Monetary Theory and Electronic Money: Reflections on the Kenyan Experience

William Jack, Tavneet Suri, and Robert Townsend

I n 2007, the leading cell phone company in Kenya, Safaricom Ltd., launched M-PESA, a short message service (SMS)-based money transfer system that allows individuals to deposit, send, and withdraw funds from a virtual account on their cell phones and that is separate from the banking system. M-PESA has grown rapidly, currently reaching more than seven million users, approximately 38 percent of Kenya's adult population, and it is widely viewed as a success story to be emulated across the developing world. Indeed, similar products have recently been launched in a growing number of countries across Africa, Asia, and Latin America, with the intent of expanding financial services to previously unreached populations.[1]

M-PESA is used not only for remittance purposes, but also to save, to purchase pre-paid phone credit and other goods and services, to pay bills, and to execute bank account transactions. However, consumers do not need bank accounts in order to use M-PESA, and Jack and Suri (2009) found it was used by more than half of the unbanked in their sample. It is used by a broad cross section of Kenyan society, but has increasingly been adopted by those at the

[1] For example, M-PESA is similar to the service called G-CASH in the Philippines.

lower end of the income distribution, as is evidenced by the steady reduction in the average transaction size since its inception. A large part of M-PESA's success is attributed to the broad and dense network of over 16,000 agents across Kenya, which provides the retail interface with consumers.

In this article, we examine the role of monetary theory in understanding this new generation of mobile banking products, especially those that, like M-PESA, do not simply provide electronic access to existing bank accounts. Deposits of money in a mobile phone-based account reflect holdings by the account owner of a commodity we refer to as e-money. Because e-money can be easily transferred from one individual to another, as long as it is expected to retain its value, it can be used in equilibrium as a means of exchange, as well as to transfer purchasing power between individuals.

Indeed, at the time of the launch of M-PESA, the service was seen as a means of overcoming the high transactions costs associated with sending cash remittances that faced the 80 percent of individuals in the economy without bank accounts.[2] But since then, e-money has been increasingly used both as a store of value and as a means of exchange, with users able to pay utility bills, make loan repayments, and even pay for taxi rides with it. The co-existence of essentially two forms of cash, even if closely related and linked, raises certain theoretical modeling issues in itself. But when one form of cash is issued by a profit-maximizing entity and the other by the central bank, further issues of competition, regulation, and coordination naturally emerge.

Although mobile banking is in its infancy in the United States, payroll cards have provided a similar payment function, albeit without the geographic reach of mobile phone communication. Funds are typically deposited by an employer into the account of an employee, who can either withdraw cash at an automated teller machine (ATM) or use the card to make purchases at stores possessing debit card machines. As in the case of mobile banking, payroll card users do not need a bank account, and transactions are executed using an existing communication network. In addition, payroll cards in the United States are generally cheaper than check-cashing services and money orders, just as M-PESA in Kenya is cheaper than most alternatives. Foster et al. (2010) describe the use of various payments systems in the United States— they find that 93.4 percent of consumers in the United States have adopted a payment card, but only 17.2 percent have a prepaid card.

This article presents a first look at how existing models of monetary theory can be used to think about the impact of mobile banking on the operations of the financial system and the implications for monetary and regulatory policy

---

[2] The original pilot program, supported by the U.K. government and the mobile phone provider Vodafone, was aimed at increasing the efficiency of microfinance products by allowing borrowers to make repayments more easily. However, by the time of the full launch, the focus had shifted to facilitating the sending of remittances more generally.

decisions that face the central bank. We are not yet in a position to develop a fully articulated model of mobile banking, but we hope this discussion will be a first step in this process. In addition, this article is not an exhaustive discussion of all models of money, but more of a focus on a subset of models that have different implications for the role of e-money in an economy.

Most theoretical models of money and credit include both a temporal dimension and some kind of generalized locational heterogeneity. Sequential trades over time require promises to be made (and kept) and records to be maintained. On the other hand, spatial separation can mean that it is not always possible for two parties to a trade to meet each other at the right time, so more complicated multilateral chains of individuals are required to effect the desired net trades.

In these environments, financial instruments such as fiat money and private debt can sometimes improve the efficiency of resource allocations by facilitating intertemporal and interspatial trades. However, equilibrium allocations may continue to be inefficient without the intervention of either a public institution (such as a central bank) or a well-regulated private agent (such as a clearinghouse).

Mobile banking has the potential to effectively reduce the distances that separate individuals, both literally and figuratively, thereby lessening the frictions that characterize models of incomplete intermediation, relaxing liquidity constraints, and reducing the need for monetary interventions. On the other hand, new liquidity constraints could arise that are binding for individuals who trade with the new financial instrument, e-money.

The article proceeds as follows: Section 1, which draws heavily on Jack and Suri (2009), provides background on the recent evolution of mobile technology and mobile banking in Kenya and on the practical operational features of M-PESA. Section 2 reviews a number of strands of the literature and discusses the specific lessons that we might learn regarding both the equilibrium impact of mobile banking and its implications for policy. Section 3 presents some empirical facts from a survey on M-PESA customers and agents that provide some insights into the implications from the models and lessons in Section 2. Section 4 concludes.

1. BACKGROUND ON M-PESA

Mobile Money in Kenya: An Introduction

Mobile phone technology has reduced communication costs in many parts of the developing world from prohibitive levels to amounts that are, in comparison, virtually trivial. Nowhere has this transformation been as acute as in sub-Saharan Africa, where networks of both fixed line communication and physical transportation infrastructure are often inadequate, unreliable, and dilapidated. While mobile phone calling rates remain high by world standards,

Figure 1 Phone Use in Kenya
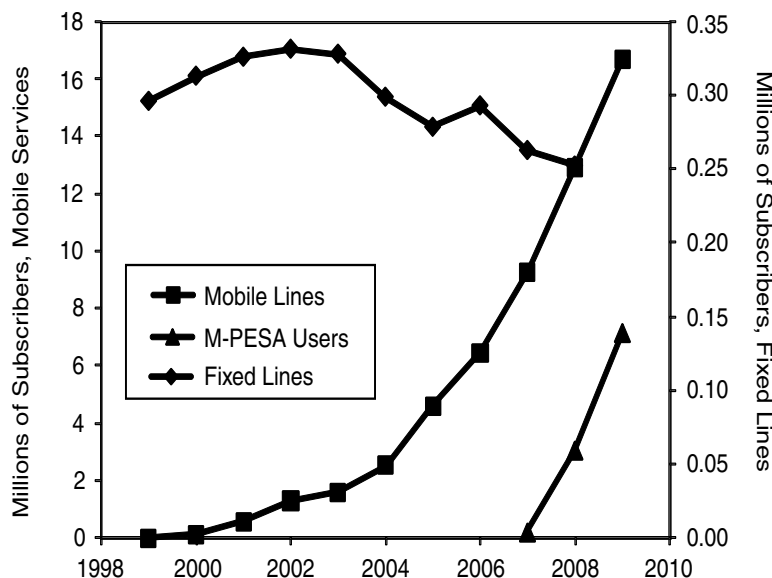
the technology has allowed millions of Africans to leap-frog the landline en route to 21st century connectivity. As the number of landlines in Kenya fell from about 300,000 in 1999 to around 250,000 by 2008, mobile phone subscriptions increased from virtually zero to nearly 17 million over the same time period (Figure 1). Assuming an individual has at most one cell phone,[3] 47 percent of the population, or fully 83 percent of the population 15 years and older, have access to mobile phone technology. In March 2007, following a donor-funded pilot project, Safaricom launched a new mobile phone-based payment and money transfer service, known as M-PESA.[4] The service allows users to deposit money into accounts linked to their cell phones, to send balances using SMS technology to other users (including sellers of goods and services), and to redeem deposits for regular money. Charges, deducted from users' accounts, are levied when e-money is sent and when cash is withdrawn.[5]

---

[3] This is not quite true, as some individuals own two (or more) phones so as to take advantage of the different tariff policies of competing providers.

[4] Pesa is Kiswahili for "money"—hence M[obile]-Money. A second mobile banking service called ZAP has since been launched, operated by Zain, the second largest mobile phone operator in Kenya. ZAP's market share remains very small at this point in time.

[5] The marginal cost of depositing and sending money is very low. These fees cover the costs of maintaining and expanding the agent network and physical infrastructure, marketing, and profits.

Figure 2 **Total Net Tariff Rates for Depositing and Sending Money by the Post Office and by M-PESA to a Registered User and to a Nonregistered User**
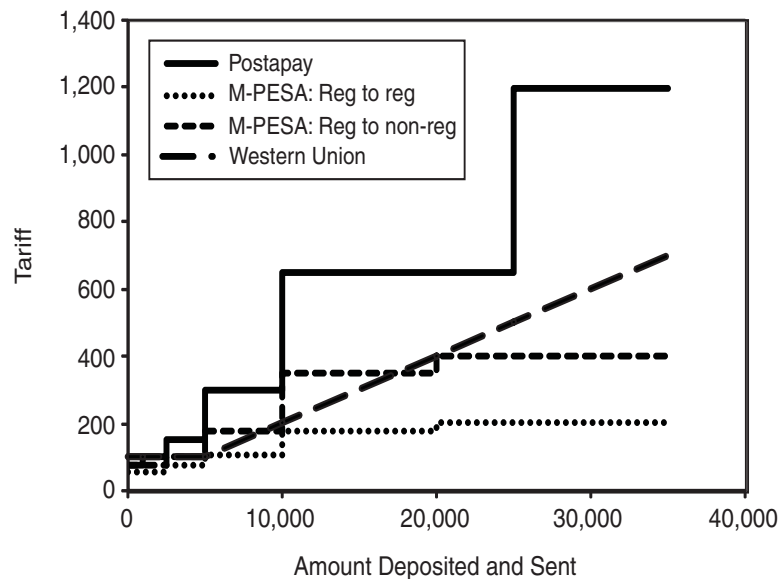
In particular, Safaricom accepts deposits of cash from customers with a Safaricom cell phone SIM (subscriber identity module) card and who have registered with Safaricom as M-PESA users. Registration is simple, requiring an official form of identification (typically the national ID card held by all Kenyans, or a passport) but none of the other validation documents that are typically necessary when a bank account is opened. Formally, in exchange for cash deposits, Safaricom issues a commodity known as "e-money," measured in the same units as money (denominated in shillings), which is held in an account under the user's name. This account is operated and managed by M-PESA and records the quantity of e-money owned by a customer at a given time. There is no charge to a customer for depositing funds into his account, but a sliding tariff is levied on withdrawals from M-PESA accounts (for example, the cost of withdrawing $100 is about $1).[6] An M-PESA user who sends e-money is charged a flat fee of about 40 U.S. cents if sending to another registered user, and a sliding fee if sending to a phone number that

---

[6] The complete tariff schedule is available at http://www.safaricom.co.ke/fileadmin/template/main/downloads/Mpesa_forms/14th%20Tariff%20Poster%20new.pdf

Figure 3 Average Daily Growth in M-PESA Registrations by Month
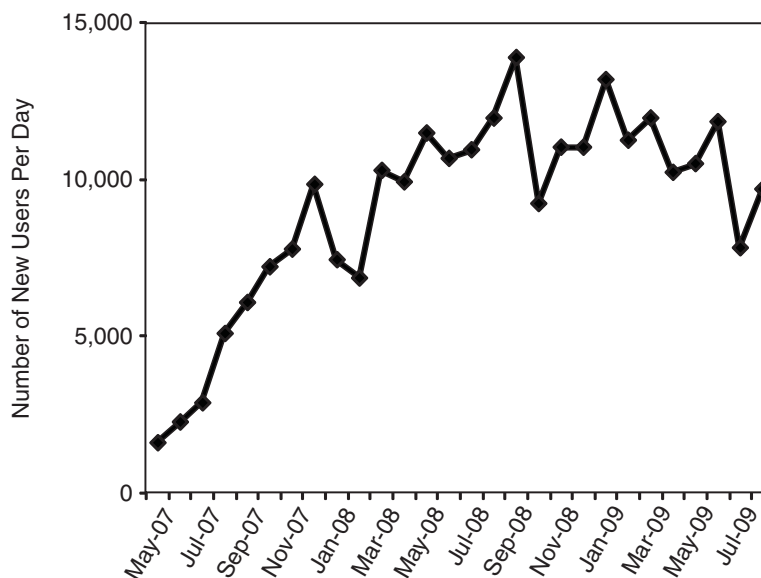
is not registered for M-PESA.[7] Figure 2 illustrates the schedules of total net tariffs for sending money by M-PESA, including the cost of withdrawing the funds incurred by the recipient, and compares these with two other money transfer services—Western Union and Postapay (operated by the Post Office). The M-PESA tariffs shown include both the sending and withdrawal fees and are differentiated according to receipt by registered and nonregistered user. Fees are charged to the user's account, from which e-money is deducted. Additional cash fees are officially not permitted, but there is evidence that they are sometimes charged on an informal basis by agents. We return to this issue below.

E-money can be transferred from one customer's M-PESA account to another's using SMS technology or sold back to Safaricom in exchange for money. Originally, transfers of e-money sent from one user to another were expected primarily to reflect unrequited, internal, within-country remittances, but nowadays, while remittances are still an important use of M-PESA,

---

[7] Nonregistered individuals can receive money sent by a registered user as long as they have a cell phone. The recipient receives a text message with a code that can be taken to an M-PESA agent who provides the cash less any fees. The fee schedule is designed so as to encourage recipients to register. Note that a nonregistered user cannot send e-money to a third individual from his phone.

Figure 4 Average Transaction Size (Kenyan Shillings): Moving Down-Market
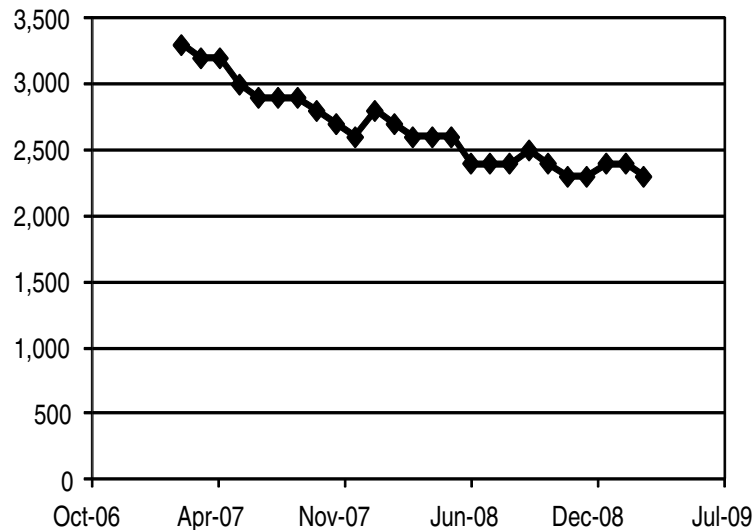
Notes: The data come directly from Safaricom.

e-money transfers are often used to pay directly for goods and services, from school fees to the wages of domestic staff.[8]

The Growth of Mobile Money

M-PESA has spread quickly and has become one of the most successful mobile phone-based financial services in the world.[9] The average number of people opening up an M-PESA account (i.e., new registrations) per day exceeded 5,000 in August 2007 and reached nearly 10,000 in December that year (see Figure 3). By August 2009, 7.7 million M-PESA accounts had been registered. Ignoring multiple accounts and those held by foreigners, this means 38 percent

---

[8] Transactions are not always limited to innocent trades. For example, there are reports of people using M-PESA to pay bribes to traffic police. Even worse, rumors have circulated in Nairobi that kidnappers are requesting ransom to be paid by M-PESA, although these rumors have not been confirmed.

[9] Similar services in Tanzania and South Africa, for example, have penetrated the market much less. See Mas and Morawczynski (2009).

Table 1 What Do Individuals Use M-PESA For?

	Fraction of Sample (Based on Multiple Responses)
Receive Money	28.40%
Send Money	25.08%
Store/Save Money for Everyday Use	14.39%
Buy Airtime for Myself	13.58%
Buy Airtime for Someone Else	8.30%
Store/Save Money for Emergencies	6.69%
Store/Save Money for Unusually Large Purchases	0.27%
Pay Bills	1.35%
Receive Money for a Bill/Else Pay Bills	0.77%

Notes: Each entry is the share of registered M-PESA users in our sample who reported the corresponding function to be the most commonly used. The bill payment service had only just started at the time of the survey and has since become rather popular.

of the adult population of Kenya had gained access to M-PESA in just over two years.

Since the launch of M-PESA, wary of regulation by the Central Bank of Kenya, Safaricom has been at pains to stress that M-PESA is not a bank. However, the ubiquity of the cell phone across both urban and rural parts of the country, and the lack of penetration of regular banking services,[10] led to hopes that M-PESA accounts could substitute for bank accounts and reach the unbanked population. Data reported in Jack and Suri (2009) suggest this is partially true, although M-PESA has been adopted by both the banked and unbanked in roughly equal proportions.[11] In addition, more recently, M-PESA users have been able to withdraw funds from their M-PESA accounts at ATMs operated by one of the commercial banks (Equity Bank) and some banks have begun to use M-PESA as their mobile banking platform. However, deposits cannot (yet) be made at ATMs, and the network of ATMs and bank branches, while growing, remain limited: In the long run they could replace agents, but both capital costs and the costs of security, operation, and maintenance suggest agents will continue to play an important role for some time.[12]

---

[10] In 2006 it was estimated that 18.9 percent of Kenyan adults used a bank account or insurance product, and by 2009 this had increased to 22.6 percent (see Financial Sector Deepening, Finaccess I).

[11] These data are from a survey fielded in late 2008. Since then, there has been some growth in the number of individuals and households with a bank account because of the expansion of such institutions as Equity and Family Bank.

[12] In 2003 there were 230 ATMs in Kenya (see Central Bank of Kenya [2003] at http://www.centralbank.go.ke/downloads/nps/nps%20old/psk.pdf). Recent data suggest there are around 1,200.

Table 2 Daily Financial Transactions, Oct. 2007–Sept. 2008

	RTGS	ACH	ATM	Mobile
Value Per Day (billion KShs)	66.3	8.5	1.0	0.1
Transactions Per Day (thousands)	1.0	39.2	180.2	107.2
Value Per Transaction (million KShs)	64.67	0.216	0.006	0.003

Notes: KShs = Kenyan Shillings.

Source: Central Bank of Kenya (2009).

The average size of M-PESA transactions has fallen over time as it has reached more of the population and has been used more extensively, as shown in Figure 4. In the two years following its introduction, the average transaction size fell about 30 percent, having started at KShs 3,300 (about $50). Most of this decline has probably been because of the expansion of take-up among the poorer individuals and households.

Table 1 shows the various types of transactions for which M-PESA is used, which include not just sending and receiving money, but also storing or saving money, purchasing airtime (the prepaid credit used for voice and text communications), and paying bills.

While the sustained growth in M-PESA registrations is notable, the volume of financial transactions mediated through M-PESA should not be exaggerated. Table 2 reports that the volume of transactions effected between banks under the RTGS (Real Time Gross Settlement) method is nearly 700 times the daily value transacted through M-PESA; and, maybe more relevant, the daily value transacted through the check system (automated clearinghouse, or ACH) is about 85 times the daily value transacted through M-PESA. Related, the average mobile transaction is about 100 times smaller than the average check transaction (ACH) and just half the size of the average ATM transaction.[13] M-PESA is not designed to replace all payment mechanisms, but has effectively filled a niche in the market.

The Agent Network

To facilitate purchases and sales of e-money, and in light of low rates of bank account coverage among a widely dispersed population, M-PESA maintains and operates an extensive network of more than 16,000 agents across Kenya. These agents are like small bank branches, often manned by a single person. As can be seen in Figure 5, the growth of this network lagged behind that of

---

[13] These data refer to a period before M-PESA could be used at ATMs.

Figure 5 Expansion of the Agent Network
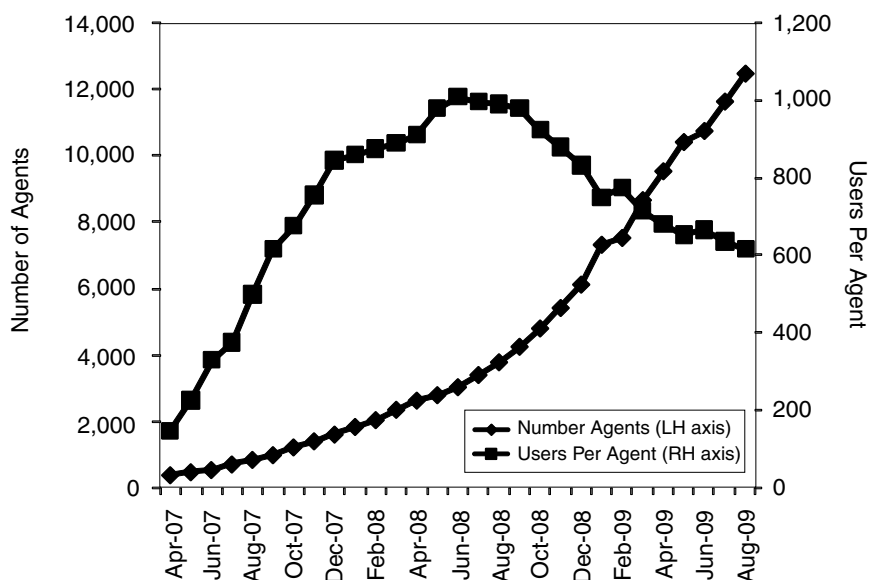
Source: Safaricom.

the customer base for the first year of M-PESA's operation, during which time the number of users per agent increased five-fold from a low of 200 to a high of 1,000. But since mid-2008, agent growth has accelerated and the number of users per agent has fallen back to about 600.

Registered M-PESA users can make deposits and withdrawals of cash (i.e., make purchases and sales of e-money) with the agents, who receive a commission on a sliding scale for both deposits and withdrawals.[14] Clearly, withdrawals of cash can only be effected if the agent has sufficient funds. But symmetrically, cash deposits can only be made if the agent has sufficient e-money balances on his/her phone. Agents face a nontrivial inventory management problem, having to predict the time profile of net e-money needs. Figure 6 shows a representation of the flow of money and e-money among individuals, Safaricom, commercial banks, and the central bank, and illustrates the core workings of M-PESA. The role of what we call the "coordinator,"

---

[14] The commission amounts are nonlinear (and concave) to the size of the transaction. Some reports suggest that in response to this, agents may encourage customers to split their transactions into multiple pieces, thereby increasing the overall commission.

Figure 6 Flows of Fiat Money and E-Money
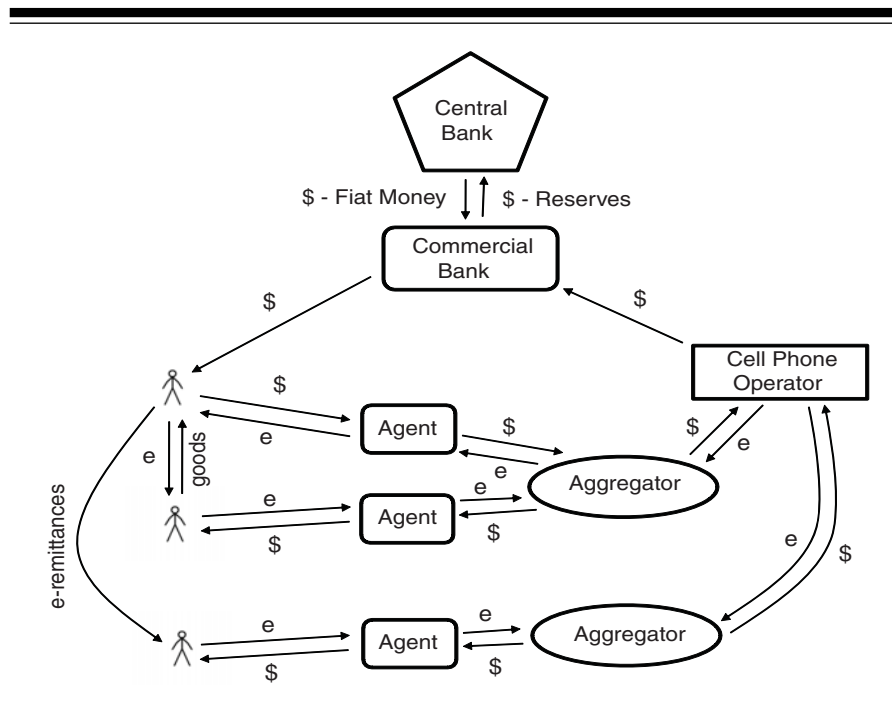
which in practice is a head office, an aggregator, or a super agent, is described in more detail below.

The network of commercial bank branches across Kenya, while growing, remains much smaller. As of November 2009, the Central Bank of Kenya[15] reports that 44 commercial banks had 849 branches across Kenya (about one branch for every 40,000 Kenyans), with 50 percent of branches concentrated in the largest (by size of branch network) four banks. As of 2008 in the United States, there were 7,086 institutions with 82,547 branches that came under Federal Deposit Insurance Corporation protection, yielding a density of bank branches about 10 times that in Kenya (whose population is about 10 percent of the United States).

In practice, M-PESA agents are organized into groups. Originally, M-PESA required that agent groups operate in at least three different locations, so that the probability of cash or e-money shortfalls could be minimized. This diversification within the group would only be effective, of course, if the inventories of money and e-money were efficiently re-allocated across agents

---

[15] http://www.centralbank.go.ke/financialsystem/banks/Register.aspx

Figure 7 M-PESA Agent Models
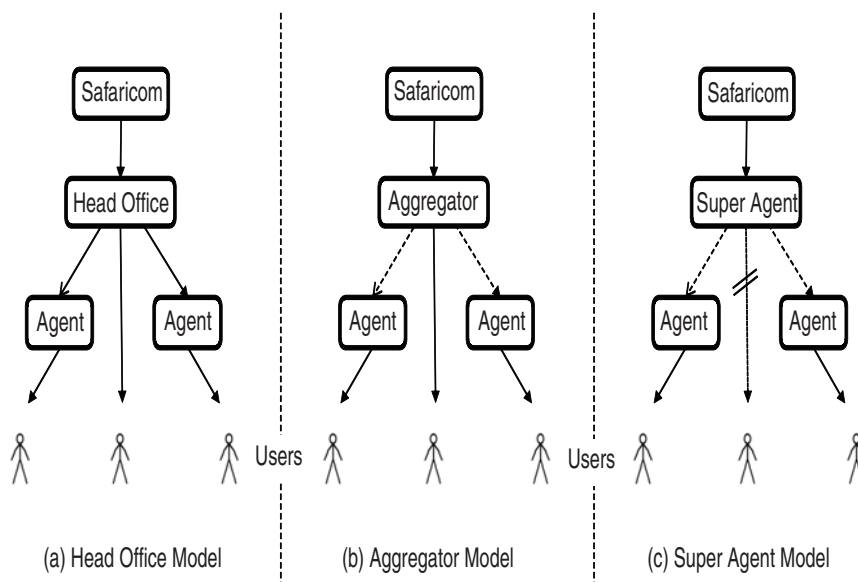
Notes: (a) The coordinating body is the "head office," which owns agents and can transact directly with customers. (b) The coordinating body is referred to as an "aggregator" and has arm's length contractual relationships with agents. (c) The coordinating body is a bank branch and is called a "super agent," but neither owns the agents nor transacts directly with customers.

in the group accordingly. There are now three agent models in operation, in which there is a central body that manages and coordinates the operations of a group of subsidiary agents. These models are differentiated with regard to the formal status of the coordinating body and the ownership structure of the group, and whether the central body conducts direct transactions with individual users, as shown in Figure 7.

In the first model, one member of the agent group is designated as the "head office," which deals directly with Safaricom, while subsidiary agents that are owned by the head office manage cash and e-money balances through transactions with the head office.[16] Both the head office and the agents can transact directly with M-PESA users. The second model is the aggregator

---

[16] M-PESA requires that each coordinating body has a bank account so that funds can be transferred easily between them. In order to open an M-PESA business, the coordinating body must have a minimum balance in a bank account, which is used to purchase initial holdings of e-money.

model, with the aggregator acting as a head office, dealing directly with Safaricom, and managing the cash and e-money balances of agents. However, the agents can be independently owned entities with which the aggregator has a contractual relationship. A final and much more recent model[17] allows a bank branch, referred to as a "super agent," to make cash and e-money transactions with agents on an ad hoc basis. However, the bank does not trade e-money with M-PESA customers. The super agent model is one example of the integration of M-PESA services into the banking system. Other developments in this vein include the ability to transfer funds, often via ATMs, between a user's M-PESA account and accounts at certain commercial banks with which M-PESA has forged partnerships. But even as M-PESA has facilitated transactions for the approximately 72 percent of user households in Jack and Suri's sample with bank accounts, it remains popular with the unbanked, of whom more than half (54 percent) used M-PESA.[18]

The cash collected by M-PESA agents in exchange for sales of e-money is either kept on the premises or deposited in the agent's (or head office's) bank account. When they wish to replenish their e-money balances, agents transfer money via the banking system to one of two bank accounts held by Safaricom. Safaricom is required to limit the quantity of e-money it issues to the amount of money it receives from agents—that is, e-money is 100 percent backed by deposits in commercial banks. However, these deposits are subject only to the regular 6 percent Kenyan Central Bank reserve requirement.

2. MODELS OF MONEY AND MEANS OF PAYMENT WITH SPATIAL SEPARATION

M-PESA's rapid expansion means that a large share of the Kenyan population now conducts at least some of their financial transactions by phone. In this section we discuss the implications of this new kind of payment system for the management of the financial system as a whole and of central bank regulatory and monetary policies in particular. To address these questions, we describe in some detail a number of models of money, the payment system, and clearing and settlement. The purpose is to focus on the features of the models that can provide insights into the operational design of mobile banking and inform policy choices facing regulators and monetary authorities. Therefore, we follow the summary of each model with a discussion of its implications for mobile banking. We proceed incrementally, beginning with simple but surprisingly

---

[17] This model started after the first round of the Jack and Suri (2009) survey.

[18] About 50 percent of households had at least one member with a bank account. Of banked households in the survey, about 60 percent used M-PESA, compared with the 54 percent of unbanked households reported above.

rich models of money, then progressively review more complex models that we believe reflect particular features of the Kenyan financial environment.

## Townsend Model of Financial Deepening and Growth

This model focuses directly on improvements in the technology of communication and links the degree of financial interconnectedness of agents with the level of economic development in a cross section and also over time. The idea is that as connectedness increases, with electronic payments connecting otherwise spatially separated agents, there is an increase in the specialization of labor, an increase in the consumption of market-produced goods, and a shift toward e-money relative to fiat money. This is the story of how financial deepening and growth are intertwined and how M-PESA could help Kenya increase gross domestic product over time at the same time as it increases monetized exchange.

Each household of type i can produce (by supplying labor) only good i, and each has a utility function over its own consumption of good i and a good it cannot produce, $i + 1$, as well as leisure. When households are in autarky, without physical or electronic contact, no trade is possible, so each household consumes all its production of good i only. In this situation, there is no need for a means of payment. In contrast, with some travel, as in the Cass and Yaari (1966) or Lucas (1980) versions of the Wicksell (1935) triangle applied many times, household i can only trade either with household $i + 1$ (whose good i values) or with household $i - 1$ (who values good i). But because of the structure of preferences (e.g., because household $i + 1$ does not value good i and instead wants goods $i + 1$ and $i + 2$), narrow bilateral *exchange* between i and $i + 1$ is in no one's self interest. This is the key lack of double coincidence of wants. Decentralized trade would give rise to autarky if it were not for valued fiat money.[19]

The timing-location is shown in Figure 8 where, in any given period, household i has two members, a shopper and a seller, who can only move horizontally to trade with households $i + 1$ and $i - 1$, respectively. Between time periods, members of household i, i even, shift down one line, and households i, i odd, stay put. Thus, debt issued by a household of type i, i even, can only be passed along to a household vertically above the issuer and so has no value. Only fiat money is used and it can have value. Specifically, one member of each household i travels to the market with $i + 1$ and purchases some of good $i + 1$ at price p_{i+1} with fiat money acquired previously; a second member travels to the market with $i - 1$ and sells some good i for money at price p_i. Note that it takes one period for goods produced and sold to come back via

---

[19] These models rule out private debt and future contracts in fiat money by assuming there are no pairings such that debt can be redeemed by the issuer. See below.

Figure 8 Trading Scheme for Paired Households
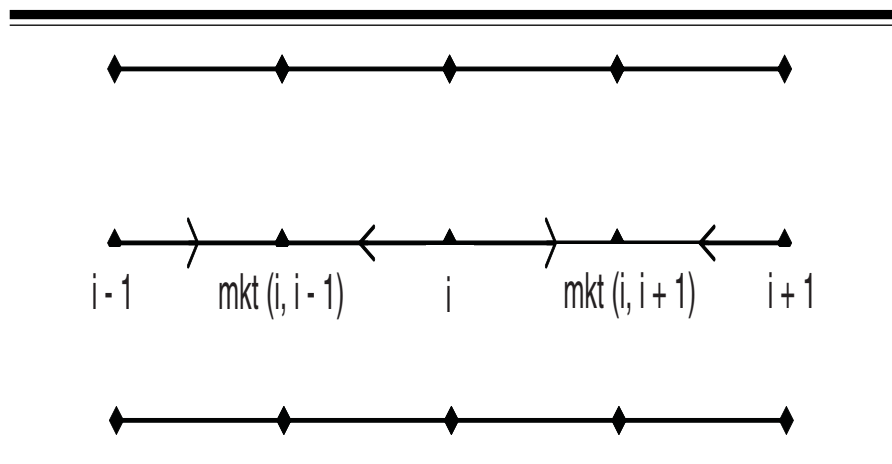
money holding in the interim as goods purchased. With constant prices across time and space and with a positive intertemporal discount rate, this makes it less beneficial to supply labor. This is a crucial aspect of this and other related models below.

In a Walrasian, centralized exchange regime with electronic debits and credits, households can now hold intraperiod debt for within-period purchases and, at the same time, send and receive electronic credits. At the end of the period, accounts are cleared. Intuitively, when one member of household i travels to market $(i, i + 1)$ to buy good $i + 1$ from household $i + 1$, it is as if that member were using a credit card (or phone) linked electronically to a central account, which will not be paid until the end of the period. The second member of household i who travels to market $(i, i - 1)$ and sells good i is paid with a credit card from household $i - 1$. At the end of the period, these electronic debits and credits are cleared and accounts must balance (we return to interperiod debt in the Lacker model below). Note that goods produced and sold can be transformed in this way to goods purchased *within the same period*, so there is no inefficiency associated with holding idle money balances. In fact, in the equilibrium of this electronic accounting system, fiat money plays no role and its price is zero. The prices of goods themselves are in some (arbitrary) unit of account. Related, though households remain separated in space, it is as if they are transacting with one another in a centralized market that ignores the spatial segmentation as far as prices and values are concerned. However, this system works only if households are allowed to overdraft their electronic accounts and there is enough commitment or punishment to make sure they honor debts accrued within the period.

In summary, if we then assume that substitution effects dominate income effects and focus on prices, the cost of consumption of the nonproduced good in terms of labor is infinite in autarky and high in the fiat money regime relative to the centralized Walrasian electronic clearing e-regime. Moving from autarky to the decentralized money regime and then to the centralized Walrasian regime, the model predicts that labor supply increases, output of the produced commodity rises, consumption of the nonproduced good rises, consumption of the produced good drops, trade volume increases, and welfare increases. If an economy has a mix of decentralized and centralized regimes, as with some fraction of "lines" (see Figure 8) using fiat money and other "lines" using Walrasian credit, and these fractions vary across countries, then per capita national income rises as financial interconnectedness increases, fiat money decreases, and per capita private debt increases, but the ratio of fiat money to income decreases and the ratio of credit to income increases. This pattern tends to be what we see in cross-sectional data. Similar comparisons are valid for an economy that is becoming more financially integrated over time, like Kenya, where forward-looking households in the fiat currency part of the economy treat financial integration into the Walrasian e-system as an exogenous random event that happens with positive probability (essentially changing the discount rate). Note, however, that thus far, in this particular model, no household needs to use multiple means of payment.

Implications for Mobile Banking

What are the implications of this kind of model of financial deepening for a system like M-PESA? It is clear that M-PESA will change the financial connectedness of the individuals in the economy, which in the model above will cause higher economic development. Therefore, the main takeaway from this model is that M-PESA can be viewed as a technological innovation that lowers trading costs or, better put, allows financial transfers (credits and debits) across agents who are still separated in space. This improves welfare, at least in the model economy without government and no vested interests in the current intermediation system (and without other heterogeneity). Fiat money and electronic payments can co-exist if some households have access to M-PESA and some do not. However, in the model, but perhaps not in the M-PESA system, the household buying goods in effect creates a net increase in e-money within the period. If e-money were essentially only a debit card, then an initial deposit of currency would have to underlie the debit transaction, undercutting this key advantage. In other words, the theory argues that we might see features of net credit creation in the functioning of the actual M-PESA system, though perhaps at an aggregated or agent level and not necessarily at the level of individual households. However, for this feature to exist there must be a (harsh) means of preventing reneging or default so that accounts actually clear at the end. Even that requires foresight of the overall equilibrium, e.g., here

the shopper knows the prices at which the seller is receiving credits. Again, we come back to this mismatch and interperiod carryovers in the other models below.

Manuelli and Sargent Turnpike Model with Currency and Debt

A closely related model of Manuelli and Sargent (2009) rationalizes the co-existence of fiat money and private credit. As in Townsend's turnpike models, agents meet in pairs and, while they have long enough relationships to undertake some efficiency-enhancing intertemporal trades via the extension of private credit, they do not stay together long enough to effect fully Pareto-efficient allocations. More specifically, time is divided into periods (think of these as "years"), each composed of four subintervals (e.g., "seasons"). Individuals meet for just half a year only, i.e., two consecutive subintervals, and then move on—some to the east, some to the west (see Figure 9). In the first subinterval of a half-year, one person in a given pair has a positive endowment of the single perishable consumption good and the other has none, and in the second subinterval these roles are reversed, giving rise to short-term (two-subinterval) private credit arrangements. However, the positive endowments in each subinterval can be either high or low (for example, $a > 0$, $b > 0$, and $a/b > 1$), while aggregate output in each half-year $(a + b)$ is constant, and each individual's annual aggregate endowment is constant, also equal to $(a + b)$. Because agents remain together for only two subintervals (one half-year), they cannot implement trades *across* half-years—that is, they cannot issue long-term debt. Fiat money plays a role in facilitating the trades that such debt would effect. Manuelli and Sargent generalize this to include labor supply, so that output is endogenous.

One interpretation of Manuelli and Sargent's model is as a generalization of Townsend's original turnpike in which endowments fluctuated with a periodicity of two and meetings lasted only one period. Instead of meeting for two periods, we can interpret Manuelli and Sargent as a model where households continue their travels after one period but remain linked electronically for two periods (though we ignore the requisite costly shipping of goods in the second period—some of the models below are more complicated so as to eliminate this flaw in our attempted interpretation). As the time and spatial limitations of communication fall (e.g., with the expansion of the network of M-PESA agents, accounts, and the use of cell phones), debts of increasingly long maturity can, in principle, be issued and repaid.

Implications for Mobile Banking and Monetary Policy

To the extent that mobile banking facilitates the operation of the private (often informal) credit market, a model that accommodates such products with

Figure 9 Turnpike Setup for Manuelli and Sargent (2009)

East-Headed Agents

(a,0)	(0,b)	(a,0)	(0,b)
(0,b)	(a,0)	(0,b)	(a,0)
1	2	3	4

West-Headed Agents
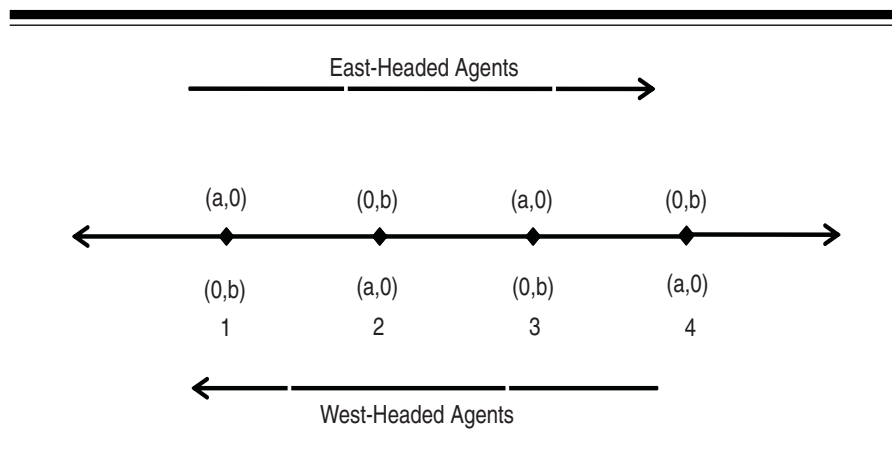
nontrivial implications for policy can be informative. To start, as in the Townsend models, the laissez faire, non-interventionist monetary equilibrium (without debt) is not Pareto optimal. Essentially, the wedge that we discussed in the earlier Cass-Yaari model, where money is earned through production and held without interest for one period, can be eliminated with intervention by paying interest on cash balances. This equates intertemporal substitution in consumption to the natural rate of time discount and ensures that no household hits a binding corner, running out of cash.

But the impact of monetary policy interventions in the form of changes to base money depends on whether private credit is allowed or, under the interpretation here, whether e-money that allows borrowing and lending is in the system. An increase in the growth rate of the money supply has ambiguous effects on the average level of output but increases the volatility of output when there are no restrictions on private borrowing and lending. However, in economies where individuals do not have access to private loan markets, say because they move on without cell phones, the results are quite different: An increase in the rate of money growth decreases mean output and has no effect on volatility of output (which remains zero). Likewise, if the economy is liberalized, or otherwise experiences a surprise innovation that allows private borrowing and lending, then prices increase and output becomes more volatile. Financial innovation is welfare-improving but intimately connected with the impact variables that central banks typically monitor or attempt to control.

As Manuelli and Sargent (2009) emphasize, the potential destabilizing effects of actual financial liberalizations are highlighted in both the academic and policy literatures. More generally, the effects of monetary policy depend on the way private credit markets are operating, even if in the process

of borrowing and lending there is no net creation of e-money. Thus, when formulating monetary policy, the central bank will need to take into account the effective change in financial regimes that M-PESA has brought with it. Indeed, in the above class of models, optimal monetary policy in terms of control over fiat base money is still relatively straightforward but not without interest. Specifically, the allocation achieved under optimal policy differs from the one associated with the corresponding economy with no locational restrictions and centralized trades permitted at time zero. While both allocations are Pareto optimal, they are not the same, implying that efficient monetary policy has redistributive consequences. Further, optimal government-issued currency continues to play an essential role even when interest is optimally paid on holdings of such currency. And, the interest-on-currency policy does not work in a way that can be replicated by free banking in a Walrasian world. Related, moving from a suboptimal policy to one with interest on currency may redistribute income and not be Pareto improving. In this model, unlike the first, e-money does not drive out fiat money nor the need for an optimal monetary policy. This is reminiscent of a class of related models of monetary management in which implementation of policy depends on the ability of agents to trade in asset markets.[20] Financial market segmentation relies on costs that may be arguably decreasing.

**Townsend's Models of Activist Monetary Policy and Money as a Communication Device**

A generalization of Manuelli and Sargent would allow credit arrangements to be used to implement trade among individuals who remain in their home location and deal with each other repeatedly over time. This is also similar to the Walrasian accounting system of the first model above, except that here again the trade is intertemporal, with borrowing and lending over time, so that any individual's balance does not have to net to zero at the end of each period. In the next set of models, credit is identifiable as direct communication and promises. Fiat money then co-exists with credit and serves as a communication device for dealing with strangers across locations.[21] But the models here feature Diamond and Dybvig (1983)-style preference shocks with patient and urgent households generating the desired intertemporal trade. Moreover, the models here deliver welfare gains from an activist monetary authority responding to shocks and managing liquidity needs. More generally, the quantity and

---

[20] For example, see Grossman and Weiss (1983), Rotemberg (1984), Romer (1987), Lucas (1990), Fuerst (1992), and Perez-Verdia (2000).

[21] See also Ireland (1994), Kocherlakota and Wallace (1998), Cavalcanti and Wallace (1999), Kocherlakota (2005), and Wallace (2005), and the review in Wallace (2000) in which outside money and inside money issued by banks with known trading histories co-exist.

Figure 10 On Each Island, a Share of Each Population Leaves to Other Islands in Period 2
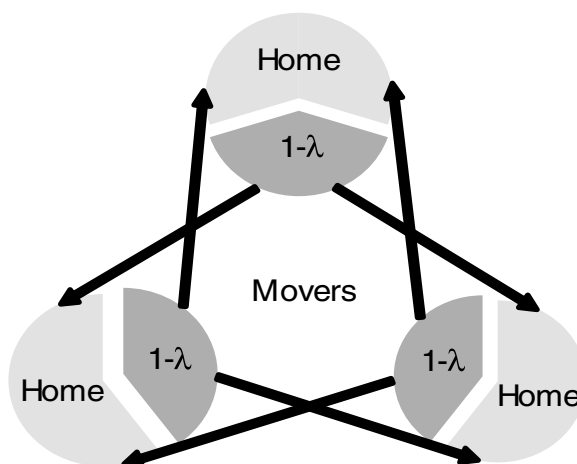
kinds of money in the system are determined optimally in an effort to compensate for missing credit and insurance markets. In this way, one can build on the platform of e-transfers to create a highly effective recordkeeping system in which electronic accounts allow for a rich variety of financial instruments.

Townsend's (1989) model envisions a scenario where there are N islands each with N inhabitants (the case of $N = 3$ is shown in Figure 10 but, more generally, N is a large number).[22] Preference shocks that are correlated among a segment of each island's residents occur in the first period. That is, some fraction of the residents are patient, in principle willing to lend, and the residual fraction are urgent, wanting to borrow. However, a share $(1 - \lambda)$ of the population of each island moves, spreading out across all the other islands in such a way that no mover encounters anyone from his home island at his new destination. This creates a problem if recordkeeping is limited to locations, that is, if there is no cross-island communication or accounting system so that only nonmovers can borrow and lend: Promises involving movers (either among themselves but going to different locations or between them and nonmovers),

---

[22] In this model, the agents all pre-commit to arbitrary tax and transfer schemes over time and to all institutions and resource allocation rules. In the language of the models, they commit to an economy-wide credit arrangement that specifies consumption and transfers to agents conditional on aggregate states and on individual specific location shifters (that are public) and individual announcements of preference shocks (private). Apart from these plans, there is no government and no distinction between private and public.

on the other hand, are not credible as they cannot be consummated at a later date.

As movers are effectively excluded from the credit market, a social planner could attempt to implement efficient intertemporal consumption profiles by asking movers at each date to report their preferences, allocating consumption accordingly. But if the information reported cannot be credibly transmitted to other islands without a recordkeeping device, then the only incentive-compatible mechanism is one that gives all movers the same level of consumption in both periods, independent of their preferences. Portable fiat money allocated to movers, and monotonically related to their first period announcements, can facilitate the transmission of information across time and space to the strangers they meet at their destinations. In this interpretation, fiat money is a portable token. By allowing side trades between individuals, monotonicity can be strengthened to linearity, delivering a price of fiat money or tokens for goods. Of course, the initial nominal price level remains arbitrary, as that is simply a matter of the denomination of the unit of account.

However, if additional periods are added to the model (e.g., another round of movers), future movers must also be allocated fiat money in order to engage in intertemporal trade. The purchasing power of each unit of money allocated to second-round movers must, for efficiency reasons, be the same as that offered to first-round movers, but the quantity will be increasing in the number of movers and the proportion who are patient.[23] As the preferences of new generations of consumers are revealed, planned consumption levels supported by allocations of fiat money in early periods may be revised. Since the purchasing power as previously explained is constant across early movers in a given period, the associated adjustment to consumption levels is effected through changes in the price level. That is, inflation eats up the purchasing power of first-round movers if it is judged that they should be getting less given what new information tells the monetary authority about the way they and second-round movers should be treated. Note that this activist policy is quite different from, say, a Friedman rule, as described in the earlier class of models above in which a constant rate of deflation can remove the distorting wedge. This is not enough here. Optimal policy is state contingent (Manuelli and Sargent [2009] anticipate such results in the concluding section of their article).

Note, however, that fiat money as tokens conveys only the information that a household has been patient in the past, not that the household has been a first- or second-round mover. If even more information, such as the dates and the nature of past transactions, was encoded in the system, then the distinction between local and inter-island accounts could disappear. That is, one

---

[23] Individuals who are patient consume more later and therefore need more units of money to confirm this to future strangers.

Figure 11 Islands with Stayers (a) and Movers (b)
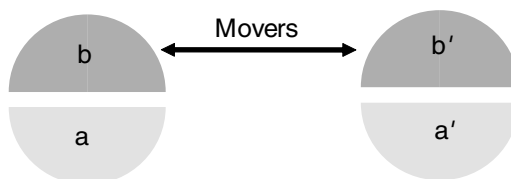
can imagine one kind of fiat money—e.g., red tokens for first-round movers, green tokens for second-round movers, and electronic accounts for those who stay home. Indeed, accounts that distinguish all these space/time transactions could be accomplished with the electronic recordkeeping that mobile technologies and markets allow, at least in principle. Indeed, with all of that, we could in theory go further and here again completely mimic the outcome of a perfect Walrasian accounting system in which changing locations per se has no consequences. The fraction of agents leaving an island would be exactly the same as the fraction arriving and, financially speaking, there would be no strangers.

Townsend (1987) generalizes this idea of multiple monies (or differentiated e-accounts) in a similar framework with four agents, spatial separation, and private information on preference shocks. In particular, suppose there are two islands with two individuals each, as illustrated in Figure 11. In period 1, agents *a* and *b* live on the left island and agents *a'* and *b'* live on the right island. In period 2, *b* and *b'* switch places, while *a* and *a'* (who are subject to shocks) remain on their home islands. Agents *b* and *b'* are risk neutral and in principle can provide insurance to agents *a* and *a'*, who are risk averse. With one good, preference shocks determine not only the degree of risk aversion but also relative patience. With two goods, there can be preference shocks for each good over time (e.g., patient for good one and urgent for good two) and an overall intertemporal shock determining utility in period 1 versus period 2.

Townsend then examines the properties of trade facilitated by alternative communication devices in this environment, both for the cases of a single good as well as for multiple goods. First, oral communication can take place only between agents in the same location and so cannot be used to convey credible information across time to strangers (if agents cannot carry tokens, commodities, or messages). The equilibrium is thus Pareto inefficient. On the other hand, tokens (money) that are appropriately distributed in period 1 can be used to verify information in period 2, helping with incentives to reveal information correctly and acting again as a technology for storing that information. The previous model provides intuition for the case of one good.

However, with two goods, one type of token may not be enough. Intuitively, one wants to convey the full history of shocks for each good in the first period, yet ensure incentive compatibility in the second when agents can turn out to be very desirous of consumption overall. For example, one type of token, say green, is handed out in period 1 given a certain realization of preference shocks, while the red token is handed out given another realization of these shocks again in period 1 (alternatively, these are different "credits" in different cell accounts). Then in the second period, the agent is required to show not just the correct number of tokens, but also the correct colored token (or have the requisite balances in a specific cell account). Indeed, much can be done even with n-commodities and m-combinations of shocks using combinations of red and green tokens (two types of e-money) as an encryption system. The point more generally is that multiple monies are used to convey the history of trade, borrowing and lending, and insurance, not simply a means of payment or transfer system.

Implications for Mobile Banking

The bottom line of these models of money as a communication device is that the better the communication of past shocks or transactions, the more efficient can be the allocation of consumption; however (with initial heterogeneity), this may be wealth redistributing. The model features tokens or fiat money but, again, portable cell devices linked to some of the account history of earlier transactions would provide similar features. To achieve an efficient allocation there can arise, as in these models, the need for active liquidity management. We can see that in a scenario where M-PESA emerges as the entity behind a large fraction of transactions, e-money could substitute for fiat money or tokens. This would not necessarily replace the need for an activist monetary policy, but it would alter that policy so that the level of tokens created on net by the financial system ideally responds to mobility and the state of demand, as would electronic credits if allowed optimally to function that way. Here a distinction between private credit and public money becomes blurred as we consider questions about optimal market design. The social good is served by having mutually agreed upon and collectively enforced rules.

Another lesson from these models is that electronic records of past transactions allow new financial instruments, in this case better borrowing/lending and insurance over space and time. Tying fiat money to e-money and thinking of both as solely facilitating payments may lead one to miss otherwise beneficial arrangements that have to do with insurance against spatial and intertemporal idiosyncratic and aggregate shocks. Indeed, under the current M-PESA system, the prices at which money trades for e-money are supposed to be fixed over time and across space; e-money and cash trade for each other one for one (as described above, however, there are nonlinearities in the transactions costs by amount traded)—yet these fees can be seen as allowing in

principle a trading price between cash and e-money that is different from one. Whether or not one wants to allow money prices and the rate of exchange of money for e-money to move with the state of local demand and inventory of the actors again begs the question of what e-money is supposed to be: a means of payment only, if it facilitates an expansion of the monetary base, or a partial substitute for missing, more centralized economy-wide insurance and credit markets.

Townsend and Wallace—Circulating Private Debt and a Coordination Problem

There is yet another way to think of money, namely as an object that, even if privately issued, appears frequently in exchange, i.e., with a high velocity. We can understand this by simply extending the model environment in the previous section to four periods with households b and b' continuing to switch locations from one period to another, back and forth, and with households a and a' remaining in a single location. Townsend and Wallace (1982) replace preference shocks with time-varying endowments of a single good, but with different profiles for the different agents, to induce the desire for intertemporal trade. They also assume there are many agents of each type in any given location to justify price-taking behavior. In one of the equilibria, the first period household b makes a deposit of goods (but could be money) to (that is, lends to) agent a, as if agent a were a bank issuing long-term debt (or at least debt payable on demand). However, household b does not hold this debt but rather moves in the second period to a different location inhabited by agent a'. At this new location, neither party is physically connected to bank a. Subsequently, in the third period, agent a' will pass the debt to b', who in turn redeems it in the last period since b' meets up with the original issuer, agent a. Note that long-term debt is also the debt that circulates, that has a high velocity. In that sense circulating debt has something to do with maturity transformation. Short-term debt (e.g., two-period debt) between agents a and b (or a' and b') not only extinguishes sooner but it also does not circulate.

With private debt transferable electronically, one has the same equilibria but with the more realistic interpretation of agent a as an M-PESA agent who issues debt (in this case in exchange for goods, not fiat money, but see below). That is, household b uses the e-money account to trade with subsequent households and, again, the e-money is netted out back to zero when a third party comes to agent a to redeem it. In this model, agent a' can also play this role as banker, or M-PESA agent, instead of a. However, without coordination, another problem emerges. The amount of e-money issued by agents a and a' has to be coordinated so as to be consistent with the overall equilibrium. If M-PESA agents a and a' are not communicating across space, then it is hard to imagine how this would happen. Townsend and Wallace refer to various
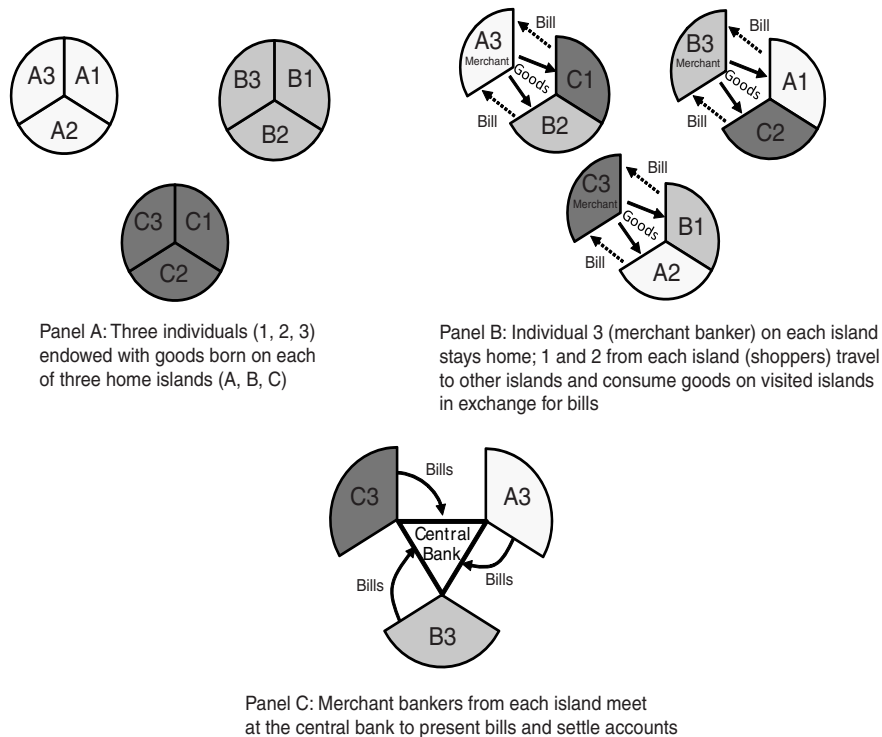
historical episodes such as the crash of markets in bills of exchange as evidence that the model with coordination problems is picking up problems that may occur in practice.

Implications for Mobile Banking

Electronic debits can be transferred across agents in spatially separated locations and have a high velocity. This seems to capture a big part of the Kenyan M-PESA reality. This comes, however, with potential coordination issues that need to be thought through. In the current model with two locations, four agent types, and four periods, one achieves the first-best with the right combination of circulating private debt and other short-term noncirculating debt. In that equilibrium, prices/interest rates are moving around over time and space and all markets in goods and financial instruments are clear. Again, fixing the price at which one object trades against another would seem to create additional problems. But even if prices were flexible, it appears that agents need to coordinate on the overall credit issue. Lack of initial coordination could show up as an over-issue or under-issue of the correct financial instrument, or of the combination of instruments that is supposed to give the correct maturity structure, showing up in turn later on as sharp movements in prices. This could even lead to doubts about the commitment or ability of agents to achieve the requisite transfers of purchasing power necessary for liquidity in intermediate periods or to ensure redemption of debt at maturity. Manuelli and Sargent ponder whether fiat money can help solve this type of coordination problem.

Lacker's Model of Clearing and Settlement and Inter-Agent Markets

Lacker (1997) focuses on clearing and settlement via a central bank and the impact of certain central bank policies such as reserve requirements and interest paid on reserves. Building on the earlier models of Townsend (1983, 1989), Lacker develops a model in which there is a large number, N, of islands, on each of which live N individuals. Each island produces a single perishable good that must be consumed on the island. This geography is illustrated for the case $N = 3$ in Panel A of Figure 12, in which the islands are labeled A, B, and C, and the individuals are 1, 2, and 3. In each period, all but one of the individuals who live on a given island travel to all the other islands at random, one to each, with one staying behind. In Panel B of Figure 12, individual 3 remains home. All individuals consume the good that is produced on the island they visit (so the one who remains consumes the good produced on his island). As in Townsend (1989), before they leave "home," travelers entrust their endowment of goods to the individual who stays behind (called the merchant banker) and is responsible for handing it out to arrivals from

Figure 12 Lacker's Model of Settlement

Panel A: Three individuals (1, 2, 3) endowed with goods born on each of three home islands (A, B, C)

Panel B: Individual 3 (merchant banker) on each island stays home; 1 and 2 from each island (shoppers) travel to other islands and consume goods on visited islands in exchange for bills

Panel C: Merchant bankers from each island meet at the central bank to present bills and settle accounts

other islands. The record of the amount entrusted to the stay-at-home agent is an individual's "deposit" and the merchant banker is thought of as operating a bank that holds his island's deposits.

As illustrated in Figure 12, each island receives a fully diversified group of visitors each period, one from each other island. If preferences and endowments were suitably fixed (e.g., if each individual consumed $1/N$ units of the good of the island s/he visited), consumption and income would balance on a person-by-person basis. However, Lacker assumes, like some of the models above, that each period the islands are hit by Diamond-Dybvig idiosyncratic independent identically distributed preference shocks that affect the urgency of consumption. All individuals from a given island get the same shock. Since each island is visited by an individual from every other island and since shocks are independent across islands, there is no aggregate uncertainty about the demand for each island's good (as N goes to ∞). However, an island that suffers a run of large urgent shocks over time consumes more over time than an island that suffers a run of small shocks.

Because goods do not move between islands, there is no possibility to directly exchange one for another. Instead, an individual purchases consumption from the merchant on the island he visits by providing a bill or check drawn on his deposit held by his own merchant who stayed at home. (This could be an electronic charge to the e-account but, again, not one paid instantaneously.) In turn, each merchant collects bills or e-credits from all other islands, one for each visitor. In any given period, some islands will consume more than they produce (i.e., issue more bills than they collect or be left with negative e-balances), while the opposite will be true for others. Intertemporal trade between islands across periods, i.e., interbank borrowing and lending of e-balances, is thus efficient.

In the final stage of the period, all the merchant bankers travel with their bills to a central location and submit them (to each other) for payment (Panel C of Figure 12). With cell technologies, physical meetings would not be necessary. Payment is effected through an accounting mechanism, with each island's account being credited and debited according to the bills or e-money presented to and by it. The residual that does not clear is carried over, in surplus or deficit.

Implications for Mobile Banking

Lacker refers to the central institution that keeps the accounts of each island as the central bank and these accounts are thought of as *reserve* accounts. However, this could equally be a private clearinghouse run by Safaricom or some other independent entity, as the model focuses on the account-keeping and clearing functions of the institution, not the issuing of money per se. Positive account balances with the institution are the liabilities of that institution, while negative balances represent overdrafts. In the model, bills are *cleared* (i.e., accepted by the clearinghouse/central bank) at the end of the period and *settled* (i.e., deposits transferred by the institution from one island's reserve account to another's) at the end of the period for across-period borrowing and lending.

Beyond Lacker's model, limits on within-period bank overdrafts with the clearinghouse/central bank can induce some banks (ones with a positive preference shock) to borrow from others. Likewise, limits on overnight overdrafts can induce residents of islands that have had a string of positive, urgent shocks to constrain their consumption below the efficient level, as they are unable to borrow enough. Lacker's model is a useful motivation for thinking about another aspect of the M-PESA system, in particular overall clearing and the related inventory management problem faced by agents. The kind of contractual conditions Safaricom might want to specify would be crucial given the reality of the actual economy in which the distinction between within-period and across-period clearing and borrowing/lending is hard to maintain.

We identify each M-PESA agent with a merchant banker in Lacker's model, although individuals are not bound to agents like residents are to islands. An M-PESA agent's trading account at Safaricom corresponds directly to the reserve account held by each bank with the central bank. To parallel the model, individuals deposit their endowments (of cash) with an agent each period, which requires that the agent hold sufficient e-money. An agent would take out an overdraft loan from Safaricom if he were required to issue e-money to a customer before having presented the equivalent amount of cash to Safaricom. Because transferring a bank note or cash is slower than transferring e-money, it seems likely that there could be demand for such overdrafts.

E-money is sent between individuals (i.e., checks are exchanged) and recipients present their e-money (i.e., checks) to agents. This happens at the end of the period in Lacker's model. If agents have enough cash to purchase the e-money from customers, their trading accounts are credited with the relevant amounts. In reality, as individuals visit the agent over the course of the day, his net demand for e-money will fluctuate and he might require short-term overdrafts from Safaricom or need to acquire cash in some other way. In the absence of such a facility, he will need to trade off the costs of holding "zero-interest reserves" (e-money balances on his trading account) against the costs of reduced trade (and commissions). Alternatively, agents could lend e-money to one another, creating the equivalent of an interbank market as envisioned in Lacker's model. Again, this might be organized by another institution (like a clearinghouse) that itself purchased e-money from Safaricom and lent it out to agents at some interest rate. Likewise, the head offices or super agents could perform this role, though neither appears to charge interest. Each head office or super agent would face a similar inventory management problem of course, having to hold enough e-money and/or cash to lend out during the day/period.

Freeman and Green's Models of Liquidity—Optimal Base-Money Management

Freeman's (1996) model and Green's (1999) reformulation, related to Townsend (1989) as exposited above, focus on getting money and circulating debt in the same setting simultaneously because of imperfect meetings between creditors and debtors. This, again, has implications for liquidity and monetary policy.

In Green's overlapping generations model, there are two types of individuals (creditors and debtors) who live for two periods each. A creditor is someone who in equilibrium will be willing to defer consumption, while a debtor will wish to borrow. We follow tradition and refer to young and old agents, simply to imply the first and second periods of the two-period, dynamic transactions profiles of the households. When young, creditors and debtors are endowed with perishable goods x and y, respectively. In the first

Figure 13 Trade with Money and Credit
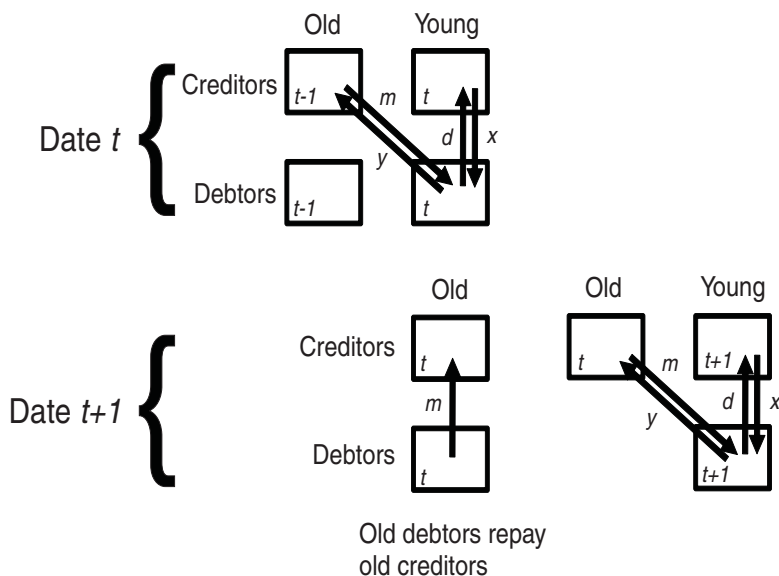
Notes: The dates within each box refer to the period of birth.

period, old creditors are endowed with fiat money and old debtors have nothing. Creditors and debtors also differ in their preferences: Creditors wish to consume when they are young and old, while debtors wish to consume only while young. Both types prefer to consume a mix of goods x and y instead of just their own.

Suppose young debtors meet young creditors first and only then go on to meet old creditors. Young debtors purchase x in return for debt d that they issue to young creditors, as illustrated in the first panel of Figure 13. Subsequently, young debtors sell their own good y to old creditors in exchange for money. At the beginning of the next period (the second panel of Figure 13), now-old (previously young) debtors settle their debts using money with now-old (previously young) creditors. Once the debt is settled the process repeats, with the now-old creditors holding money and the new young cohorts endowed with goods.

Nontrivial monetary dynamics can arise when creditors and debtors do not necessarily meet at the "right" time. With various waves of movers, old agents either arrive late or leave early: In particular, a fraction $(1 - \delta)$ of debtors arrive late and a fraction γ of creditors leave early. This naturally complicates the debt settlement process and can lead to inefficiencies. In

**Figure 14 Late-Arriving Debtors and Early-Leaving Creditors**
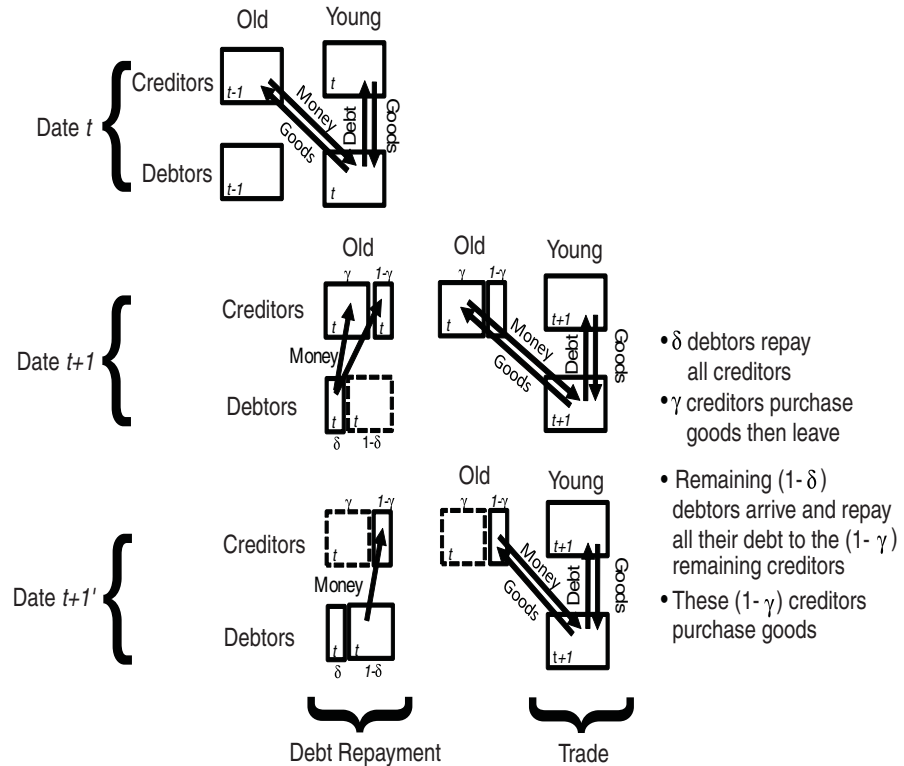
Notes: In the figure, debt settlement and trade between the young and old occur in the same period, $t+1$.

particular, while efficiency requires that all creditors consume the same quantity of goods when old (purchased from young debtors), those who leave early may in equilibrium consume *less* than this amount while those who leave later consume more. Thus, early-leaving creditors can end up facing liquidity shortages that constrain trade.

This situation is illustrated in Figure 14, in which period $t+1$ is divided into two subperiods ($t+1$ in the second panel down, and $t+1'$ in the third panel). Agents not in the market at the relevant moment in time are shown as boxes with broken lines. It is assumed that the creditors are fully diversified at $t+1$, holding debt issued by each and every old debtor. At the beginning of $t+1$, all the old creditors are present, as are a fraction, δ, of old debtors. The debtors settle their debts and each creditor receives a share, $\delta < 1$, of

Figure 15 Trade in Money and Credit with a Secondary Debt Market
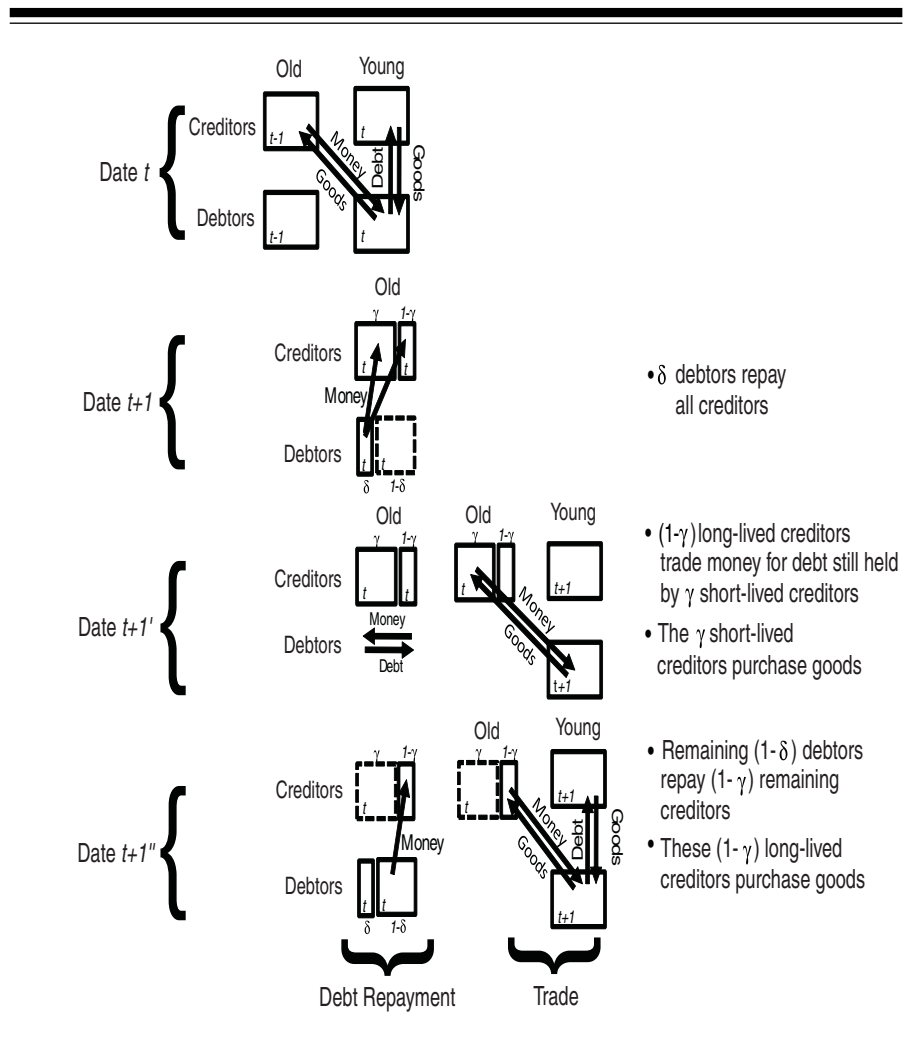
the amount owed to him. At date $t + 1'$, the late-arriving old debtors are able to settle their debts in full with old creditors who remain. Early-leaving old creditors will consume less than their efficient level of consumption and late-leaving ones will consume more.

An alternative scenario illustrated in Figure 15 allows old creditors to exchange debt. Period $t + 1$ is now divided into $t + 1$, $t + 1'$, and $t + 1''$. At the beginning of $t + 1$ (the second panel), early-arriving debtors repay their debts to all creditors. As before, all creditors continue to hold outstanding debt issued by debtors who have not yet arrived. Next, at time $t + 1'$ (the third panel), early-leaving creditors sell their remaining holdings of debt to

long-lived creditors in exchange for money. If $\delta < \gamma$, with a relative scarcity of debtors and an abundance of early-leaving creditors, the creditors spend all of their money on debt, which has a price less than one. Early-leaving creditors then purchase goods from young debtors and quit the market. In the final panel, at time $t + 1''$, late-arriving debtors settle their debt with those creditors who remain, including the debt those creditors bought from early-leaving creditors. They then purchase goods from young debtors. The amount of consumption they enjoy is $(1 - \delta)/(1 - \gamma)$ times the efficient level. Thus, if not enough debtors show up in time, even if creditors trade their debt, the allocation is inefficient.

As Freeman observed, a central bank can remedy the inefficiency by issuing money to some or all creditors and then withdrawing it from circulation later, say by taxing young creditors as they enter the second period. The important issue is that new money is issued to creditors and is issued *before* the early-leaving creditors depart.

Implications for Mobile Banking

In Green's version of the model, debtors consume nothing in the second period of their lives, but creditors do. The only reason for old debtors to come to the market is to pay off their debts. So if e-money allows them, or some of them, to do this without coming to the market, then the share that are "late" is smaller. In the extreme case, there would be no late-arriving debtors and no liquidity problems for creditors. But if some old debtors still didn't pay off their debts in time (maybe because they couldn't find an agent with e-money), then it would be possible that early-leaving creditors wouldn't have enough money to finance the efficient level of purchases from young debtors.

These models can inform our thinking about mobile banking in a couple of ways. First, focusing on the reduction in transactions costs associated with transferring e-money, mobile banking might reduce the proportion $(1 - \delta)$ of debtors who "arrive late" and the proportion of creditors, γ, who "leave early." Debtors who previously had to physically meet their creditors in order to settle their debts can now settle them with e-money and no longer need to be present. On the other hand, even if not all debts are repaid at the beginning of the period (i.e., if there remain some late-arriving debtors), the existence of e-money could relax the liquidity constraint faced by early-leaving creditors and make central bank intervention less necessary.

However, it is overly simplistic to assume that mobile banking allows individuals to send *money* costlessly: it allows them to send *e*-money costlessly (or at least at low cost) but they must acquire it first. A more complete model would thus include individuals holding optimal mixes of money and e-money and would describe the production process whereby each is converted into the other. In practice, this conversion is effected by M-PESA agents who simply

Table 3 The Problems Consumers Have Had with Agents

	Most Used Agent	Closest Agent
Agent Gave Less Money/ E-Money Than I Was Owed	2.63%	2.80%
Agent Charged Me to Deposit	1.11%	1.68%
Agent Overcharged Me	1.17%	1.84%
Agent Undercharged Me	0.52%	0.78%
Agent Was Absent	0.74%	0.91%
Agent Refused to Perform the Transaction	0.80%	0.61%
Agent Was Unknowledgeable	1.04%	1.57%
Agent Was Rude	3.66%	6.02%
Agent Had No E-Money/Not Enough E-Money	43.60%	22.81%
Agent Had No Cash/Not Enough Cash	34.78%	51.33%
Other	9.95%	9.65%

perform the role of a technological black box—a black box that is sometimes out of service.

Although this feature is not part of the Freeman and Green set-up, if money and e-money are both used in equilibrium, then a "late-arriving debtor" might correspond to an individual who is otherwise "on time" and has sufficient financial resources (money and/or e-money) to repay his debts, but who is frustrated in not being able to find an M-PESA agent with sufficient e-money or money.[24] Similarly, an early-leaving creditor in this environment could be one who has in fact been repaid, say in e-money, but who must use money to purchase the consumption good.[25] If he cannot find an M-PESA agent with sufficient cash, then he could be liquidity constrained as above. First, if he is stuck with e-money but must trade with money, he will suffer a loss equal to his excess e-money holdings. On the other hand, even if he cannot find an M-PESA agent to trade with, he might trade his e-money with another creditor for cash, just as early-leaving creditors sell their debt to late-leaving creditors for cash in the third panel of Figure 15. But such trades must take place between locationally proximate agents, and if there is an excess supply of e-money locally, the allocation of consumption might remain inefficient.

---

[24] Whether the debtor needs to find an M-PESA agent with money or e-money will depend on what form of financial wealth the debtor has on hand and how the creditor wishes to be paid. This in turn will depend on the specific features of the two monies.

[25] Again, whether the creditor needs money or e-money depends on how the seller wants to be paid.

Table 4 Unable to Deposit Cash (No E-Money) or Unable to Withdraw Cash (No Cash)

	Most Used Agent	Closest Agent
Have You Ever Been Unable to Deposit Money at this Agent?		
Yes	6.63%	6.22%
No	93.37%	93.78%
Total	100%	100%
Have You Ever Been Unable to Withdraw Money from this Agent?		
Yes	6.63%	15.33%
No	93.37%	84.67%
Total	100%	100%

3.   LINKING THEORY WITH DATA: RESULTS FROM HOUSEHOLD AND AGENT SURVEYS

In this section we present data that speaks to some of the issues raised by the models of money summarized above, especially as regards shortages of e-money and cash and whether there are indeed credits in the system because of the operational logistics of agents as described in Section 1. These data, some of which are reported in Jack and Suri (2009), derive from a survey of 3,000 households and 250 M-PESA agents in Kenya in late 2008.[26] We focus on issues related to M-PESA agents as reported by consumers and the agents themselves, as motivated by the models.

First, 10 percent of all consumers reported facing at least one problem with the agents they had visited. Of those who reported problems, Table 3 shows the breakdown of the problems they had. By far, the most common problems are agents' lack of cash and e-money. The first four rows in the table, in fact, suggest that in some cases agents have been able to increase the price of e-money by varying the fees they charge consumers. This is an important implication of the models discussed above—fixing prices for cash and e-money will require an accompanying policy stance. However, the penultimate two rows in the table confirm that this strategy is employed nowhere near enough to clear the market.

In addition, in the survey, consumers were specifically asked if they were either unable to deposit money or unable to withdraw money from the

---

[26] Part of these data form the basis of a confidential report issued by Financial Sector Deepening to the Central Bank of Kenya. In addition, Jack and Suri (2010) look at some of the microeconomic risk-sharing impacts of M-PESA. Other papers have also looked at the more microlevel impacts of e-money on currency demand (for example, see Fujiki and Tanaka [2009]).

Table 5 How Often Do Agents Run Out of E-Money?

	Fraction
More Than Once a Day	3.2%
Once a Day	6.4%
Once a Week	14.0%
Once a Month	5.6%
Once Every Three Months	1.2%
Once Every Six Months	0.4%
Less Often Than That	12.0%
Never	57.2%

M-PESA agent closest to them or from the agent they used the most. Table 4[27] shows that approximately 6 percent of M-PESA users are unable to deposit money with an agent, i.e., the agent does not have any e-money to give the consumer in return. Also, as many as 15 percent of consumers were unable to withdraw money from the closest agent, i.e., the agent had no cash to give the consumer in exchange for e-money.

In the survey of agents themselves, respondents were asked how often they run out of e-money and how often they run out of cash—these results are reported in Tables 5 and 6. On average, about 29 percent of agents run out of e-money once a month or more frequently and indeed a nontrivial fraction (14 percent) run out about once a week. Similarly, about 26 percent of agents run out of cash once a month or more frequently than that and about 10 percent run out once a week (and, in fact, about 8 percent run out on a daily basis). Clearly there are liquidity issues, both in terms of cash as well as in terms of e-money. This is anticipated from the discussion of Lacker (1997), Freeman (1996), and Green (1996)—models in which such liquidity constraints are evident.

Safaricom initially required all M-PESA agents to pre-purchase e-money before they could trade it for money to the public. If an agent runs out of e-money, he is required to purchase more, either from Safaricom or from the public when they redeem cash, before being able to take cash deposits from the public. This suggests that there are no credits or debits involved between agents and Safaricom and therefore no role for a formal settlement system. Indeed, even as the agent model has evolved, this feature has been maintained, at least with respect to the relationship between the "coordinating bodies" of Figure 7 and Safaricom. On the other hand, cash and e-money transactions between agents and their head offices or aggregators need not remain in continuous balance, and the parties can operate in a net credit or

---

[27] Note that the questions asked for Tables 3 and 4 are quite different. Table 3 asks about the main consumer-reported problems with agents while Table 4 asks about the incidence of two specific problems.

Table 6 How Often Do Agents Run Out of Cash?

	Fraction
More Than Once a Day	3.2%
Once a Day	8.4%
Once a Week	10.0%
Once a Month	4.8%
Once Every Three Months	1.2%
Once Every Six Months	0.4%
Less Often Than That	22.4%
Never	49.6%

debit position vis-à-vis each other. These imbalances are of little concern for the head office model (Panel A of Figure 7), to the extent that the agents are owned and controlled closely enough that internal financial arrangement of the group does not affect its viability. However, the more arm's length relationship between agents and an aggregator (Panel B) suggests that chronic imbalances with such a group could prove problematic.

We note that while the potential exists for persistent financial imbalances within a group under the aggregator model, in principle M-PESA users on the one hand and Safaricom on the other do not face any risk associated with the bankruptcy of any particular agent or agent group, as deposits of cash are matched at the level of the coordinating body by transfers of e-money and vice versa. If an individual user finds that all agents within a reasonable distance go out of business, he will likely face a liquidity constraint unless he is able to use his e-money to directly purchase goods and services.

The survey asked agents how they pay for e-money when they request it from their head office. In well over half the cases, agents receive transfers from their head offices without any immediate corresponding payments (see Table 7).

Similarly, agents were asked how they get cash for M-PESA transactions when they run out. As reported in Table 8, in more than half the cases agents do not immediately exchange e-money for cash received.

The statistics in Tables 7 and 8 suggest that credit arrangements, explicit or otherwise, between agents and their head offices or aggregators appear to be widespread. We do not know the maturity of these credits but, given the large number of agents reporting them, it is possible that at least some of them are longer than simple overnight positions. Indeed it seems inevitable that there may be a nontrivial amount of credit in these transactions as the "supply chain" of e-money involves exchanges in spatially and temporally separated markets, leading naturally to one-way transfers of either cash or e-money. Given the mechanics of M-PESA, these credits cannot be issued between agents and individual users or between the coordinating body and Safaricom. But, the

Table 7 How is E-Money Paid For When an Agent Requests It?

	Fraction
It is a Direct Purchase of E-Money From the Head Office (Involves a Cash Transfer)	36.2%
Receive a Direct Transfer From the Head Office with No Concurrent Payment	31.2%
Receive a Direct Transfer From the Head Office with No Payment Required	18.1%
Other	12.6%
Refused to Answer	2.0%
Total	100%

evidence indicates that such net credits/debits do exist between agents and their coordinating bodies. Again, all the models above allow for credit and debits, which are often welfare improving. In addition, some of the models above illustrate the welfare-improving nature of broader financial integration, which such credits and debits would encourage.

4.   CONCLUSION AND FURTHER MODELING

The most successful version of mobile banking in Kenya (and perhaps the world), M-PESA, is—quite literally—everywhere. In many cases, the scenarios envisioned in existing monetary theory models appear to match the reality of M-PESA and, as such, these models promise to inform decisions taken by both Safaricom in managing M-PESA and the central bank in managing the Kenyan economy. The empirical evidence presented, from surveys of both M-PESA users and agents, further serves to illustrate the importance of these lessons.

For example, just as the central bank may intervene to relax liquidity constraints, it is arguable that Safaricom should actively manage "e-liquidity" by issuing e-money that is at times, in some locations, unbacked by money deposits, assuming that such activism would be costless and allowed by the central bank. In fact, the data suggest that some M-PESA agents are engaging in such e-liquidity management already (for example, when they receive e-money transfers from their head offices without a corresponding transfer back of cash). This has implications, of course, for the measurement and meaning of monetary and debt aggregates. Improved systems, however, require that the company have better information on net demands for e-money across agents than the agents themselves have, or at least be better able to act on this information without the space/time coordination problems that the models suggest. Not surprisingly, Safaricom has been changing their agent model over time to better deal with cash and e-money liquidity issues. Time series

Table 8 How Do Agents Get Cash When They Run Out of It?

	Fraction
Redeem From Head Office in Exchange for E-Money	17.6%
Redeem Direct Transfer of Cash From Head Office with No E-Money Exchange	20.4%
Use Own Savings	42.8%
Other*	18.0%
Don't Know	1.2%
Total	100%

Notes: *Of which 27 percent is from "sale of credit card," 27 percent is "wait for deposits," 21 percent is "borrow from management," and 11 percent is "from the other business."

and geographically disaggregated data on fluctuations in demand would be useful for further evaluating these issues and making improvements to their system.

On the modeling side, understanding the operations of the M-PESA agent network seems key to the development of an overall comprehensive model of e-money. For example, modeling the decisions and constraints of agents would potentially allow us to endogenize the timing patterns assumed in some of the existing models. Frictions that impede efficient and immediate reallocations of money and e-money balances across agents would thereby replace these timing assumptions as the fundamental source of liquidity constraints. Similarly, realistic heterogeneity across consumers—for example, in terms of phone ownership, access to mobile coverage, safety of the local environment, frequency of market access, access to M-PESA agents—could be modeled more explicitly.

REFERENCES

Cavalcanti, Ricardo, and Neil Wallace. 1999. "A Model of Private Bank-Note Issue." *Review of Economic Dynamics* 2 (January): 104–36.

Cass, David, and Menahem E. Yaari. 1966. "A Re-Examination of the Pure Consumption Loans Model." *Journal of Political Economy* 74 (August): 353–67.

Central Bank of Kenya. 2009. Presentation at Conference on Banking and Payment Technologies, East Africa, Nairobi (February).

Diamond, Douglas, and Philip Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (June): 401–19.

Financial Sector Deepening. FinAccess Reports, various. Kenya: FSD.

Foster, Kevin, Erik Meijer, Scott Schuh, and Michael A. Zabek. 2010. "The 2008 Survey of Consumer Payment Choice." Federal Reserve Bank of Boston Public Policy Discussion Paper 09-10 (April).

Freeman, Scott. 1996. "The Payments System, Liquidity, and Rediscounting." *American Economic Review* 86 (December): 1,126–38.

Fuerst, Timothy S. 1992. "Liquidity, Loanable Funds, and Real Activity." *Journal of Monetary Economics* 29 (February): 3–24.

Fujiki, Hiroshi, and Migiwa Tanaka. 2009. "Currency Demand, New Technology and the Adoption of Electronic Money: Evidence Using Individual Household Data." IMES Discussion Paper Series (December).

Green, Edward J. 1999. "Money and Debt in the Structure of Payments." Federal Reserve Bank of Minneapolis *Quarterly Review* 23 (Spring): 13–29.

Grossman, Sanford J., and Laurence Weiss. 1983. "A Transactions Based Model of the Monetary Transmission Mechanism." *American Economic Review* 73 (December): 871–80.

Ireland, Peter N. 1994. "Money and the Gain From Enduring Relationships in the Turnpike Model." Federal Reserve Bank of Richmond Working Paper 94-07.

Jack, William, and Tavneet Suri. 2009. "Mobile Money: The Economics of M-PESA." Working Paper.

Jack, William, and Tavneet Suri. 2010. "The Risk Sharing Benefits of Mobile Money." Working Paper.

Kocherlakota, Narayana. 2005. "Discussion of 'From Private Banking to Central Banking: Ingredients of a Welfare Analysis.'" *International Economic Review* 46 (May): 633–6.

Kocherlakota, Narayana, and Neil Wallace. 1998. "Incomplete Record-Keeping and Optimal Payment Arrangements." *Journal of Economic Theory* 81 (August): 272–89.

Lacker, Jeffrey M. 1997. "Clearing, Settlement and Monetary Policy." *Journal of Monetary Economics* 40 (October): 347–81.

Lucas, Robert E., Jr. 1980. "Equilibrium in a Pure Currency Economy." *Economic Inquiry* 18 (April): 203–20.

Lucas, Robert E., Jr. 1990. "Liquidity and Interest Rates." *Journal of Economic Theory* 50 (April): 237–64.

Manuelli, Rodolfo, and Thomas J. Sargent. 2009. "Alternative Monetary Policies in a Turnpike Economy: Vintage Article." New York University Working Paper (June).

Mas, Ignacio, and Olga Morawczynski. 2009. "Designing Mobile Money Services: Lessons from M-PESA." *Innovations* 4 (April): 77–91.

Perez-Verdia, Carlos. 2000. "Transaction Costs and Liquidity." University of Chicago, PhD dissertation.

Romer, David. 1987. "The Monetary Transmission Mechanism in a General Equilibrium Version of the Baumol-Tobin Model." *Journal of Monetary Economics* 20 (July): 105–22.

Rotemberg, Julio J. 1984. "A Monetary Equilibrium Model with Transactions Costs." *Journal of Political Economy* 92 (February): 40–58.

Townsend, Robert. 1983. "Financial Structure and Economic Activity." *American Economic Review* 73 (December): 895–911.

Townsend, Robert. 1987. "Economic Organization with Limited Communication." *American Economic Review* 77 (December): 954–71.

Townsend, Robert. 1989. "Currency and Credit in a Private Information Economy." *Journal of Political Economy* 97 (December): 1,323–44.

Townsend, Robert, and Neil Wallace. 1982. "A Model of Circulating Private Debt." Federal Reserve Bank of Minneapolis Staff Report 83.

Wallace, Neil. 2000. "Knowledge of Individual Histories and Optimal Payment Arrangements." Federal Reserve Bank of Minnesota *Quarterly Review* 24 (Summer): 11–21.

Wallace, Neil. 2005. "From Private Banking to Central Banking: Ingredients of a Welfare Analysis." *International Economic Review* 46 (May): 619–31.

Wicksell, Knut. 1935. *Lectures on Political Economy*. London: MacMillan.