BORRADORES DE ECONOMÍA

A Mixed Duopoly in the Provision of Payment Services

By:
Carlos A. Arango-Arango
Yanneth Rocío Betancourt-García

No. 1280
2024

BANCO DE LA REPÚBLICA COLOMBIA

A Mixed Duopoly in the Provision of Payment Services

Carlos A. Arango-Arango Yanneth Rocío Betancourt-García[*]

Abstract

In this paper, we study the coexistence of cash and electronic payments introducing some distortions in the payments markets to understand the widespread use of cash, specially in emerging countries. Following Lagos and Wright (2005) we model explicitly some frictions in the exchange process considering money as essential. We introduce in this theoretical framework, theft and informality (measured by tax evasion), as factors affecting cash usage and, therefore competition with an electronic payment method. In this paper, segmentation in the payments market is considered by introducing heterogeneity in the seller's side, assuming different levels of productivity to explain the preference for cash or for electronic payments. Considering the above, the provision of the electronic payment platform is modeled under three different market structures to identify the effects of the distortions comparing the results with the social planner solution. In the first case, the electronic payment platform is provided by a public firm as a free service; in the second case a private monopoly provides the platform at a positive cost, and in the third case the conditions for the existence of a mixed duopoly are derived. The existence of a public provider in the electronic payments market could lead private networks to provide these services at a lower cost than in the monopoly case, increasing the coverage of digital payments and reducing cash usage, which implies gains in social welfare. This paper gives a theoretical basis and key insights to the discussions regarding public provision of new payment services when the market is already served by private suppliers.

Keywords: Cash, payment methods, payments services, electronic payments, instant payments.

JEL Classification: E40, E41,E42,E44

[*]Principal Researcher and Advisor, Banco de la República de Colombia. Email: carangar@banrep.gov.co; ybetanga@banrep.gov.co. This research took place through the visiting program under the Bilateral Assistance and Capacity Building for Central Banks (BCC), financed by SECO and the Graduate Institute of Geneva (IHEID). The authors thank Cyril Monnet for the academic supervision of this paper, as well as Professors Cedric Tille, Jean Charles Rochet and Fernando Alvarez for their valuable comments and suggestions.We also appreciate the comments of the participants and reviewers of the seminars at the IHEID, Banco de la República, CEMLA and the LAJCB Conference, 2023.

Duopolio Mixto en la Provisión de Servicios de Pago

Carlos A. Arango-Arango Yanneth Rocío Betancourt-García[1]

Resumen

Este artículo analiza teóricamente la coexistencia de dos medios de pago, como el efectivo y los pagos electrónicos, considerando algunas distorsiones en el mercado de pagos. Siguiendo a Lagos y Wright (2005), se modelan explícitamente algunas fricciones existentes en el proceso de intercambio. En este marco teórico, que considera al dinero como esencial, se introduce el robo y la informalidad (medida por la evasión fiscal), como factores que afectan el uso de efectivo y, por tanto, la competencia con otro medio de pago. Adicionalmente, se considera la segmentación en el mercado de pagos mediante la heterogeneidad en el lado de los vendedores, suponiendo diferentes niveles de productividad, para explicar la preferencia por el efectivo o por los pagos electrónicos.

Los efectos de estas distorsiones se modelan bajo tres estructuras de mercado diferentes en la provisión de los pagos electrónicos, las cuales se comparan con los resultados del planificador social. En el primer caso, la plataforma de pago electrónico es proporcionada por una empresa pública como un servicio gratuito; en el segundo caso, un monopolio privado proporciona la plataforma a un costo positivo, y en el tercer caso se analiza la existencia de un duopolio mixto en la provisión de estos servicios de pago. Se demuestra teóricamente que la existencia de un proveedor público en el mercado de pagos podría llevar a las redes privadas a proporcionar estos servicios a un costo menor que en el caso de un monopolio privado, aumentando la cobertura de los pagos digitales y reduciendo el uso de efectivo, lo que implica ganancias en el bienestar social. Este artículo proporciona una base teórica que puede ayudar a los debates actuales sobre la provisión pública de nuevos servicios de pago cuando el mercado ya cuenta con proveedores privados.

Palabras clave: Efectivo, medios de pago, instrumentos de pago, servicios de pago, pagos electrónicos, pagos instantáneos, duopolio mixto.

Clasificación JEL: E40, E41, E42, E44.

[1] Los autores son, respectivamente, Investigador principal y Asesora, de la Subgerencia de Sistemas de Pago y Operación Bancaria, Banco de la República, Colombia. Correos electrónicos: carangar@banrep.gov.co; ybetanga@banrep.gov.co. Este documento se realizó bajo el programa de *Visiting del Bilateral Assistance and Capacity Building for Central Banks (BCC)*, financiado por SECO y el *Graduate Institute of Geneva* (IHEID). Los autores agradecen la supervisión académica de Cyril Monnet, así como los valiosos comentarios y sugerencias de los profesores Cedric Tille, Jean Charles Rochet y Fernando Alvarez. Se agradece las sugerencias de los comentaristas y asistentes a los seminarios del IHEID, Banco de la República, CEMLA y de la Conferencia de LAJCB-2023.

1 Introduction

Rapid improvements in information technology and communications have led to a growing transformation in financial services, especially payment systems have undergone substantial innovation over the years. New payment methods, platforms, and channels have been introduced, and for retail payments innovations have addressed the entire value chain, closing service gaps and increasing the speed of payments.

The payments industry is experiencing a shift driven by the need for faster, cheaper, and more convenient means of payments. In many countries, central banks and monetary authorities have recently promoted faster payment systems (FPS) to provide a real-time, 24/7 fund-transfers (BIS, 2016). Over 60 jurisdictions have already launched FPS and a number of countries are planning to implement such systems in the coming years (CPMI-FPS, 2021).

Recently, digitalization in retail payments in emerging economies have increased rapidly, partially induced by the Covid-19 pandemic. Although strong growth in FPS volumes was observed between 2020 and 2022 in some Latin-American countries, such as Mexico, Brazil, Colombia, Chile, and Perú, some problems still prevail in terms of provision, access, acceptance and use of electronic instruments (FIS, 2020).

Despite these advancements, cash continues to dominate as the default mode of payment in the low-value retail segment, due mainly to a lack of financial inclusion and to the absence of alternative, secure, low-cost payment methods. Around the world, 1.7 billion adults are tied to cash as their only payment method, as they do not have a transaction account (World Bank, 2017). In Colombia, around 75% of individuals prefer cash to other payment instruments (Banco de la República, 2022), and the demand for cash (real pesos per capita) has grown annually above 5 % in the last 15 years (Arango-Arango et al, 2020).

The existence of market failures in the provision of electronic payments, the lack of interoperability between platforms and segmentation in the markets, have maintained the dominance of cash. The above added to the benefits of using cash, such as those related to tax evasion, could lead cash to continue being preferred in transactions, despite the existence of certain costs like theft. These distortions and inefficiencies in the payments market can generate a mix of payments methods that does not correspond to the socially optimal provision.

As payment services are essential to the well-functioning of the economy, governments and central banks face key policy questions. In order to guarantee a safe, efficient, and attractive payment system, policymakers have to weigh their role in the regulation of the payments ecosystem and their involvement in the provision of payment infrastructures. In this sense, they are asking if public provision of new payments technologies (as FPS) may fill up the gaps left by the private sector, to bring the economy closer to the public policy objectives. Indeed, a public provider may promote competition, innovation, low costs, and the entry of new players in the payment's ecosystem.

In this sense, in the recent past many emerging economies have opted for the public provision of FPS as a bet to leverage on interoperability, network effects and scale economies at the same time as solving coordination failures in the market for faster payments. This is the case of Brazil with its FPS PIX, which in a very short time, was able to bring onboard two thirds of its adult population (Duarte et al., 2022). In other jurisdictions, as in the colombian case, the private sector has developed instant payment solutions based on closed networks (e.g. Transfiya) that create a fragmented market of walled mobile payment gardens. However, this system, as many private FPS in other economies, has a limited number of payment service providers as participants, lacks interoperability with other payment rails, and it is limited to few use cases such as person to person and hardly make inroads into other payment flows, such as person to merchants, excluding a large share of retail transactions (Arango-Arango, et al. 2021).

Public provision of an FPS infrastructure is relevant as a policy mechanism to induce the private sector to enhance its fast payment solutions. The competition of private and public infrastructures is not strange to the provision of payment services. For large value payment systems, in Europe the public utility TARGET competes with EURO1, the private one, and in the US Fedwire (public) competes with CHIPS (private). In the case of retail payment systems, the Fed has announced the implementation of an FPS (FedNow) that would compete with RTP, the private FPS launched by the Clearing House. In Colombia, CENIT, the public ACH owned by the central bank, competes with ACH Colombia, the private provider of interbank transfers in batch.

This paper develops a theoretical model to analyze the provision of an electronic payment method, that compites with cash, under different market structures. From this approach we find some key insights to the policy discussions regarding public provision of a payment platform, when the market is already served by private suppliers. It also makes an important contribution to the understanding of how public provision shall improve the payments ecosystem in terms of safety and efficiency, increasing the social welfare by lifting the existing inefficiencies.

2 General Framework

The New Monetarism Theory is an appropriate framework to study these topics because it treats money as essential. This macroeconomic approach explains the behaviour of the institutions that facilitate the exchange process, like money, banks and financial intermediaries using microfoundations. It also allows to explicitly model the frictions that can exist in trade. A relatively new subbranch of this theory studies the economics of payments, analyzing the payments systems and giving theoretical tools to study frictions in payments markets.

Following this approach, in this paper we model the payments market introducing some distortions, that can affect the agents´decisions about which means of payment prefer to use in their transactions. In the Lagos & Wright setup, we introduce two alternatives means of payments: cash and an electronic payment instrument, modeling some frictions that can exist in these markets. Agent's payment decisions, about which payment method they prefer, depends on the costs and benefits that each one has for the different agents in the economy (buyers and sellers). These decisions are influenced by some factors like theft, informality, safety, confidentiality, among others, which are involved in cash and electronic transactions, and can affect the general equilibrium outcome in the payments market and therefore the social welfare.

Although, both sides of the payments market, buyers and sellers, can be affected by the same or different distortions, in this paper we model some of them for simplicity and to focus on only one side of the market: the seller´s side. Here, we introduce theft only in the seller's side, as a cost for sellers to accept cash payments because they can be robbed with some probability. Even though, buyers' decisions can be also affected by theft, we do not introduce it in the model. We consider another distortion in the seller´s side, tax evasion, which can be very important for merchants´decisions, especially in emerging economies, where some merchants could prefer accept cash to evade taxes and remain in the informal sector. In this context, tax evasion could be interpreted as a measure of the informality level in the economy.

The differences between agents is also important to understand why people choose one method or another to make their payments, depending on their own characteristics or preferences. Although heterogeneity could exist on both sides, agent heterogeneity is modeled here only on the sellers' side, which is defined by their productivity level, assuming that small merchants (which have a low level of productivity) may prefer to accept cash payments because they are less expensive than electronic payments or because of the possibility of evading taxes.

Considering all these factors and assumptions in a theoretical model, following Lagos and Wright setup, we identify the inefficiencies and distortions in the payments market by comparing the first-best solution (without frictions) with the equilibrium solution in the cash market and in the electronic market with frictions in the seller´s side of these markets. We model these markets separately to analyze different market structures in the provision of the electronic payment platform. First, we consider the case when a public provider offers the electronic platform as a free public good, which means that there is no charge to the sellers or buyers for using this platform. Then we model the case where there is a private monopoly that maximizes its profits, given a cost of provision of the platform and a fee charged to the final users, in this case we assume that only sellers have to pay for using the private platform. Finally, we find the conditions that must be satisfied to have a mixed duopoly in the provision of a payment system, where a public and a private providers could coexist and offer the electronic platform at the lower cost.

3 The Model: Basic Setup

Following the Lagos and Wright benchmark, our basic setup has the following main assumptions, which are shown in Figure 3.1. There are two types of agents, buyers and sellers, and there is a measure one of both. Time is discrete and continues forever $t = 0, 1, ...$ and all agents discount the future at the rate β. Each period is divided into two subperiods (day and night).

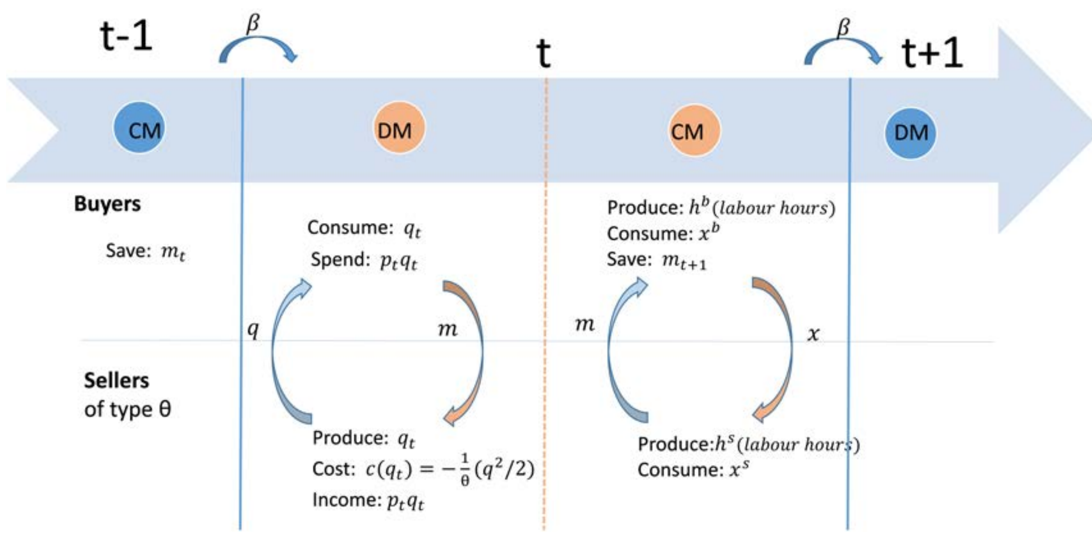

Fig. 3.1: Basic Setup

During the day, trade occurs in a decentralized market (DM) between buyers and sellers. Some agents, known as sellers, produce a good q but they don't consume in this period. The good q, is specialized and non-storable, and it is produced at a cost $c(q;\theta)$, which is increasing and convex: $c'(q;\theta) > 0, c''(q;\theta) > 0$ and $c(0;\theta) = 0$. We suppose that this function is given by: $c(q;\theta) = \frac{1}{\theta} q^2/2$, where θ is a measure of the seller's productivity (the higher θ the less costly it is to produce the good) and we assume θ is uniformly distributed over the interval $[0, 1]$. There are other agents in this market, buyers, who want to consume but they do not produce the consumption good (q) during this period. By consuming the good they get a level of utility, which is given by an increasing and concave function $u(q)$, where $u'(q) > 0$, $u''(q) < 0$ and $u(q) = 0$. In this period (DM) the price of the consumption good (p_t) is determined in a Walrasian market.

During the night, there is a centralized market (CM) where both sellers and buyers can produce and consume. Both agents produce a general good x, which is not storable. Agents are endowed with labour (h) and the production function is linear, so that one unit of labour produces one unit of good $h = x$. They have a linear utility function $U(x, h) = x - h$. Agent's utility functions are separable across subperiods.

The central bank provides money in a quantity $M \in \mathbb{R}_+$, which is intrinsically useless, divisible and storable. In this model, money has a value in terms of the good q, which is denoted by ϕ, so that, the real value of money in each period is given by: $\phi_t M_t$. Also, the real price of the good q in DM is given by $\phi_t p_t$, where p_t is the nominal price of it. The central bank adjusts the money supply M_t by making transfers to buyers in CM. The money supply grows at a rate γ, so that in a stationary equilibrium the amount of real balances is constant $\phi_t M_t = \phi_{t+1} M_{t+1}$, and inflation is given by $\phi_t / \phi_{t+1} = \gamma$.

3.1 First Best solution

We solve the model to find the first best solution, assuming that there is a social planner and there are not distortions in the payments markets (there are not theft, not tax evasion and not agent's heterogeneity). The first best solution is obtained by solving the following optimization problem for the society without frictions. Since preferences are linear in the CM for all agents, all the redistribution effects do not have any impact on aggregate social welfare. So the planner's problem is given by:

$$\max_{\{q(\theta)\}} u\left(\int q(\theta)d\theta\right) - \int \left[\frac{1}{\theta}\frac{q(\theta)^2}{2}\right] d\theta.$$

From the first order conditions with respect to an arbitrary $q(\theta)$, we obtain,

$$\theta u'(q^*) = q^*(\theta)$$

where $q^* = \int q^*(\theta)d\theta$. Given that θ has a uniform distribution $\int_0^1 \theta d\theta = \frac{1}{2}$ and integrating the first order condition over all θ, we find the optimal quantity q^* from the following equation:

$$\frac{u'(q^*)}{2q^*} = 1 \tag{3.1}$$

which implies that the marginal utility for all the buyers have to be equal to the average marginal cost of the sellers (given that they do not have the same cost). Therefore, the disutility generated by the production of the good must be equal to the utility that it gives to other agents in the society.

3.2 Two payment methods with frictions

In the basic setup, we introduce two payment methods: cash, issued by the central bank; and an electronic payment method, for which sellers have to pay for a terminal to accept it. First, we consider the case where the electronic payment platform is provided by a public authority as a free public good, which implies that there is no cost for final users to use it. Then, we consider the private provision of the platform by a monopolistic firm, which offers it at a positive cost. Finally, we discuss the mixed duopoly case, where a private firm would compete with a public provider, who can charge a price equal to the marginal cost for the use of the platform.

3.2.1 Public Provision

In this case, we assume that the electronic payment method is offered by a public provider like a free public good, without cost for the final user. In DM, a buyer who meets a seller with a terminal can pay with the electronic method for the amount of the purchase. Although buyers are constrained by their budget to make their purchases, this constraint is the same for both payment methods, which means that it does not make any difference between the two, for this reason, this constraint is not modeled here. However, when buyers use cash, they are constrained by their cash holdings, so that they face a cash in advance constraint. Meanwhile, when they pay digitally, they are not constrained and the electronic platform provider can enforce the repayment of the buyer in the following period (CM). For this reason, sellers requires the electronic platform because buyers are anonymous and lack of commitment.

For simplicity, we analyze each market separately to identify the effects of frictions for each payment method. We assume that there is a market for each payment method and that in each market the other payment instrument is not accepted (e.g. a buyer holding cash not trade in the market where only the electronic payment instrument is accepted). In each market, buyers and sellers maximize their value functions to choose the optimal quantity of goods they want to buy or sell, given the amount of the corresponding payment method they hold. We introduce some frictions on each market, cash payments are subject to theft, unlike electronic payments and cash payments are used to evade taxes because they can be hidden, while receiving electronic payments leads to paying taxes (see Figure 3.2.).

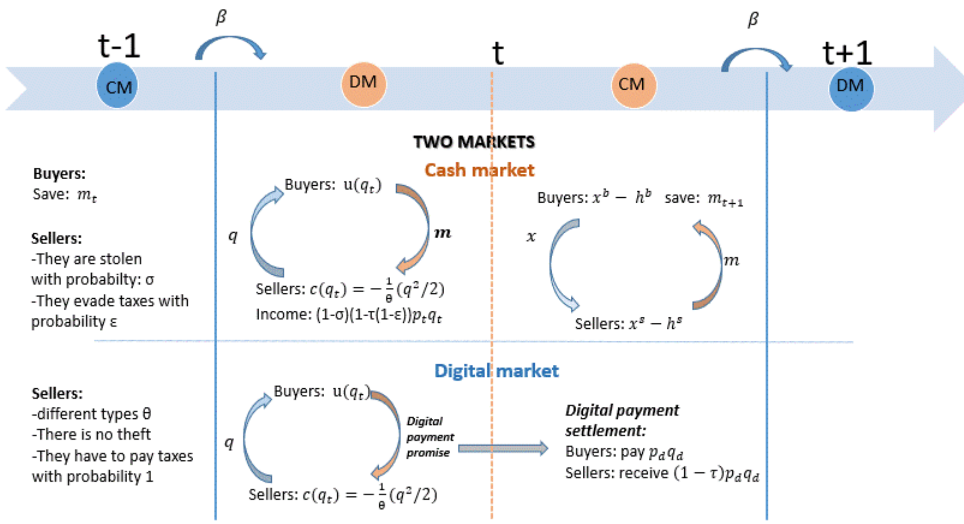

Fig. 3.2: Two payment methods with frictions

We assume that at the end of CM, cash can be stolen from sellers with a probability σ. Therefore they enter to DM in the next period with a fraction $(1-\sigma)$ of the money they obtained in the previous period. Tax evasion is modeled here as a benefit of cash transactions, assuming that with probability ε sellers can evade and not pay taxes (without being caught) when they accept cash. By accepting electronic payments sellers have to pay taxes τ with probability 1.

In our model buyers have no cost of accessing the electronic payment. Therefore, buyers will always hold this instrument. We suppose that sellers can only accept one payment method in DM, and they choose which one to accept. We guess and verify that sellers with productivity level below a threshold $\bar{\theta}$ prefer to accept only cash, while sellers with a higher productivity prefer to accept electronic payments. This translates in an endogenous market segmentation for payments: a fraction $\pi = \bar{\theta}$ of sellers accept cash while the remaining $1-\bar{\theta}$ only accept electronic payments. The treshold $\bar{\theta}$ is given by the productivity level that leaves the seller indifferent between accepting cash or electronic payments, so that the profits for the seller are the same in both markets.

It will be useful to define the mean productivity in both markets: in the cash market the mean productivity is given by: $\int_0^{\bar{\theta}} \theta d\theta = \frac{\bar{\theta}^2}{2}$, while in the market with electronic payments it is $\int_{\bar{\theta}}^1 \theta d\theta = \frac{(1-\bar{\theta}^2)}{2}$.

The cash market:

Buyers' problem: Given that all buyers have access to both payment instruments and that we do not assume any friction on the buyers´side, they are constraint in this market only by the amount of sellers who accept cash $(\pi = \bar{\theta})$. So that, buyers choose the optimal amount of money to maximize the value function in the centralized market $W(\tilde{m})$, considering the value function in the decentralized market, where they can trade with probability π in the cash market and with probability $(1-\pi)$ in the electronic market.

$$W(\tilde{m}) = \max_{x,h,m} x - h + \beta\left[\pi V_{cash}(m) + (1-\pi) V_{electronic}(m)\right]$$

s.t.

$$x - h = \phi\tilde{m} - \phi m$$

where:

$V_{electronic}(m)$ is the value function in DM next period when they trade with electronic payments:

$V_{electronic}(m) = \max_{q_d} \quad u(q_d) - \phi p_d q_d + W(m)$

$V_{cash}(m)$ is the value of cash in DM next period and is given by:

$V_{cash}(m) = \max_{q} \quad u(q) + W(m - pq)$

subject to $\phi pq \leq \phi m$

Using quasi-linearity, we can write: $W(m - pq) = -\phi pq + W(m)$.

And from the first order condition with respect to q we know that:

$$u'(q) - (1+\lambda)\phi p_{t+1} = 0 \qquad (3.2)$$

Substituting these expressions in $W(\tilde{m})$, we can solve the problem for m and get the following first order condition,

$$-\phi_t + \beta\left(\pi\lambda\phi_{t+1} + W'(m)\right) = 0$$

where the marginal value of unspent cash is $W'(m) = \phi_{t+1}$. Substituting λ from 3.2, we get the Euler-equation in the cash market:

$$\frac{\phi_t}{\beta\phi_{t+1}} = 1 + \pi\left(\frac{u'(q_{t+1})}{\phi_{t+1}p_{t+1}} - 1\right) \qquad (3.3)$$

Sellers' problem: Because sellers face theft with a probability σ when they accept cash, they only receive a fraction of what buyers paid $(1-\sigma)\phi pq$. They can also evade taxes with probability ε if they receive cash. So that, they receive a fraction $(1-\tau(1-\varepsilon))$ of cash payments and they pay a fraction τ on taxes with probability $(1-\varepsilon)$. The seller's problem is the same no matter which event comes first (theft or tax evasion). Given the real price in the DM is ϕp, a seller of type θ supplies q to maximize the following expression:

$$\max_{q} (1-\sigma)(1-\tau(1-\varepsilon))\phi pq - c(q;\theta)$$

Using $c(q) = \frac{1}{\theta} q^2/2$, we get the optimal supply:

$$q(\theta) = \theta(1-\sigma)(1-\tau(1-\varepsilon))\phi p \tag{3.4}$$

Therefore more productive sellers (high values of θ) are also "larger" in the sense that they have more sales.

The total supply in this market is given for the aggregate quantities $q(\theta)$ offered by the sellers of low type $\theta \leq \bar{\theta}$:

$$\int_0^{\bar{\theta}} q(\theta)\, d\theta = (1-\sigma)(1-\tau(1-\varepsilon))\phi p \int_0^{\bar{\theta}} \theta d\theta = \frac{\bar{\theta}^2(1-\sigma)(1-\tau(1-\varepsilon))\phi p}{2}$$

- Market clearing requires that the above expression equals total demand q_t and solving for the (real) price we obtain:

$$\phi_t p_t = \frac{2q_t}{\bar{\theta}^2(1-\sigma)(1-\tau(1-\varepsilon))} \tag{3.5}$$

- Substituting this equation in the Euler equation (3.3) we get the equilibrium allocation q in the cash market:

$$\frac{\gamma}{\beta} = 1 + \bar{\theta}\left[\frac{u'(q)}{2q}\bar{\theta}^2(1-\sigma)(1-\tau(1-\varepsilon)) - 1\right] \tag{3.6}$$

where we have used $\pi = \bar{\theta}$ and in steady-state we have: $q_t = q_{t+1} = q$ and $\gamma = \frac{\phi_t}{\phi_{t+1}}$.

Rewritten this expression we have:

$$\frac{u'(q)}{2q} = \frac{\gamma - \beta\left(1-\bar{\theta}\right)}{\beta\bar{\theta}^3(1-\sigma)(1-\tau(1-\varepsilon))} \tag{3.7}$$

The electronic market:

Buyers' problem: In the electronic market, buyers' maximizes their value function in DM to optimize the quantity of the good demanded in this market q_d.

$$\max_{q_d} u(q_d) - \phi p_d q_d$$

with first order condition,

$$u'(q_d) = \phi p_d \tag{3.8}$$

Sellers´problem: Sellers in the electronic market do not face theft $\sigma = 0$, but they cannot evade taxes $\varepsilon = 0$. They pay a fraction τ on taxes for sure to the government and they receive only a fraction of the payment $(1-\tau)\phi p_d q_d$. In the DM with electronic payment a seller with productivity θ maximizes the following objective function :

$$\max_{q_d} (1-\tau)\phi p_d q_d - \frac{1}{2\theta} q_d^2$$

from which we obtain the following first order condition:

$$q_d(\theta) = \theta(1-\tau)\phi p_d \tag{3.9}$$

Therefore, the total supply in the electronic market from the sellers with high productivity (high type $\theta \geq \bar{\theta}$):

$$\int_{\bar{\theta}}^{1} q_d(\theta)\, d\theta = (1-\tau)\phi p_d \int_{\bar{\theta}}^{1} \theta d\theta = (1-\tau)\phi p_d \frac{\left(1-\bar{\theta}^2\right)}{2}$$

- Using market clearing and equalizing this expression with the total demand q we obtain:

$$\phi_t p_{d_t} = \frac{2q_{d_t}}{\left(1-\bar{\theta}^2\right)(1-\tau)} \tag{3.10}$$

- Substituting this equation in the Euler equation we get the equilibrium allocation q_d in the electronic market:

$$\frac{u'(q_d)}{2q_d} = \frac{1}{\left(1-\bar{\theta}^2\right)(1-\tau)} \tag{3.11}$$

Optimal level of sellers in each payment market $\bar{\theta}$: To find the level $\overline{\theta}$ where sellers are indifferent between accepting cash or electronic payments we equalize the sellers' profits in both markets:

- Profits in cash market: $(1-\sigma)(1-\tau(1-\varepsilon))\phi pq - \frac{1}{2\theta}q^2$
- Profits in electronic market: $(1-\tau)\phi p_d q_d - \frac{1}{2\theta}q^2$

substituting the optimal quantities q and q_d in each market from equations (3.4) and (3.9) and ϕp, ϕp_d from equations (3.5) and (3.10), we obtain[2]:

$$\bar{\theta} = \sqrt{\frac{q}{q+q_d}}. \tag{3.12}$$

The equilibrium conditions in this setup with a electronic platform without cost, heterogeneity in the sellers' side, theft of cash and tax evasion are given by:

[2] So that: profits in cash are given by $\frac{2\bar{\theta}q^2}{\bar{\theta}^4}$, and profits in electronic market are: $\frac{2\bar{\theta}q_d^2}{\left(1-\bar{\theta}^2\right)^2}$

1. Stationary equilibrium allocation in the cash market (equation 3.7):

$$\frac{u'(q)}{2q} = \frac{\gamma - \beta\left(1 - \bar{\theta}\right)}{\beta\bar{\theta}^3 (1-\sigma)(1-\tau(1-\varepsilon))}$$

2. Stationary equilibrium allocation in the electronic market (equation 3.11):

$$\frac{u'(q_d)}{2q_d} = \frac{1}{\left(1-\bar{\theta}^2\right)(1-\tau)}$$

3. Optimal threshold of sellers in each market (equation 3.12):

$$\bar{\theta} = \sqrt{\frac{q}{q+q_d}}$$

Comparing these results with the optimal condition (equation 3.1) we find that the frictions introduced in our model creates some distortions in the equilibrium allocations for both markets.

In the cash market, the distortions are given by inflation γ, theft σ and tax evasion ε; while in the electronic market only taxes have a direct impact τ on the equilibrium allocations. Moreover, given that there is segmentation in the seller's side, these frictions have an indirect impact in the traded quantities in both markets via the level of indifferent sellers between the two markets $\bar{\theta}$.

The equilibrium allocation in the cash market q is affected positively by the amount of taxes τ and the probability of evasion ε, so that if taxes increase and the probability of evasion is high, the sellers prefer to accept cash and the quantity traded in this market increases. The impact of an increase in inflation γ and theft σ on the equilibrium allocation in the cash market is negative because these are costs for the sellers of accepting cash. Finally, if the level of sellers who are indifferent between the two payment methods is high, so that more sellers prefer to accept cash then the quantity traded in the cash market increases.

In the electronic market, the equilibrium allocation is affected negatively by the level of taxes and the amount of sellers that accept cash. It means that if taxes are high less sellers will accept electronic payments and the quantity traded decreases. The same happens if more sellers prefer to accept cash.

The impact of inflation and theft on the level of sellers that are indifferent between the two markets $\bar{\theta}$ is negative. It means that less sellers accept cash when the costs of inflation and theft are high. In contrast, when the level of taxes and the probability of evasion are high, sellers prefer to accept cash, then the quantity of sellers accepting cash increases.

3.2.2 Private Monopoly Provision

Now we assume that there is only one private provider of the electronic payment platform, who charges a positive fee $\psi > 0$ directly to final users. We assume that the buyers do not face any cost to use this platform, so that all of them have access to both payment methods (cash and electronic), which implies that there is not distortion in the buyers' side. In this case, only sellers who want to access to this platform are charged with the fee. Now, the existence segmentation in the payments market is affected by the platform's fee, so that the amount of sellers who accept cash depends on this fee $\bar{\theta}(\psi)$. For simplicity, we assume that the electronic platform's fee does not depend on the quantity that is traded in this market q_d, so it is fixed and maximizes the monopolist's profits.

The problem for buyers and sellers in the cash and the electronic market in terms of the quantities traded does not change. So the equilibrium conditions in both markets are maintained (equations 3.7 and 3.11), only the number of sellers in each market changes $\bar{\theta}$, because the sellers' profits in the electronic market are affected directly by the fee and both market's profits are affected indirectly by $\bar{\theta}(\psi)$.

The profit in the cash market is:

$$(1-\sigma)(1-\tau(1-\varepsilon))\phi pq - \frac{1}{2\theta}q^2 = \frac{2\bar{\theta}(\psi)\,q^2}{\bar{\theta}(\psi)^4}$$

while the profit in the electronic market is:

$$(1-\tau)\phi p_d q_d - \frac{1}{2\theta}q_d^2 - \psi = \frac{2\bar{\theta}(\psi)\,q_d^2}{\left(1-\bar{\theta}(\psi)^2\right)^2} - \psi$$

Equalizing both expressions we obtain:

$$\psi = \frac{2\bar{\theta}(\psi)\,q_d^2}{\left(1-\bar{\theta}(\psi)^2\right)^2} - \frac{2\bar{\theta}(\psi)\,q^2}{\bar{\theta}(\psi)^4}$$

From the Euler's equation in both markets we substitute q and q_d to get:

$$\psi = \frac{\bar{\theta}(\psi)}{2}\left[(1-\tau)\,u'(q_d) - \frac{\bar{\theta}(\psi)\,\beta(1-\sigma)(1-\tau(1-\varepsilon))\,u'(q)}{\gamma-\beta\left(1+\bar{\theta}(\psi)\right)}\right]^2 \tag{3.13}$$

Monopoly problem: The private provider chooses a fee ψ to charge to sellers of type $\theta \geq \bar{\theta}$, such that it maximizes its profits. The monopoly faces some costs related to the provision of the infrastructure and all the services needed to make electronic payments (as the provision of terminals and some operating costs). From equation (3.13), we have an expression for ψ in function of $\bar{\theta}$, so that solving the following problem for $\bar{\theta}$ is equivalent to solve for ψ:

$$\max_{\bar{\theta}} \left(1-\bar{\theta}\right)(\psi - cost)$$

$$s.t.\psi = \frac{\bar{\theta}(\psi)}{2}\left[(1-\tau)\,u'(q_d) - \frac{\bar{\theta}(\psi)\,\beta(1-\sigma)(1-\tau(1-\varepsilon))\,u'(q)}{\gamma-\beta\left(1+\bar{\theta}(\psi)\right)}\right]^2$$

From the first order condition with respect to $\bar{\theta}$ and equation 3.13 we find an expression for the optimal fee $\hat{\psi}$, which is a polynomial in $\bar{\theta}$ and depends on all the parameters of the model. By making some assumptions for the parameters[3] and taking a CRRA utility function[4], we solve this expression for $\hat{\psi}$. We find an inverse relationship between $\hat{\psi}$ and the maximum level of sellers that accept cash $\bar{\theta}$, as expected. It means that if the private monopoly platform charges a high fee $\hat{\psi}\uparrow$, less sellers will accept electronic payments ($(1-\bar{\theta})$ decreases). Figure 3.3. shows that if the level of taxes is high (τ increases) sellers will prefer to use cash ($\bar{\theta}$ increases). From figure 3.4. we observe that if inflation increases more sellers will prefer electronic payments instead of cash ($\bar{\theta}$ decreases).

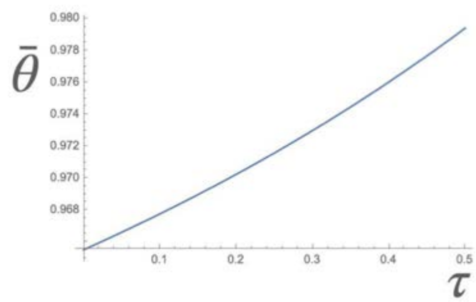

Fig. 3.3: Threshold of sellers and taxes

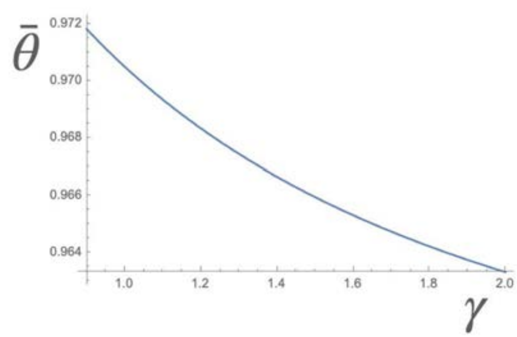

Fig. 3.4: Threshold of sellers and inflation

These results imply that if the electronic platform is provided by a private

[3] The assumptions we made are: $\rho = 0.8$, $\tau = 0.2$, $\beta = 0.9$, $\varepsilon = 0.2$, $\sigma = 0.5, \gamma = 1.02$ and $cost = 0.5$.

[4] $u(q) = \frac{q^{1-\rho}}{1-\rho}$ where ρ is the coefficient of relative risk aversion.

monopoly, which charges a fee to the sellers, the equilibrium allocations continue being suboptimal like in the public provision case, given that the distortions in both markets are the same. However, in the case of a private provider the equilibrium quantities are distorted by the fee, via the amount of sellers who are indifferent between the two markets ($\bar{\theta}$), because sellers face another cost if they want to access to the electronic platform. Now, more sellers prefer to accept cash ($\bar{\theta}$ increases) given that the private electronic platform is more costly than the public one. It means that if the public provider does not charge any fee (or a low fee) to the sellers for using the electronic platform, it is more efficient for the economy to have only the public platform for electronic payments. However, given that the private provider faces some costs to provide the electronic platform it has no way to compete with the free public platform at zero cost.

3.2.3 A Mixed Duopoly in the provision of electronic payments

In the mixed duopoly setup, the private provider competes with the public (government) provider. In this case, we suppose that the government charge some fee to the final users of the electronic platform (it is not a free public good). We assume that the public provider sets the fee ψ_g equals to the cost of providing the payment service $cost_g$. Since there is no reason to believe that the public cost ($cost_g$) equals the private cost ($cost$) , we have to consider several possible scenarios.

When the public technology is expensive, $\psi_g = cost_g > \psi$, the government cannot provide a viable payment service to compete with the private one, so that in this case there cannot exist a mixed duopoly, and only the private monopoly will exist.

When the public technology is in a medium range, $\psi_g \in [cost, \psi]$, the government can contest the market to the private payment service. Since $\psi_g \leq \psi$, the private provider has no choice but to match the public fee as otherwise the private provider would have no demand. If the private provider matches the public fee, we assume that it obtains half the demand for electronic payment service: $\left(1 - \bar{\theta}(\psi_g)\right)/2$, where $\bar{\theta}(\psi_g)$ is the threshold of merchant's type when the fee is ψ_g. However, the private provider can do better by setting its fee below but close to ψ_g, so that the private provider captures the entire market with profit (almost) equal to $\left(1 - \bar{\theta}(\psi_g)\right)(\psi_g - cost)$. In this case, the public provider while usefully contesting the market only gets some residual demand, if any. The market is split between the two providers only in the knife-edge case where $cost_g = cost$.

Finally, if the public provider uses a better technology than the private provider ($cost_g < cost$), the government can become the sole provider of electronic payments. It is however difficult to envisage that this could be the case.[5]

To conclude, a mixed duopoly in the electronic payments market will be possible only if both platforms (the public and the private) charges the same fee to sellers, such that both providers coexist in the market. Although the costs of

[5] If this was possible, the government could always "give" its technology to the private provider, while competing with it at the same time.

both providers could be different, it could be possible only when one provider offers the platform at the lower cost and the other in order to compete adjust its costs and charges the same fee. Then, the optimal result for the payments market is the provision of the payment platforms by a mixed duopoly, given that the competition of a public provider could lead to the monopoly private provider to reduce its costs and charge a lower fee to the sellers, which should be the same for both platforms and it has to be equal to the lower cost of provision (the marginal cost). This result also implies a welfare gain because the electronic platform will be provided at a lower cost than in the monopoly case, and more people will be willing to use this payment method.

4 Conclusions

This paper aims to contribute to the literature on the provision of payment services in markets with public and private operation. Electronic payments, in general, can have significant impacts on means of payment usage when there exist theft and tax evasion. Segmentation in payments systems limits the benefits of the existence of electronic payments given the incentives that some merchants have regarding the use of cash, mainly because the existence of some distortions in this market, as tax evasion. Although, cash theft could limit these incentives, it also imposes distortions on payments outcomes and welfare. Policy instruments including direct policies on theft and tax evasion are important to solve these market failures.

Moreover, the existence of a public provider in the electronic payments market could lead private networks to provide these services at a lower cost, increasing the coverage of digital payments and reducing cash usage. However, ceteris paribus, the lack of interoperability between platforms perpetuates some of the distortions associated with cash and limits the digitalization of payments. Future research would analyze how interoperability between platforms could impact means of payment usage and welfare. Furthermore, because the electronic payment markets are affected by economies of scale it is useful for policymakers to take into account these characteristics in the modeling.

References

Arango-Arango, C., A. Ramírez-Pineda y M. Restrepo-Bernal, (2021). Person-to-business instant payments: Could they work in Colombia? Journal of Payments Strategy & Systems, Volume 15 / Number 4 / Winter 2021-22, pp. 376-397(22).

Arango-Arango, C, F., Arias, N. Rodríguez, N. Suárez y H. Zárate (2020). “Efectivo y pagos electrónicos”. Ensayos sobre Política Económica No. 93, Banco de la República.

Arango-Arango,C., Y. R. Betancourt-García, M. Restrepo-Bernal y G. Zuluaga-Giraldo, (2021). “Pagos electrónicos y uso del efectivo en los comercios colombianos 2020". Borradores de Economía, No. 1180, Banco de la Republica, Colombia.

Aruoba, S. Borağan, 2021. "Institutions, tax evasion, and optimal policy," Journal of Monetary Economics, Elsevier, vol. 118(C), pages 212-229.

Bech, M., U. Faruqui, F. Ougaard, C. Picillo (2018). "Payments are a-changin' but cash still rules". BIS Quarterly Review, March 2018.

BIS. (2016). "Fast payments – Enhancing the speed and availability of retail payments". November, Committee on Payments and Market Infrastructures.

CPMI-FPS. (2021). Developments in retail fast payments and implications for RTGS systems.

Duarte, A., J. Frost, L. Gambacorta, P. K. Wilkens and H. S. Shin, (2022). "Central banks, the monetary system and public payment infrastructures: lessons from Brazil's Pix". BIS bulletin, No. 52.

FIS. (2020). "Flavors of Fast".

Graeme, G. and J. Wright, (2007). "Competing Payment Schemes", Journal of Industrial Economics, Wiley Blackwell, vol. 55(1), pages 37-67, March.

Gu, C., Mattesini, F., Monnet, C., & Wright, R. (2013). "Banking: A New Monetarist Approach". The Review of Economic Studies, 80(2 (283)), 636–662.

Lagos, R., & Rocheteau, G. (2009). "Liquidity in Asset Markets with Search Frictions". Econometrica, 77(2), 403–426.

Lagos, Ricardo, Guillaume Rocheteau, and Randall Wright. 2017. "Liquidity: A New Monetarist Perspective." Journal of Economic Literature, 55 (2): 371-440.

Lagos, R., & Wright, R. (2005). "A Unified Framework for Monetary Theory and Policy Analysis". Journal of Political Economy, 113(3), 463–484.

Mohammed, A.L. (2020). "Informality, Frictional Markets and Optimal Policy". Job Market Paper.

Nosal, Ed. & Rocheteau, G. (2011). "Money, Payments, and Liquidity," MIT Press Books, The MIT Press, edition 1, volume 1, number 0262016281, December.

Reporte de Infraestructura Financiera. (2022). Banco de la República, Colombia.

Weinberg, J. A., (2002). "Imperfect competition and the pricing of interbank payment services", Economic Quarterly, Federal Reserve Bank of Richmond, issue Win, pages 51-66.

World Bank, (2017). "Fast payment systems: Preliminary Analysis of Global Developments". December, 2017.