give much more flexibility for new entrants wishing to contest the market. In addition, there is a growing need for policymakers to ensure there is a level playing field across all players, whether they are large or small, whether they have one type of license or another, and whether they are banks, telcos or other types of players. We can therefore expand the list of regulatory issues above to include:

4. *Preventing regulatory arbitrage*. Ensuring that there is no unjustified regulatory arbitrage opportunities between different types of financial service license holders.
5. *Preventing anti-competitive practices by dominant players*. Preventing larger players from exploiting their scale advantage to lock out smaller competitors, by driving towards interconnection of platforms (interoperability) and precluding pricing below cost (anti-dumping).
6. *Preventing mobile operators´ abuse of essential service elements under their exclusive control*. Ensuring that mobile operators do not use their market power in the communications market and their control over the telecoms numbering range to gain unfair control over financial service providers who must use the telecoms services of mobile operators.

Each of these six issues are dealt with separately in the rest of this paper, focusing on the economic rather than the legal arguments.[10]

Box: Key resources on regulation of branchless banking and mobile money

Alexandre, Claire, Ignacio Mas, and Daniel Radcliffe (2011). "Regulating new banking models that can bring financial services to all." Challenge, Vol. 54, No. 3, May-June.

Chatain, Pierre, Raul Hernandez-Coss, Kamil Borowik, and Andrew Zerzan (2008). "Integrity in Mobile Phone Financial Services: Measures for Mitigating Risks from Money Laundering and Terrorist Financing." Working Paper 146. Washington, DC: The World Bank.

di Castri, Simone (2013). "Mobile Money: Enabling regulatory solutions." Mobile Money for the Unbanked report, February. London: GSMA.

Dias, Denise, and Katharine McKee (2010). "Protecting Branchless Banking Consumers: Policy Objectives and Regulatory Options." Focus Note 64, September. Washington, DC: CGAP.

Khiaonarong, Tanai (2014). "Oversight Issues in Mobile Payments." Working Paper 14/23, July. Washington, DC: International Monetary Fund.

Klein, Michael, and Colin Mayer (2011). "Mobile banking and financial inclusion: the regulatory lessons." Policy Research Working Paper Series 5664, May. Washington, DC: The World Bank.

Lyman, Timothy, Gautam Ivatury, and Stefan Staschen (2006). "Use of Agents in Branchless Banking for the Poor: Rewards, Risks and Regulation." Focus Note 38, October. Washington, DC: CGAP.

Lyman, Tim, Mark Pickens, and David Porteous (2008). "Regulating Transformational Branchless Banking: Mobile Phones and Other Technology to Increase Access to Finance." Focus Note 43,

[10] One important area that is not specifically covered here is consumer protection regulation. Adequate consumer protection ought to be an integral part of a competitive system but is probably not itself a precondition for competition to emerge.

January. Washington, DC: CGAP.

Mas, Ignacio (2012). "Transforming access to finance in developing countries through mobile phones: creating an enabling policy framework. Banking & Finance Law Review, Vol. 27, No. 2, January.

Staschen, Stefan, Ahmed Dermish, and Lara Gidvani (2012). "Regulatory Impact Assessment Methodology: Towards Evidence Based Policy Making in Financial Inclusion." Bankable Frontier Associates Concept Note, September.

Tarazi, Michael, and Paul Breloff (2010). "Nonbank E-Money Issuers: Regulatory Approaches to Protecting Customer Funds." Focus Note 63, July. Washington, DC: CGAP.

Tarazi, Michael, and Tilman Ehrbeck (2011). "Putting the Banking in Branchless Banking: Regulation and the Case for Interest-Bearing and Insured E-money Savings Accounts." Mobile Financial Services Development Report 2011. Geneva: World Economic Forum.

Agent regulations for cash in/out

Under typical branchless banking regulations, there is a presumption that there is a principal-agent relationship between an authorized financial institution and the (collection of) retail establishments that offer deposit and withdrawal (or cash in/out) services to their clients. Need that be the case, or can cash in/out transactions be viewed in the same terms as any other merchant transaction it engages in.

A cash-in (cash-out) is an electronic transfer of money from the store (customer) to the customer (store), for a corresponding transfer of physical cash in the inverse direction. It is an exchange of two forms of monetary value between the store and the customer. The bank enters the deal as the trusted supplier of the technology platform that underpins the electronic side of the transaction: the bank guarantees that the store (customer) is the rightful owner of the necessary amount of e-money, and records the transfer of ownership of the e-money to the customer (store). In other words, the bank is an enabler of the cash in/out transaction, but not a party to it. The store only handles its own stock of cash and electronic money balances, and never the financial institution´s or any of their customers´.

This is very different to the process for depositing at a bank branch, where the teller takes physical possession of the customer´s cash in the name of the bank, and the customer´s account is credited as being owed more by the bank. A deposit at a bank branch results in a net increase in the bank´s liabilities (deposits), while a cash-in at an agent does not since value was merely transferred between account-holders.

The relationship between a customer and a store when they are performing a cash-in transaction is in fact very similar to their relationship when the customer goes to the same store to buy rice for cash (see Figure 1). No one would claim that the store is an agent for the rice manufacturer, as long as the store is selling its own inventory of rice. The store has a mere client relationship with the rice manufacturer or distributor, in so far as the store has to buy stocks of rice from time to time, with which to trade on its own account. Why would the same not apply if the store sells (and buys) its own stock of electronic money that is issued and managed (but is not owned) by a financial institution?

The regulatory insistence on seeing a principal-agent relationship for cash in/out results in two common regulatory mandates: *(i)* financial institutions are required to sign agency contracts with such retail players, and *(ii)* the bank retains legal responsibility for all the

actions committed by its cash agents. These are onerous requirements, which raise significant logistical and legal problems for financial institutions. Financial institutions are, realistically, in no position to manage the risk from the actions of thousands of independent players. Also, larger institutions who wish to develop their own proprietary cash in/out networks can deny agency status to independently-managed shared agent network, thereby locking them out and making it more difficult for shared agent networks to emerge. For all these reasons, these two requirements create significant barriers to the development of healthy agent networks.

Figure 1: The analogy between store-based cash in/out and everyday merchant transactions

(a) Normal merchant transaction: exchanging rice for cash

(b) Agent transaction: exchanging electronic value for cash (and *vice versa*)

One way to address these limitations is to permit contractual separation between account issuance/management (performed by financial institutions) and cash in/out services (performed by independent retail outlets). That´s in fact how most indirect distribution models work: the rice company and the Coca Cola company do not have direct contracts with each and every store that sells their rice and soda bottles. Independent distributors can emerge that buy rice and soda bottles from the manufacturers, and resell them to their own retail base.

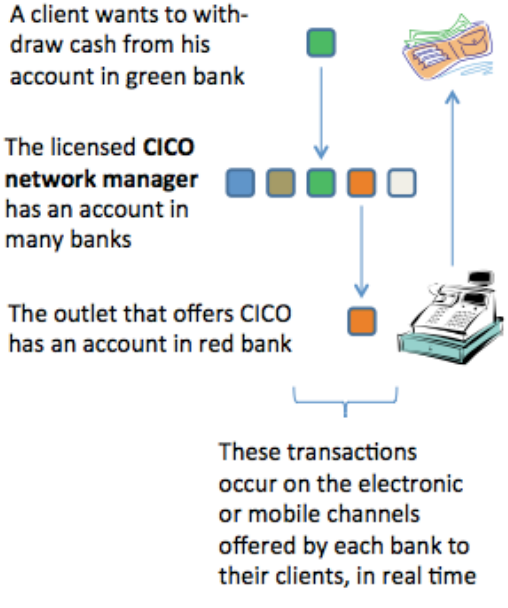
In theory, any individual or business with a bank account and with access to an electronic (or mobile) transactional capability from that account could offer cash in/out services for that bank, without necessarily requiring the explicit agreement of the bank. It´s not that the bank would not be involved: it would be endorsing the transaction insofar as it would occur on its technology platform. In that situation, the bank would vouch for the propriety of the electronic transaction, but could not vouch for the correct completion of the offsetting cash transaction – but that´s already a common risk that we all accept every time we walk into a retail outlet to buy goods.

It can be argued that exchanging cash for rice is not the same as exchanging cash for electronic value for two main reasons:

- *Consumer protection risks*. A cash in/out transaction gives rise to special consumer protection risks, in so far as transactions might be larger than typical merchant transactions and customers may not be familiar with the process for undertaking them. Hence, customers engaging in cash in/out transactions in a retail environment may require special support in ensuring that they are adequately informed, treated in a fair manner, and have proper recourse mechanisms if they feel aggrieved.

- *Technology platform risks*. Rice is a fairly standard commodity: while a store may decide to stock a few varieties and qualities of rice, there is no reason for it to hold stock from each rice manufacturer. In the case of electronic money, who the issuer is matters, if the store´s customers hold electronic money from different financial institutions and there is no effective interoperability between their platforms. In this case, each store would need to hold balances with and use the mobile banking capability provided by each financial service provider it wants to offer cash in/out services for. This is too burdensome, so stores are likely to look to independent agent network managers to act as a one-stop-shop technology provider and clearinghouse across the various banks and e-money issuers they want to work with. Stores would operate out of a single account with their agent network manager, possibly using only the technology platform managed by the independent network manager, and the agent network manager would then offset the transaction with the customer on the customer´s financial institution platform. This is represented in Figure 2. But the introduction of these independent network managers will create new technology and counterparty risks of their own, insofar as they provide the technology platform on which the store performs cash in/out transactions with customers.

Figure 2: Operations of an independent cash in/cash out (CICO) network



These two sources of risk could be addressed by creating a new license category, issued by the banking regulator, for independent agent network managers (IANM). IANM license holders would be free to contract with any store they like and to offer cash in/out services for any financial institution they have a (corporate) customer account with. But IANM licensees would have to abide by specific consumer protection obligations, such as:

- Obligation to post, in each outlet, a notice with the prices of their cash in/out services, list of issuers served, and customer care hotline. (There may be an obligation to maintain uniform pricing of cash in and cash out across all issuers, in order to avoid customer confusion.)
- Obligation to investigate and to report to the regulator on customer complaints, and to maintain a blacklist of stores where frequent problems have been reported.
- Each outlet is subject to the possibility of on-site inspection, and to be terminated as a cash in/out outlet, by the regulator.

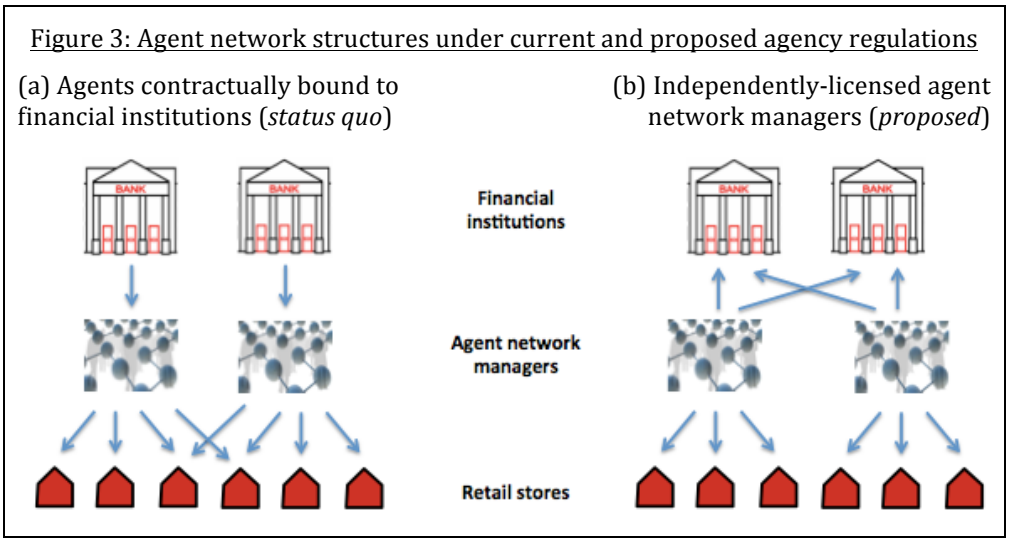
Licensed IANMs would also have to abide by regulations and be supervised by the financial regulator in regards to the technology platform and financial clearing procedures they use with their stores. This could be based on very straightforward checklists to ensure that their

systems meet the required standards of security, privacy and integrity, and that they do not create large counterparty risks.

The introduction of this licensing regime can be expected to change the dynamics for agent network development quite substantially, in the way represented schematically in Figure 3. The left-hand panel shows the typical regulatory arrangement where financial institutions are required to sign up agents contractually: the financial institutions pick their own network agent managers (or they perform this function in-house), and their network managers in turn pick retail stores to act as their sub-agents. There may be an overlap in the agent base of different institutions, as some stores may sign up contracts with competing financial institutions (through their agent network managers) in order to meet the needs of more of their customers. But few individual stores are likely to want to sign up with *every* financial institution. Independent agent network managers may wish to sign up every financial institution, but they may find that some institutions –likely the larger ones— want to maintain a proprietary agent channel and refrain from contracting with them.

The right-hand side panel of Figure 3 shows the likely arrangement under the proposed IANM licensing regime proposed above. Each IANM can now offer cash in/out for any financial institutions it wishes, and no financial institutions can block them from doing so. (It may be necessary to prescribe in regulation that financial institutions cannot deny service as corporate customers to licensed IANMs.) IANMs are also each likely to sign up different stores, because there is no longer a reason for any single store to sign multiple agency contracts.

Figure 3: Agent network structures under current and proposed agency regulations



Therefore, the proposed policy reform would have the following benefits:

Issue	*Benefits under proposed reform*	*Vs. current situation*
Speed to critical mass	Retail cash in/out outlets could gain critical mass faster, since they would be able to serve any customer of any financial institution from the outset, without any possibility of them being excluded by the larger banks as they	Cash in/out outlets are often stuck with only one, or maybe a few financial institutions, and hence are not able to serve all their customers within their service area.

	pursue their own agent network.	
Shared agent networks	Shared agent networks would naturally emerge, because that would be the natural vocation of all stores and IANMs. All it would take is for the IANM to become a corporate customer of each financial institution.	Shared agent networks could only emerge after protracted negotiations between financial institutions, with the bigger ones having little incentive to participate.
Geographic coverage	Local retail players in underserved areas could set up as IANMs and rally stores to bring cash in/out services to their communities. Scale and breadth of geographic reach would no longer be so important.	Financial service providers generally have a national scope, and hence tend to seek out national players who can help them build agent networks. This naturally limits the participation of more localized players in building agent networks.
Branded competition at agent level	Each licensed IANM would logically seek to brand their agents, and differentiate their brand based on the reliability, accessibility and service quality of its outlets.	The brand of the provider takes primacy over the brand of the cash in/out outlet or the agent network manager. Thus, all outlets for the same institution are competing with no possibility of brand differentiation between them.

Even if this regulatory approach were enacted, financial institutions would still be able to develop their own agent networks as they do today. The reform would open up the possibility of an additional or alternative path to development of cash in/out outlets. Its value is premised on creating a path for new entry models into and spurring more competition within the cash in/out business.

Progressive KYC

There has been good progress towards adopting more risk-based approaches to KYC in selected countries. This is implemented by allowing for a graduated approach to KYC. The objective is to require burdensome proof-of-identity and other documentation only as and when it is justified by the kind of usage that the customer is likely to give to that account.

Authorities should treat the tailoring of KYC rules as an ongoing journey rather than a one-off reform. As the experience with branchless banking systems builds up, there ought to be a more thorough and granular understanding of the risks involved, and this should lead to a progressive fine-tuning of account opening requirements and transactional limits. This would permit a greater variety of sales approaches and customer growth strategies by institutions wishing to onboard poorer people.

Moreover, the burden of KYC compliance imposed on financial institutions ought to be, at least in principle, inversely related to the amount of public infrastructure that exists around identity. As governments adopt national ID schemes, build online registries enabling real-time ID checks, and invest in security systems that make ID counterfeit and theft more

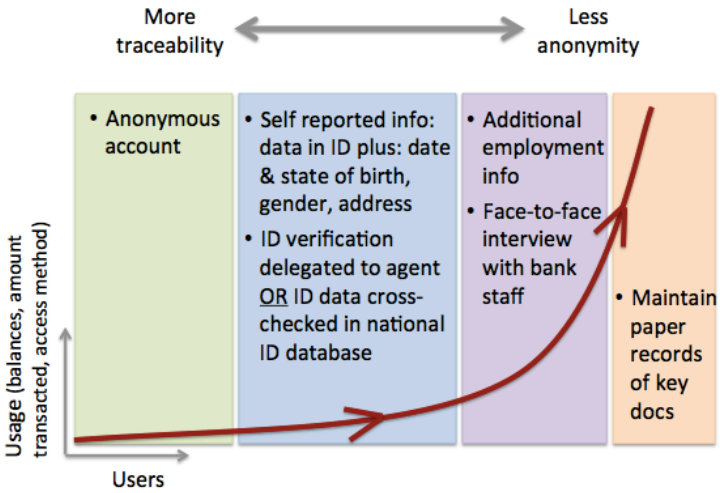
difficult, KYC restrictions should adapt to reflect the growing reliance of technology solutions to the question of identity.

The purpose of creating differentiated KYC requirements is not so much to fit different population segment into "KYC tiers" based on expected financial service usage levels, but rather to avoid unnecessary upfront account opening costs for people who are as yet uncertain about how and how much they might use them. By allowing for a deferral of some of the more onerous KYC elements, these customers can learn about and experiment with the service, and only conduct fuller KYC once they have discovered the value of the account for them and are ready to use it more fully. Thus, a regime of progressive KYC should be seen fundamentally as one permitting a deferral of KYC.

The reality is that most electronic account issuers experience inactivity levels in the 30-60% range, and among active accounts it is not unusual to find that half keep low balances of up to $50. Given this situation, it seems unnecessarily costly and counterproductive to require that new customers conduct stringent KYC upfront.

Figure 4 shows schematically how the Mexican progressive KYC regime works.[11] Four KYC levels are mapped loosely against the typical curve that one sees relating customer numbers to account usage statistics. The arrows on the curve are a reminder that the curve can also be interpreted as a personal customer progression, in which case the four KYC levels can be thought of as potential "upgrades" on any account. The four levels are differentiated by clear account size and transaction volume limits (not shown).

Figure 4: Schematic of a progressive KYC regime



In Mexico as elsewhere, KYC levels vary according to the following parameters: *(i)* the amount and nature of information that customers need to report and that the system needs to capture; *(ii)* the documents customers need to bring to verify that information (e.g. ID, proof of address), and how these documents are verified if at all (online checks, face-to-face

[11] Per the resolution that modified article 115 of the Law of Credit Institutions, issued by the Ministry of Finance on 12 August 2011, available here.

interview with bank staff, or delegated to agents); and *(iii)* whether documents need to be captured and stored by in paper form or only digitally.

In setting graduated KYC levels, regulators generally focus on the magnitude of money laundering and other illicit usage risks presented by different classes of users, based on their characteristics and usage levels. But the other consideration they need to take into account as they tackle illicit financial activities is striking the right balance between *(i)* eliminating anonymity by enforcing tougher KYC policies, and *(ii)* promoting greater traceability of transactions by inducing more transactions to happen electronically. There is an implicit trade-off between these two approaches: tougher KYC requirements that result in higher exclusion will make untraceable cash transactions more common and socially acceptable. As suggested by the arrow at the top of Figure 4, the lower KYC tiers on the left of the chart should be designed primarily to induce more participation in the electronic payments system in an attempt to increase traceability, while the higher KYC tiers on the right of the chart should be designed to tackle anonymity more forcefully.

E-money licensing

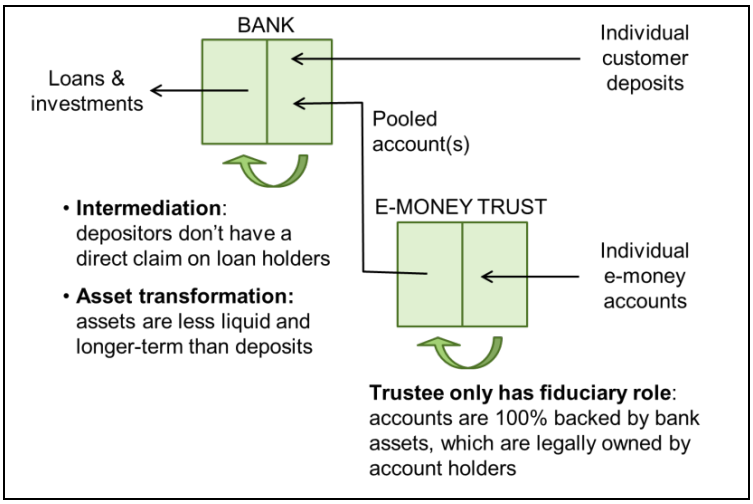
It would seem odd to seek financial inclusion by excluding certain players, and yet some countries have done just that by precluding the direct participation of mobile operators as licensed financial service providers. There is a need to create a more diversified financial service licensing regime that permits a broad range of players to participate in ways that suit their business requirements while still ensuring the stability and integrity of the financial system.

Some countries have created a special license, typically known as an e-money issuer (EMI) license, to enable non-banks (such as mobile operators) to operate branchless banking services. An EMI faces draconian restrictions on what it can do with any money they hold on behalf of their customers: typically they need to put it in a trust or special-purpose vehicle, and the trust can only invest in safe assets such as liquid bank deposits. While the EMI is the legal issuer of customer accounts, it has very little room for discretion on how to invest the funds that reside in those customer accounts.

Thus, EMIs can be thought of as a *pass-through* mechanism to deliver money into the banking system, as depicted schematically in Figure 5. In return for such restrictions on its activities, and recognizing the intrinsically low-risk nature of its operations, EMIs face significantly lower capital requirements and are exempted from much of the prudential regulations that apply to banks.

EMIs do not engage in two of the main functions of banks: intermediation (reinvesting depositors´ funds as loans) and asset transformation (investing liquid or short-term deposits into longer-duration assets). But they do perform two functions that banks also do: managing individual customer accounts (including the cash deposits and withdrawals associated with those) and managing payments (transfers between accounts and between institutions).

Figure 5: E-money as a pass-through into the banking system



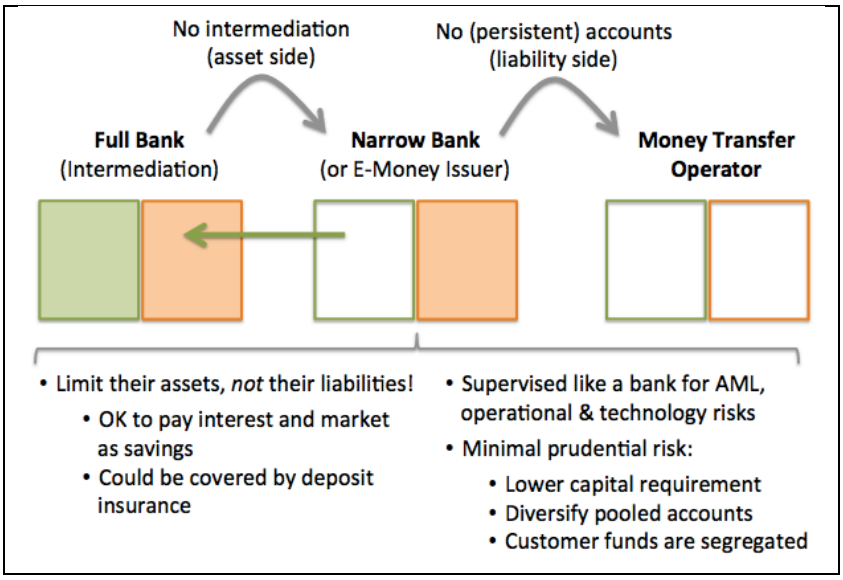
Where they exist, EMIs generally are regulated under the payment system law rather than the banking law, and supervised by the payments systems rather than the bank supervision unit at the financial authority. This, and the fact that most EMIs are in fact used by customers primarily to effect payments, has created the view that they are fundamentally about payments – or at most a "light form" of banking. However, this takes a narrow, overly prescriptive view of their potential role in financial inclusion.

It is more appropriate to reinterpret e-money issuance less as an alternative to banking and more as a form of banking – in particular, as *narrow banking*. A narrow bank is one which issues deposits but does not intermediate the corresponding funds. It is therefore like a normal bank on the liability side of its balance sheet, but carries none of the riskier banking activities on the asset side. This distinguishes it from payment service providers and money transfer operators, which do not issue accounts at all and hence do not carry any deposit (bank-like) liabilities.

Figure 6 then lays out a licensing regime that has three pillars rather than the traditional two of banking and payments. Full banks can issue deposits, intermediate the funds they raise through their deposits (i.e. loan them out), and accept a maturity mismatch between their assets and liabilities. Narrow banks can issue the same types of liabilities as banks and facilitate payments from those accounts, but they cannot on-lend the funds or engage in maturity transformation. Money transfer operators cannot issue accounts or otherwise raise funds from the general public (other than on a temporary basis, to accommodate the timing mismatches between when payments are requested and collected).

Therefore, the extended licensing regime can be understood as a segregation of certain components of a bank´s typical balance sheet. As you move from left to right in Figure 6, narrow banks have limitations on their asset side relative to full banks, and money transfer operators (and other payment system providers) have limitation on their liability side relative to full or narrow banks.

Figure 6: The three licensing pillars for financial inclusion



Reinterpreted in this way, EMI-issued accounts would constitute deposits entirely analogous to those issued by banks. Accordingly, EMIs –as narrow banks— would be subjected to regulation and supervision designed to mitigate operational and technology risks associated with deposit-taking and record-keeping, but would be exempted from the more burdensome prudential supervision practices since they incur no direct prudential risks.

This interpretation of EMIs as narrow banks would have several implications relative to how they are generally treated today:

- EMIs should be free to market e-money accounts as a savings product, whether or not they are labeled as banks by regulation.

- E-money accounts should be able to accrue interest, just like full bank accounts do. Of course, in practice e-money accounts will always be low yielding as the pooled funds backing the stock of e-money earn low interest because they can only be invested in safe assets.

- E-money accounts could be subject to the same deposit insurance regime and benefit from the same legal protections as full bank deposits. EMIs could contribute to the deposit insurance fund in the same way and on the same basis as full banks do. (This would need to be done in a way that avoids double-charging of the deposit insurance premium: if the EMI pays a fee on the basis of the value of its deposits, then there shouldn't be another levy on the pooled funds backing the e-money balances which are sitting in bank accounts.)

- EMIs should be able to offer credit and insurance products to their customers, but with the requirement that they be offered and funded by a separate legal entity than the entity that holds the funds backing their e-money accounts. The EMI trust or special-purpose vehicle should not be exposed to any credit or insurance risk, but the EMI could market other financial prodcuts that complement their own basic e-money offering.

In many jurisdictions where EMIs exist, they are not authorized to do one or more of the above. In the recent Peruvian e-money law, they are precluded from paying interest. In Kenya, mobile money operators are not even allowed themselves to collect the interest on pooled accounts, even to cover their own costs, and they cannot market e-money accounts as a savings product.

To reinforce that EMI accounts have the same level of rights and protections as full bank accounts, EMIs themselves could be given the designation of banks, though in a way that makes a clear distinction between them (as narrow banks) and full banks.[12] Financial inclusion ought to be more than making it easier for people to pay and receive money; providing safe methods for people to store value and manage their money would seem to be at least equally important. Therefore, it seems unnecessarily restrictive to designate institutions designed to specifically address financial inclusion gaps by the product they are mostly providing today.

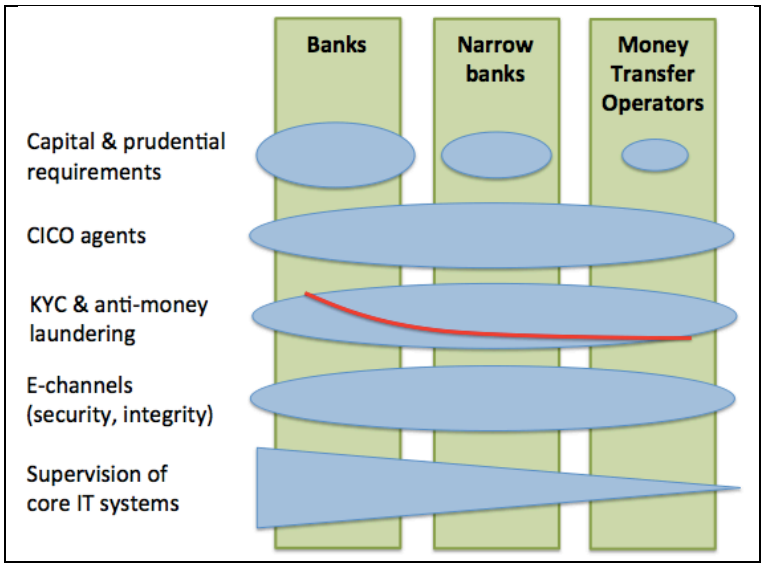
Preventing regulatory arbitrage

Whether there are two main license categories (banks and payment providers) or three as proposed here (full banks, narrow banks and payment providers), the issue then arises as to how to craft the regulations such that one category is not unduly favored over another. There of course ought to be differences in license conditions; otherwise there might as well just be one, in a return to the old days where banks alone were expected to do the whole job. The question is whether the differences are justified based on the risks raised by each type of license holder.

In general, the question is one of *quid pro quo*: a particular license holder ought to benefit from more lenient regulation than the other categories only in exchange for accepting tougher restrictions in another way which already addresses specific risks. Figure 7 shows a schematic framework for applying this principle, based on the three-license setup we proposed above. The framework is not complete, but we use it here to illustrate how it can be used across five general classes of regulatory issues:

Figure 7: Framework to guide where to differentiate regulations across different license categories

[12] Mexico is one country that has created the legal figure of a narrow bank, with substantially reduced capital requirements relative to full banks. The term proposed for EMIs in India by an RBI-appointed Committee on Comprehensive Financial Services is the "Payment Bank." Though it feels entirely appropriate to call them banks for the reasons stated above, highlighting their payment role is perhaps taking a narrow, overly prescriptive view of what commercial strategies licensees will follow.



- *Capital & prudential requirements*. Because only full banks intermediate funds and engage in maturity transformation, they present much higher prudential risks than either narrow banks or money transfer operators. Accordingly, they should be subject to higher capital requirements and more stringent prudential supervision standards. This is *not* a case of unjustified over-regulation of banks: the reason why they need to be more prudentially regulated is precisely because they are *not* subject to the specific scope-of-business limitations that narrow banks have (not to intermediate funds) or that money transfer operators have (not to issue accounts and raise open-ended deposits from the public). Because the various license types accept different restrictions on the types of operations that they can engage in, it is entirely legitimate to assign them different capital and prudential requirements. Not doing so would in fact create regulatory arbitrage opportunities. Thus, in the top row in Figure 7, we represent such regulations as three distinct bodies, decreasing in size from left to right.

- *Cash in/out agents*. All three license types require mechanisms for taking cash from the public and returning it at a later point in time or to somebody else. In other words, they all need to interface between cash and electronic value. It is not clear why we should hold a different standard of financial safety and consumer protection if you are handing cash at an agent outlet in order to deposit it into your *own* bank account, fund an e-money account so that you can transfer it on to someone else, or directly send the money to someone else. Therefore, agency regulations for cash in/out should be the same across all license categories – or at least between full banks and narrow banks (or EMIs). Not doing so –for example making it tougher for banks to sign up and manage cash in/out agents— would create regulatory arbitrage opportunities, because it would hinder banks ability to compete for value-storage services with EMIs, and for money transfer services with both EMIs and banks. Thus, in the second row in Figure 7, we represent these regulations as a single body across the three license types.

- *KYC and anti-money laundering*. To the extent that all three licenses can be used to shift money around, here again, in principle, regulations should be harmonized

across all three license categories. However, there is a point of pragmatism insofar as banks are likely to be shifting larger transactions than the other categories. This is an argument for creating a progressive KYC regime as described above; if one exists, then it can be applied across the board because the graduated KYC regime will itself be applied differently by each type of license holder. Only if there is no progressive KYC regime will there be a justification for differentiating the KYC regime across license types. But this would be suboptimal, because it would create regulatory arbitrage across license types for transactions of the same type and size. Thus, in the third row in Figure 7, we represent such regulations as a single body which incorporates the graduated KYC regime following the usage curve depicted in Figure 4.

- *Security of electronic channels*. Regulators worry mainly about three aspects of electronic transaction security: proper user authentication, data encryption and privacy. This follows a similar argumentation to the previous point. It may be justified to introduce different standards of electronic transaction security, but this should depend on the size and nature of the specific transaction involved, rather than on the license held by the provider. A given money transfer operation should be held to the same standard of electronic security, whether it is operated by a bank, EMI or money transfer operation. The overall regulatory framework should be the same, but because banks will handle much larger transactions, they will apply much higher standards *in practice*.

- *Supervision of core IT systems*. The previous point addressed security in the front-end and delivery channel, whereas here we look at the safety and soundness of the back-end information technology (IT) systems. All players will operate some kind of core IT platform and hence all should be subject to effective regulation and supervision around technology and operational risks. But the sophistication of supervisory practices will vary by license type, since some licensees are restricted from engaging in certain types of activities. Banks will have the most sophisticated IT platforms, commensurate with their broad scope of operations and risks they need to manage; mobile money operators will have the simplest since they do not manage persistent accounts on behalf of their customers. Thus, it is legitimate for the extent of supervision of these systems to be differentiated, since they present different risks both qualitatively and quantitatively. Thus, in the bottom row in Figure 7, we represent the regulations as a single but tapering body.

Preventing anti-competitive practices by dominant players

Network-based markets, of which electronic payments is one, are characterized by economies of scale both on the supply side (due to the existence of fixed costs that can be amortized over more users and transactions) and on the demand side (due to network effects, which increase the value of the service to individual users the more users are on the network). This confers strong advantages to the larger players.

Dominant players can entrench their market-leading position in network-based businesses through two main practices. First, they can refuse interconnection with smaller players, thereby undermining the commercial value of the smaller networks. Second, they can engage in *dumping*, which is the practice of lowering prices below costs in an attempt to

drive smaller players with less deep pockets out of business, at which point they can again raise prices and reap monopoly rents. We look at the rationale and potential regulatory responses to these practices below.

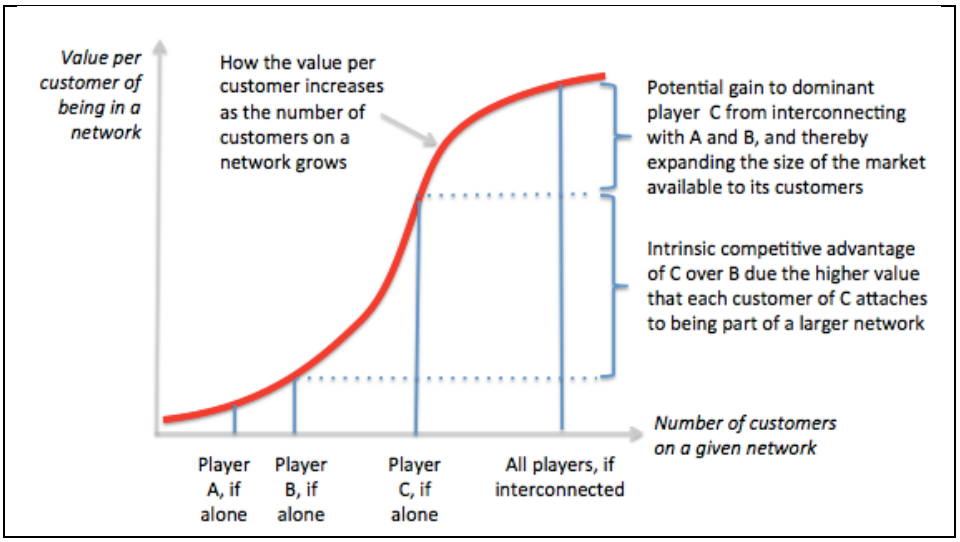
Interconnection

Early movers and larger, more established players with a marked scale advantage can therefore use their size to minimize the threat of competitive entry by keeping their customers on a closed network and thereby not sharing network effects with their competitors.

Regulators can thwart this situation by promoting interconnection of networks, by which we mean connecting the back-end systems of the various providers so that customers of one provider can communicate or transact directly with customers of a different provider. There is a strong economic justification for seeking to spread network effects across all players in the industry by interconnecting them. First, by pooling each providers´ customers into a single interconnected market, total-market network effects are maximized. Thus, in a static sense, each player could benefit from interconnection, as long as the total net gains from creating market-level network effects are shared in an economically fair manner across all the providers. Second, a market in which all players are interconnected creates more room for new entrants to contest the market, as they can attach themselves to the market-level network effects. Thus, requiring interconnection can help to reduce the barriers to entry that are inherent in network-based businesses.

It is always hard for competitors to decide to work together on some key aspects of their business. It usually comes down to whether the players involved opt to maximize the total size of the pie by joining up their network effects, or just their slice of the pie by insulating their customers from their competitors. Figure 8 shows the two opposing considerations that a dominant player (labeled C in the Figure) faces. On the one hand, by not interconnecting with its smaller competitors it can lock in more value per customer than its competitors can, and with that extract higher profit or undercut the pricing of competitors. On the other hand, by interconnecting it would create additional value for its existing customers, who could benefit from being exposed on the same network to the customers of the other players.

Figure 8: How interconnection can maximize market-level network effects, while potentially undermining the larger player´s dominance



Thus, interconnection agreements may or may not occur through the normal competitive interplay. There is therefore a justifiable role for regulators and policymakers to find ways to move market players on a path towards interconnection.

However, the way in which interconnection occurs –how the interconnection service is defined, how it is priced, how the technical network interfaces are specified, what specific shared infrastructures are used to implement interconnection— can have significant consequences for market development. Hasty or injudicious regulatory impositions on interconnection can kill off incentives for investment, innovation and growth, especially if it is mandated at an early stage of development of a market when there is still substantial uncertainty about customer value, business models and cost of service provision.

Regulators are therefore well advised to take a cautious but determined approach to interconnection. They should always rely on market negotiations between players in the first instance, though they should find ways to cajole players so that fair market arrangements emerge. Potential tools for the regulator might be:

- Actively promoting industry fora that help frame industry discussions on interconnection.
- Unambiguously stating a target date by when dominant players must have achieved interconnection with smaller players. This could be supported by establishing a clear process and formula for designating those players who are deemed to have "significant market power." (This requires first defining the market, and then specifying the market share threshold that constitutes significance.)
- Arming themselves with the legal power to intervene and arbitrate in situations where a player designated as having significant market power is denying interconnection or offering unduly harsh interconnection terms to smaller players.

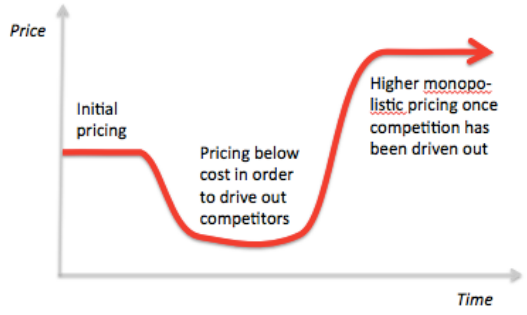
Anti-dumping

Another way in which dominant players can exploit their scale is through their relatively stronger financial –rather than necessarily market— position. If they believe that they are able to sustain bigger losses than their competitors before going bankrupt, they might force

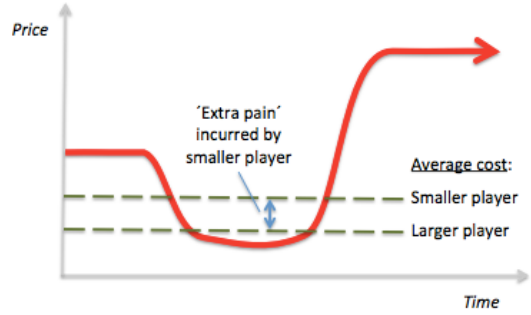
the entire industry –themselves and their competitors—into severe losses by cutting prices below costs. They might do that in the hope of driving their less well-funded competitors out of business; when that happens, the deeper-pocketed player can then raise prices to a higher level than at the beginning due to the then-prevailing lack of competition, thereby more than recouping the losses it incurred during the price war. The practice of dumping is illustrated schematically in the left-hand panel of Figure 9.

Figure 9: How dumping works

(a) Reducing prices below cost to drive out financially weaker competitors



(b) Scale-driven cost advantages favor dominant players engaging in dumping



In industries where fixed costs are large, such as in network-based payment systems, a dumping strategy will in fact hurt the larger players less than the smaller ones on a per-unit basis. That is because larger players can amortize fixed costs over more users and transactions, and hence typically have lower average unit costs than smaller players. This is illustrated by the distance between the two horizontal lines in the right-hand panel in Figure 9. Therefore, dominant players have a greater capacity to engage in dumping behavior, both because of their greater financial muscle and their lower unit cost base.

There is a long tradition in competition policy of making dumping illicit, and competition authorities need to watch dominant operators carefully to ensure they do not incur in such injurious pricing behavior.

While there are several examples of price wars having occurred in the mobile money arena, they have usually been triggered by smaller competitors eager to gain market share, rather than by larger players wishing to drive out competition.

Preventing mobile operators´ abuse of bottleneck service elements under their exclusive control

From a telecoms regulation point of view, mobile money is another instance of a value added service and those tend to receive a light regulatory treatment. All the specific regulations that pertain to the safety and soundness of all forms of mobile money –who can issue accounts, conditions of service, data security and privacy standards, supervisory treatment, consumer protections, etc.— should be the domain of the banking regulator. The division of responsibilities between the two regulators is crystal clear.

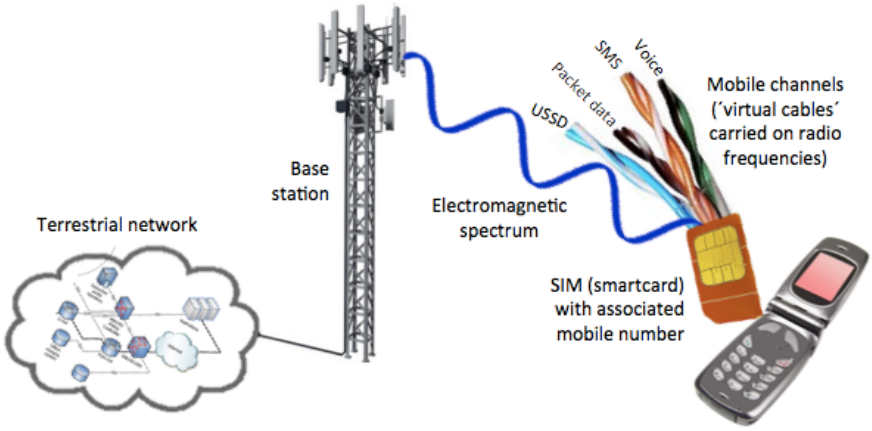
But mobile operators´ participation in retail payments presents competitive challenges which banking and telecoms operators will need to monitor closely and address jointly. The

problem is that mobile operators are both component suppliers and direct competitors to financial institutions wanting to offer mobile financial services. There is a risk that mobile operators transfer market power from their core telecoms market to the emerging retail mobile payments market, in such a way as to effectively shut banks out of mobile payments. In the future, most financial services can be expected to have a mobile component, so such a situation would have severe implications for competition in the financial inclusion space.

This is quite understandably spooking many banking regulators into preventing mobile operators from playing directly in the payments/banking space. That's unfortunate; mobile operators should be encouraged as much as any other type of organizations to push the frontiers of financial access, while guarding against any potential abuses of market power on their side. We need to allow mobile operators to contest the mobile financial services market without dominating it.

To begin with, authorities need to identify specifically those components of mobile communication services over which mobile operators have bottleneck control and which are essential for the provision of mobile financial services. Figure 10 shows the service elements of a mobile communications network, focusing on the elements on the customer access side, since those are more unique to mobile operators and have greater potential to create bottlenecks because they are more costly to duplicate.

Figure 10: Key service elements of a mobile telecoms network



National telecoms policies empower mobile operators as stewards of two scarce public resources: designated parts of the electromagnetic (or radio) spectrum and the unique telecoms numbering range. Mobile operators' delegated control over these resources give them an exclusive ability to create and manage mobile telephony and data connectivity services. It is then generally left to national telecoms regulators, in conjunction with competition authorities and the courts, to foment the development of a competitive telecoms market despite the inherent lack of contestability of the market.

If mobile operators are not subjected to effective competition policies, they can impact the development of mobile financial services in two ways. If, as a result of monopolizing tendencies, mobile connectivity services are generally expensive and of low quality, it will affect financial institutions as much as anyone else using mobile services. Addressing this issue is squarely the responsibility of the telecoms regulator, and is beyond the scope of this

paper. But even if that´s not the case, mobile operators may choose to specifically withhold service from or offer discriminatory terms to financial institutions, if the mobile operator competes with them at the financial services layer. Addressing this issue will require close coordination between the banking and telecoms authorities.

It is therefore entirely appropriate to place certain specific restrictions on mobile operators wishing to offer any kind of regulated financial service, as part of their financial license conditions. These restrictions could potentially relate to non-discriminatory (or *equal*) access to mobile connectivity services, and the use of mobile numbers as a transaction routing mechanism. We look at each of these issues in turn below.

Equal access rules on mobile communication channels

Mobile operators with access to portions of the electromagnetic spectrum can provide certain communications (or bearer) channels which any financial service provider that wants to offer mobile financial services over wide-area cellular networks must use. It is because of this near-absolute dependence of financial institutions on mobile operators´ communications services that we refer to mobile channels as an essential or *bottleneck* service element.

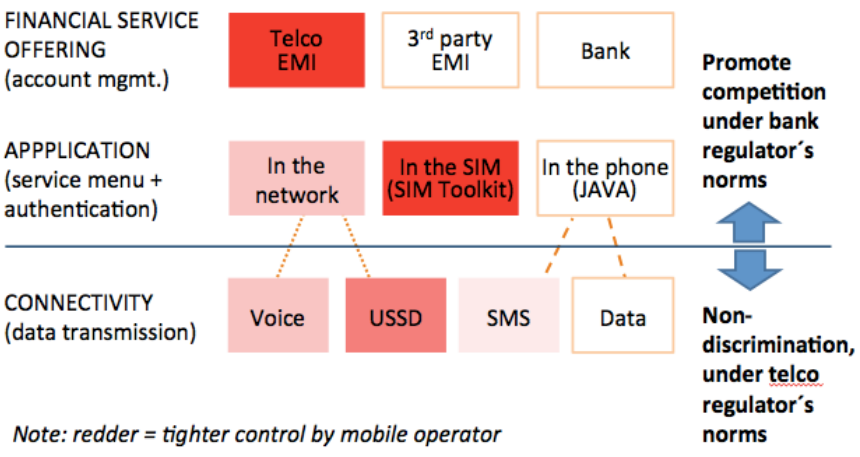
Consider the options that a financial institution would have if it was denied access to mobile channels or wished to circumvent mobile operator´s services. One set of options would be for the financial institution to buy spectrum and become a mobile network operator, or to buy wholesale access to mobile channels and become a mobile virtual network operator (MVNO) such as Equity Bank is doing in Kenya. However, these may not be possible if there is no spectrum on offer by the government, and no local mobile operator is willing to provide MVNO host services. These options also represent daunting increases in business scope and require deep pockets. Another set of options would be to use non-cellular wireless protocols, such as satellite (which is very expensive), wifi (which does not offer ubiquitous coverage), or other wide-area wireless technologies running on unlicensed spectrum (which generally operate at high frequencies with poor indoor coverage and lack scale in the distribution of terminals). These options would clearly handicap the financial institution relative to competitors who do have access to cellular channels.

It is therefore important to ensure that mobile operators do not abuse their control of such an essential service element. This becomes a distinct concern when a mobile operator itself wants to offer mobile financial services, since in that case financial institutions are both customers and competitors of the mobile operator. Figure 11 shows a simple three-level value chain for mobile financial services. Mobile operators that offer mobile money services under an EMI license and other financial institutions are competitors at the top (financial service provision) layer. But financial institutions need to buy access to mobile connectivity services represented in the bottom layer of Figure 11from those same telcos.

A mobile operator offering mobile money services then faces an inherent business conflict: whether to support its core mobile connectivity offering by selling the necessary bearer services to banks, or whether to impede effective competition to its new financial services offering by denying access to, degrading the quality or increasing the pricing of the bearer services used by other financial institutions.

The extent of this conflict depends on the nature of the mobile channel, and in particular on how widespread the use of the channel is commercially and how tied it is to a particular application environment under the control of the mobile operator (middle layer in Figure 11).

Figure 11: Managing bank-telco co-opetition



The more common channels, such as voice, SMS (the protocol underlying text messaging) and packet-data (under various flavors such as GPRS, EDGDE, 3G or HSDPA), are broadly made available by mobile operators under standard commercial offers. In this case, it should be fairly easy to establish and monitor a requirement of non-discriminatory access by banks and third-party providers of mobile money services to these channels. Since the market for these bearer services is sufficiently large and lucrative for the mobile operators, we can expect them not to over-price their voice and SMS services or to degrade the quality of their service specifically to lock out financial institutions from using them to construct competing mobile money services.

The challenge is with a less common bearer service called USSD (Unstructured Supplementary Service Data), which is not widely commercialized by most operators. (You may recognize it as the service you use when you are asked to dial a sequence of numbers starting with star and ending with hash after you buy a scratch card.) The session-based nature of USSD presents two strong advantages over SMS as a channel for mobile financial services: it lends itself to implementing network-based menus which makes it easier to use, and it entails no storage of messages anywhere which makes it more secure. It may also be more feasible than using voice in countries where voice tariffs are still expensive, or packet-data in countries where most people still use simple phones. For banks without access to the SIM card, USSD may be the only realistic option. In this case, simply stating non-discriminatory access to the USSD service may not suffice as long as USSD is used primarily for banking services, since MNOs may price it specifically to preclude competition from emerging in mobile financial services.

Because of the risk of anti-competitive behavior by mobile operators, they should only be licensed as EMIs if there is a policy framework in place that precludes them from abusing their bottleneck control over mobile channels (and especially USSD) for anti-competitive purposes.

This is not something that the banking regulator can do on its own, as promoting the competitive supply of bearer services is the domain of the telco regulator. Therefore, the banking and telecoms regulators must work out an integrated regulatory regime for mobile operators that offer mobile money services *before* the banking regulator grants EMI licenses to mobile operators to enter the financial services space.

Telecoms regulations need to ensure that mobile operators who themselves are engaged in mobile financial services are obliged to offer mobile communications services to any other financial institution that requests access, on non-discriminatory and cost-oriented terms. This means that such mobile operators: *(i)* cannot deny service to any financial institutions; *(ii)* must offer mobile services to their in-house or related EMI on the same terms as to any other financial institution; and *(iii)* must price mobile services to financial institutions on the basis of cost plus a reasonable profit mark-up.

Peru has spearheaded this approach. The new e-money law passed in 2013 required that mobile operators had to offer all EMIs access to mobile channels on a non-discriminatory basis, and mandated the telecoms regulator, OSIPTEL, to issue regulations to safeguard this principle. Accordingly, OSIPTEL has now taken the internationally unprecedented step of regulating EMIs' terms of access to mobile channels.[13] Unfortunately, this regulatorily-enforced policy of non-discriminatory is only extended to EMIs, not to banks, which creates the kind of regulatory arbitrage situation we discussed in an earlier section.

Following the Peruvian example, non-discrimination can be enforced by requiring that all contracts between mobile operators and financial institutions be cleared in advance by the telecoms regulator, and that all contracts be made public. If a financial institution and a mobile operator fail to agree on access terms within two months, the financial institution can request the telecoms operator to set the terms based on other existing contracts and an analysis of costs involved. Maximum periods for implementation of new contracts also need to be set by regulators.

The Peruvian approach is to let the market determine pricing in the first instance, though the telecoms regulator is prepared to make a determination in case of a dispute or failure to reach agreement between the parties. The telecom regulator´s declared willingness to step in is designed to enhance the bargaining power of financial institutions in front of mobile operators who are their essential suppliers. The Indian telecoms regulator, TRAI, is considering taking a more direct approach and prescribing a price for USSD.[14] Though in the end the result may be similar to that reached under the Peruvian approach, it does seem more appropriate for the regulator to give commercial negotiations a chance before intervening.

---

[13] Eletronic Money Law No. 29985 was approved by Parliament in January 2013, and the implementing regulations were issued by the Ministry of Finance under Supreme Decree 090 in May 2013. OSIPTEL´s regulation on e-money issuers' access to telecommunications services were issued in September 2013 under Resolution 126.

[14] See TRAI´s "Consultation Paper on USSD-based Mobile Banking Services for Financial Inclusion," dated 20 September 2013, available here.

This type of coordinated pro-competitive vigilance by the financial and telecoms regulators will be essential until such time when everyone has a smartphone and the phone can perform locally all security and menu presentation services which are now under the control of the mobile operator (through the USSD channel or an application in the SIM). At that point, banks will have credible options to build their own mobile financial services with minimal control by the mobile operators.

Equal access rules on the SIM card

Another asset mobile operators control uniquely is the SIM (Subscriber Identity Module) – a smartcard with a very small form factor, which fits into a slot on all GSM phones. The SIM uniquely identifies every mobile user, and links that user to his mobile phone number. Only the mobile operator has access to the applications and the memory embedded within the chip in the SIM. Mobile operators view the SIM technically as the end-point of their mobile communication infrastructure. In fact, in many countries the SIM is legally the property of the mobile operator, not the user that keeps it in his phone.

Beyond customer authentication, access to the SIM provides benefits in terms of data transmission security, since SIMs can contain pre-loaded security keys that can implement end-to-end data encryption from the mobile handset all the way to the transaction authorization server. Access to the SIM might also enhance the usability of services, since the SIM controls the on-the-phone menu onto which mobile money can be incorporated directly. Thus, access to the SIM would make it easier for any financial service providers to develop secure applications that operate locally on any mobile phone.

While some operators do "rent out" space on their SIMs to banks and other financial institutions on purely commercial terms, there is no precedent worldwide for establishing equal access rules to the application and memory space within the SIM. While having access to the SIM confers some clear advantages, it is not the only way of offering mobile financial services, since the application can reside on the network (accessible through voice or USSD channels) or on the phone (in the form of a downloadable Java app) instead of on the SIM (these options are depicted in the middle layer of Figure 11). Thus, the SIM does not fully constitute an essential service.

Thus, regulators can mitigate operators´ stark competitive advantage in deploying financial services on their proprietary SIMs by ensuring that financial institutions have fair access to appropriate mobile channels such as USSD that do not require special access to the SIM in order to create an acceptable security and usability framework.

Bottleneck control over the numbering range

A third element under the strict control of mobile operators is the unique mobile number (in technical parlance, MSISDN) that they assign to every SIM they issue. Their uniqueness stems from the existence of a national numbering plan, which is usually designed and managed by the telecoms regulator. Mobile numbers have become a quasi-universal unique identifier of individuals, one that people are familiar with. They can therefore be very useful to route money transfers as an alternative to longer bank account numbers which people largely do not remember.

Mobile operators could in theory leverage their control over mobile numbers to benefit their mobile money operation by bundling various types of transactions that are addressed to the same phone number. Product bundling occurs when two or more distinct services need to be purchased together, even if the customer wishes to buy only one.

In most mobile money schemes, there are two common types of money transfers. The first one is an entirely electronic transaction, where value moves from one digital account to another. A second form of money transfer is used to send money to those who do not have an electronic account, and consists of sending a unique alphanumeric code (technically: a one-time password) to the recipient, typically by SMS. The recipient can collect the cash at any agent of the provider on presentation of the code.

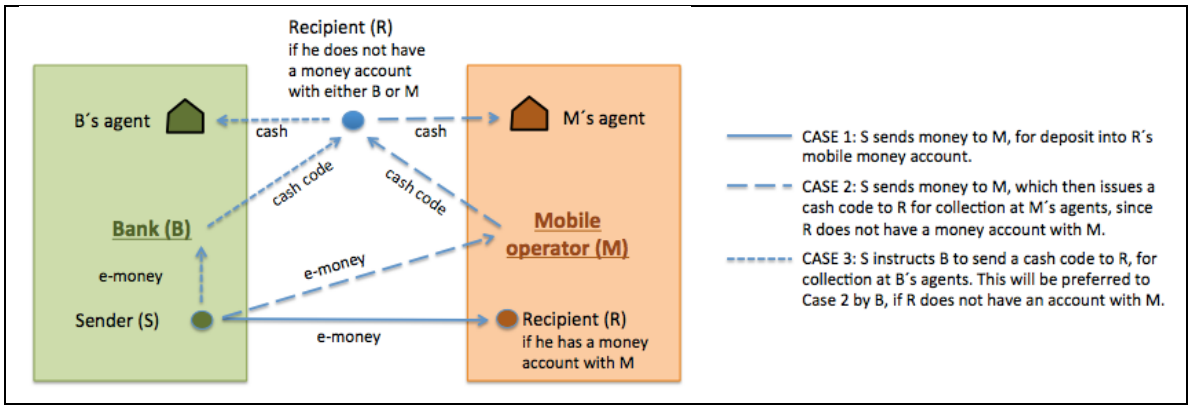
These two services are bundled whenever the sender has no control over whether a money transfer will take one form or another. To the extent that the two services are priced differently to the sender (as is the case with M-PESA in Kenya), the sender cannot be certain of the price it will be charged for a given money transfer unless he knows whether the recipient has an account or not; the sender cannot issue a "pay electronically only" instruction. This bundling of the two services may not be too much of a problem among purely retail customers of the mobile money service. But it will create competition issues in the wholesale market.

Suppose that a customer S (for sender) of bank B wants to send money to a non-customer R (for recipient) by addressing it to R´s phone number which belongs to mobile operator M. Suppose, further, that M operates its own mobile money scheme under which it bundles the two services mentioned above. Bank B has no way of knowing whether R has a mobile money account with M or not; when B passes its customer R´s payment instruction to M, M will pay electronically into R´s mobile account if R has one, otherwise it will instruct R to collect the cash at one of M´s agents.

The problem is that bank B would want the money transfer to be processed by M only if R has a mobile money account with M. If it does not, then B would want to issue its own cash code to R via a simple text message, which R can claim at any agent of R – not M. M is in effect using its control over its mobile money accounts to favor its own agents. M´s bundling strategy forces B to have to concede the over-the-counter cash collection business to M, at the detriment of B´s own agents, if it wishes to interconnect with M at all. To put it slightly differently, because of M´s service bundling, B has the uncomfortable choice of doing the right thing by its customers by letting them send money to mobile money customers of M, or doing the right thing by its agents and always sending a cash code to anyone who is not a customer of B.

Figure 12 represents the situation schematically. Bank B would like to handle the transaction through case 1 or 3, depending on whether R has a mobile money account with M or not, but mobile operator M bundles cases 1 and 2.

Figure 12: Handling of transactions sent to the phone number of a non-customer



The proper policy prescription, especially if M is deemed to be a player with significant market power, would be to force M to unbundle the two money transfer products, such that sending networks have the choice of using their electronic money transfer service (to digital accounts) or their cash-by-code service (for collection at agents), or both. M should be required to notify sending networks whether or not a mobile money account is associated with the recipient´s phone number.

So would this embolden the innovator-entrepreneur?

Let´s now go back to the entrepreneur we referred to at the beginning of this paper, who has an innovative idea for a money management application that he believes better meets the financial needs of the informal majority in his country. If he faced the type of regulatory regime proposed above, would he have a better chance of testing and potentially rolling out his service than if he faced the current best practice regulatory regime? There is good reason to think that he would.

He would have a choice of financial licenses that more closely matched what he wanted to do. In particular, he would be able to get a "narrow banking" license which would exempt him from tough capital requirements but without restricting his ability to reward his customers with interest, bundling in a credit offer, or marketing the service as savings.

He wouldn´t have to spend so much effort complying with KYC requirements for new customers who he is not sure if or when they will use his service much. He can defer such KYC costs to such time when his customers are willing to commit to the service.

He would be able to benefit from all existing cash in/out networks, without having to build his own. All he would have to do is to convince each independent agent network manager to become a corporate customer of his narrow bank. He might sweeten the deal to get their attention by offering them monetary incentives, but he wouldn´t have to worry about any of the logistics involved in getting his brand of electronic money exchangeable everywhere where there are agents.

He would draw comfort from knowing that he can expect a competitive level playing field. He wouldn´t have to worry about some competitors being unduly able to do what he does under more favorable regulatory terms, simply because they have a different type of license than his. He also wouldn´t have to worry about the larger, more established competitors taking unfair advantage of their size, because the financial regulator´s stance would

strengthen his hand when he is seeking to negotiate interconnection on fair terms with them, and their pricing actions would be under close regulatory scrutiny. And he wouldn´t have to worry about large mobile operators squeezing him on the provision of telecoms services in order to contain the competitive threat he poses to their mobile money services, because the telecoms regulator would enforce non-discrimination.

None of this would of course guarantee his success. But it would probably go a long way to embolden him to try his idea in an incremental, controlled way. And this –encouraging smaller, more innovative players to try their hand at finding new ways of financially including people— should be a part of regulators´ policy objectives, alongside the preservation of system stability and safety of customers´ funds. The goal of financial inclusion will not be achieved without more competition and innovation.