BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2014-22

# Understanding the Cash Demand Puzzle

by Janet Hua Jiang and Enchuan Shao

Bank of Canada Working Paper 2014-22

June 2014

# Understanding the Cash Demand Puzzle

by

**Janet Hua Jiang[1] and Enchuan Shao[2]**

[1]Funds Management and Banking Department
[2]Currency Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
jjiang@bankofcanada.ca
eshao@bankofcanada.ca



ISSN 1701-9397 

# Acknowledgements

For their comments and discussions, we would like to thank Jonathan Chiu, Ben Fung, Scott Hendry, Miguel Molico, Héctor Pérez Saiz, Lukasz Pomorski, Francisco Rivadeneyra, Gerald Stuber, Liang Wang, Russell Wong, and seminar participants at the Bank of Canada, University of Iowa and University of Hawaii. We would also like to thank Glen Keenleyside for his excellent editorial assistance.

# Abstract

We develop a model to explain a puzzling trend in cash demand in recent years: the value of bank notes in circulation as a percentage of GDP has remained stable despite decreasing cash usage at points of sale owing to competition from alternative means of payment such as credit cards. The main feature of the model is that cash circulates between economic activities where the substitutability between cash and other means of payment is uneven. Our model predicts that, once credit expands beyond a certain level, agents adjust their cash management practices in response to further credit expansions, causing the velocity of cash to slow down, so that the demand for cash can remain flat despite diminishing cash transactions.



*JEL classification: E41, E51*
*Bank classification: Bank notes; E-money; Credit and credit aggregates*



# Résumé

Nous élaborons un modèle permettant d’expliquer une tendance curieuse en ce qui a trait à la demande de monnaie observée ces dernières années : la valeur des billets de banque en circulation, exprimée en pourcentage du PIB, est demeurée stable, et ce, malgré la diminution de l’utilisation d’espèces aux points de vente sous l’effet de la concurrence d’autres modes de paiement comme les cartes de crédit. Ce modèle est caractérisé par le fait que l’argent circule entre des secteurs d’activité économique au sein desquels la substituabilité entre les espèces et les autres modes de paiement est inégale. Selon les prévisions de notre modèle, lorsque l’utilisation du crédit atteint un certain niveau, les agents revoient leurs pratiques de gestion de l’argent comptant en réaction à toute nouvelle expansion du crédit. Cela provoque un ralentissement de la vitesse de circulation des espèces, de sorte que la demande de monnaie reste constante en dépit de la baisse des transactions réglées comptant.



*Classification JEL : E41, E51*
*Classification de la Banque : Billets de banque; Monnaie électronique; Crédit et agrégats du crédit*

# 1 Introduction

The retail payment landscape has undergone significant changes in the past few decades with the emergence of various new payment instruments. In particular, cash has been losing ground to other means of payment at the point of sale. For example, according to Arango et al. (2012), in Canada, cash accounted for nearly 80% of the volume and more than 50% of the value of point-of-sale transactions in the early 1990s; in 2011, these numbers dropped to about 40% and less than 20%, respectively.[1] The reduction in cash usage has been picked up by an expanded usage of competing payment methods. For instance, during the same period, the share of credit card payments rose steadily from about 10% to more than 20% in volume, and from less than 40% to 50% in value. Given the diminishing popularity of cash usage at points of sale, one would expect the demand for cash to decrease. However, the total demand for cash, measured as the value of bank notes in circulation as a percentage of GDP, has remained more or less flat at around 3.5% in the past three decades (see Figures 4 and 5 in Fung et al. 2014).

Similar trends are also observed in other countries. For example, Bailey (2009) documents that, in the United Kingdom, the value of bank notes in circulation as a percentage of GDP rose slightly from 2.4% in the mid-1990s to 3.2% in 2009, while cash transactions fell gradually in terms of both volume and value. Recently, Bagnall et al. (2014) have observed the surprising resilience of cash demand in a group of industrial countries, including Austria, Australia, Canada, Germany and the United States. At the same time, there is evidence that cash transactions have been crowded out by other means of payment. For example, Wang and Wolman (2014) use the scanner data from a large discount chain store in the United States from April 2010 to March 2013, and find that the fraction of cash transactions fell at a rate of between 1.3% and 3.3% per year, depending on the size of transactions. The Reserve Bank of Australia conducted two consumer payments use diary studies in 2007 and 2010, respectively. During the period, the share of the number (value) of cash transactions decreased by 6% (5%) (see Bagnall et al. 2011). In addition, the Reserve Bank of Australia Payments System Board Annual Report (2013) shows that from 2004 to 2013, cash usage measured by the value of cash withdrawals grew much more slowly than household consumption.

The main purpose of this paper is to construct a model to reconcile these puzzling observa-

[1] There are in general three approaches to estimate the value of cash transactions at points of sale: using ATM withdrawals as a proxy, gathering information from consumer payments surveys and using retailer scanner data. Compared with the other two methods, ATM withdrawals are a less-direct measurement of cash usage, but it is relatively easy to construct time series. Arango et al. (2012) approximate the value of cash transactions in the retail sector by the value of ATM withdrawals.

tions. The model has two main features. The first is that substitutability between cash and other means of payment is uneven across different economic activities. While alternative means of payment directly compete with cash in many point-of-sale transactions, they are less-ideal substitutes for cash in the underground economy, bars, casinos, or in activities where agents desire anonymity or where unbanked or underbanked agents are involved. We capture this asymmetry by modelling an economy with two sectors: a cash-credit sector, where both cash and credit cards can be used as means of payment, and a cash-only sector, where credit is not accepted.[2] The cash-credit sector resembles point-of-sale transactions where cash is subject to intensive competition from other means of payment; the cash-only sector captures cash-intensive activities, where it is hard for other means of payment to compete with cash.

The second main feature of our framework is to emphasize that, during each cycle that cash circulates in the economy, it is likely to be used for a sequence of transactions, where cash and other means of payment may have different extents of substitutability. In particular, in our model, some cash is used first in the cash-credit sector and then again in the cash-only sector. In other words, some agents receive cash revenues in the cash-credit sector to spend in the ensuing cash-only sector: imagine taxi drivers acquiring cash from passengers to dine in cash-only restaurants, bakeries receiving cash from customers to purchase ingredients at local farmers' markets, farmers selling produce in cash to pay unbanked temporary workers, firms using cash revenues to pay suppliers, etc.[3]

Our model predicts that, once credit expands beyond a certain level, agents adjust their cash management practices in response to further credit expansions, causing the velocity of cash to slow down, so that the demand for cash can remain flat despite diminishing cash transactions. The intuition is as follows. As credit expands, agents who are buyers in the cash-credit sector hold less cash (to spend in that sector). At the same time, sellers in the cash-credit sector, who rely on cash revenues in that sector to finance their purchases in the cash-only sector, have to acquire more cash in advance to make up for the shortfall of cash receipts in the cash-credit sector. Compared with cash acquired by buyers, which is used in both sectors, cash acquired by sellers has a lower velocity because it is only used in the cash-only sector. Following a credit expansion, cash demand by agents who are sellers in the cash-credit market, which has a lower velocity, constitutes a higher fraction of the total demand for cash. As a result, the overall

[2] In the real world, many people use credit cards not for the borrowing function but rather for the payment function. Our model of credit is consistent with this fact since agents do not revolve their debts in the model.

[3] According to a study by the Canadian Federation of Independent Business (2011), 13% of small and medium enterprises engage in cash transactions with their suppliers. On average, 2% of payments to suppliers by small and medium enterprises are in cash; the percentage is higher at 6% for hospitality businesses.

velocity of cash decreases, and the total demand for cash stays constant even if the value of the cash transactions falls.

More generally, our study contributes to the monetary theory literature by showing the importance of modelling how cash circulates between economic activities featuring non-uniform substitutability between cash and alternative means of payment. Compared with standard monetary models, such as the overlapping-generations model of Samuelson (1958), the cash-in-advance model in Lucas and Stokey (1987), and the monetary search model in Lagos and Wright (2005), our model has very different implications about how credit expansion affects cash velocity, allocation and money demand.

Standard monetary models predict that, in the absence of precautionary demand, the velocity of cash is fixed in each trading cycle, and, therefore, is not affected by credit expansions. In our model, cash functions as a medium of exchange in two subsequent markets. Some part of cash is first used in the cash-credit sector and then reused to finance spending in the cash-only sector; this part of cash has a higher velocity. Our model predicts that credit expansions (beyond a certain level) reduce the amount of cash circulating in both sectors, or the part of cash that involves a higher velocity. Consequently, the overall velocity of cash decreases in response to credit expansions. As for the demand for cash, in a standard model where credit competes with cash in a single market subject to a credit limit (see Gu et al., 2013), credit expansions through relaxed credit limits reduce the demand for cash monotonically, until cash is completely driven out of circulation, a result incompatible with the recent trend in cash demand. In contrast, our model is able to capture the puzzling observation that total cash demand remains flat despite diminishing cash transactions. Regarding allocation, the standard Lagos and Wright (2005) model predicts that, in a monetary economy, the allocation is determined by the inflation rate and is invariant to the credit level. Our model generates a different prediction, suggesting that credit expansions may affect the allocation in a monetary economy by expanding (reducing) trading in the cash-credit (cash-only) sector.

The rest of the paper is organized as follows. Section 2 describes the environment. Section 3 characterizes the equilibrium allocation and examines the effect of credit expansions. Section 4 concludes. In the appendix, we discuss an alternative way to model credit expansions.

## 2 Economic Environment

Time is discrete and runs from $0$ to $\infty$. Each period $t$ consists of three sequential competitive markets. The first is a settlement market where agents settle debt obligations and adjust money holdings. Following the settlement market, there is a cash-credit market where both cash and credit can be used for transactions. The third market is a cash-only market where only cash is accepted as the means of payment.[4] As discussed in the introduction, the structure reflects the different extent of substitutability between cash and credit across different economic activities. Credit competes directly with cash in the cash-credit market, but not in the cash-only market.

There are two types of infinitely-lived agents, each of measure two. Type I agents are active in all three submarkets. Type II agents are active only in the settlement market and the cash-only market, and do not participate in the cash-credit market. There is a single perishable good in each market: good $x$ in the settlement market, good $q$ in the cash-credit market and good $y$ in the cash-only market.

In the settlement market, all agents can consume or produce good $x$. The utility from consuming $x$ units of the settlement market good is $x$. If $x < 0$, it means that the agent produces and incurs disutility. At the beginning of the settlement market, type I agents experience a preference shock that determines their preferences in the cash-credit market. With probability $1/2$, they become buyers, who would like to consume in the cash-credit market but cannot produce. With an equal probability, they become sellers in the cash-credit market, who can produce but do not want to consume. The preference shock is $i.i.d.$ across agents and time. For a buyer, the utility from consuming $q$ units of the good is $v(q)$, with $v' > 0$, $v'' < 0$ and $v'(0) = \infty$. The coefficient of relative risk aversion, $\sigma_q = -qv''(q)/v'(q)$, is less than $1$ everywhere. For a seller, the cost of producing $q$ units of goods is $q$. There exists $q^*$ such that $v'(q^*) = 1$. In the cash-only market, type II agents are endowed with a productive technology and can produce $y$ units of good at a cost of $y$, but do not want to consume. Type I agents cannot produce but would like to consume; their utility from consumption is $u(y)$, with $u(0) = 0$, $u' > 0$, $u'' < 0$ and $u'(0) = \infty$. The coefficient of relative risk aversion, $\sigma_y = -yu''(y)/u'(y)$, is less than $1$ everywhere. There exists $y^*$ such that $u'(y^*) = 1$. Assume that $y^* < q^*$, or the cash-only sec-

[4]The environment differs from the standard Lagos and Wright (2005) or Rocheteau and Wright (2005) model in that, in each trading cycle, there are two (instead of one) consecutive decentralized markets (in the first market, both cash and credit are used; in the second market, only cash is used) following the centralized market (which corresponds to the settlement market in our model). Berentsen et al. (2005) study a model with a centralized market followed by two decentralized markets. In their model, cash has to be used in both decentralized markets, and the focus of their study is the distribution of money balances and the non-neutrality of money.

tor is small relative to the cash-credit sector. The discounting factor between two consecutive periods is $\beta \in (0, 1)$.

In the following, we use subscripts "1" and "2" to refer to type I agents and type II agents, respectively. The subscripts "$s$" and "$b$" denote type I sellers and buyers, respectively. Type I agents' lifetime utility is

$$\mathcal{U}_1 = \sum_{t=0}^{\infty} \frac{1}{2}\beta^t \{[x_{bt} + v(q_{bt}) + u(y_{bt})] + [x_{st} - q_{st} + u(y_{st})]\}.$$

Type II agents' lifetime utility is

$$\mathcal{U}_2 = \sum_{t=0}^{\infty} \beta^t (x_{2t} - y_{2t}).$$

The structure of the economy implies that there are potential gains from trade between type I buyers and type I sellers in the cash-credit market, and type I and type II agents in the cash-only market. There are two payment instruments: money and credit. We assume that buyers can pay for good $q$ in the cash-credit market in money or credit, and type II agents only accept money for payment in the cash-only market. The stock of money, $M_t$, grows at a constant gross rate, $\gamma \equiv M_{t+1}/M_t$. We assume that $\gamma > \beta$. The expansion (contraction) of money supply is implemented as lump-sum transfers to (taxes on) type I agents in the settlement market in the amount of $\tau = (\gamma - 1) M$. Credit is extended in the cash-credit market, and credit balances are settled in the next period's settlement market. There is a credit limit such that the debt level in the settlement market cannot exceed $\ell \geq 0$. Credit expansion is modelled as an increase in $\ell$.[5]

In each trading cycle, type I agents engage in the following activities. In the settlement market, they adjust their money holdings after learning about the realization of their preference shocks, and settle credit balances incurred in the previous cash-credit market. In the cash-credit market, they trade among themselves for good $q$ using either cash, credit or both. In the cash-only market, they trade with type II agents, using cash to purchase good $y$ from type II agents. Type II agents engage in the following activities. They sell good $y$ to type I agents in the cash-only market, and spend money earned in the cash-only market to purchase good $x$ in the settlement market. The timing of events is summarized in Figure 1.

[5]In the appendix, we discuss the alternative specification that credit expansion manifests as an increase in the proportion of buyers with access to credit cards. Our main results continue to hold under that specification.

# 3 Characterization

In this section we characterize the equilibrium allocation and discuss the effect of expanding credit usage. The value functions are denoted as $W$ in the settlement market, $V$ in the cash-credit market and $U$ in the cash-only market. The initial money holding is $m$ in the settlement market, $\hat{m}$ in the cash-credit market and $\tilde{m}$ in the cash-only market. To simplify notation, we omit the time subscript. In the case where variables in the current period and the next period are involved, we label variables in the next period with a subscript "+." We first analyze decisions in the settlement market, then the cash-only market and finally the cash-credit market.

## 3.1 Settlement Market

### 3.1.1 Type I Agents

In the settlement market, type I agents settle credit/debt balances incurred in the cash-credit market in the previous period, and adjust cash holdings to carry into the cash-credit market after the realization of their preference shocks. Normalize the price of the settlement market good, $x$, to 1, and use $\phi$ to denote the value of a dollar measured in units of $x$. The value function of a type I agent who enters the settlement market holding $m$ units of money and $a$ units of IOUs (or credit extended to others in the previous period's cash-credit market) with a preference realization of $j \in \{b, s\}$ is

$$W_j(m, a) = \max_{x_j, \hat{m}_j} x_j + V_j(\hat{m}_j)$$
$$\text{s.t. } x_j + \phi\hat{m}_j = \phi(m + \tau) + a,$$

where $V_j(\hat{m}_j)$ denotes the value function of carrying $\hat{m}_j$ dollars into the cash-credit market, $x_j$ is consumption in the settlement market, and $\tau$ is the lump-sum transfer of money from the government. Note that type I agents may start the settlement market with different portfolios depending on their preferences in the last period's cash-credit market, and may choose different $\hat{m}$ depending on the realization of their preferences in the current period's cash-credit market. The value function can be rewritten as

$$W_j(m, a) = \max_{\hat{m}_j} \phi(m + \tau - \hat{m}_j) + a + V_j(\hat{m}_j). \tag{1}$$

The first-order condition is

$$\phi \geq \frac{dV_j\left(\hat{m}_j\right)}{d\hat{m}_j}, \text{ with equality if } \hat{m}_j > 0. \tag{2}$$

For type I agents, the marginal value of money in the settlement market is

$$\frac{dW_j\left(m, a\right)}{dm} = \phi,$$

and the marginal value of IOUs is

$$\frac{dW_j\left(m, a\right)}{da} = 1.$$

Note that owing to linear preferences in the settlement market, the choice of money holding $\hat{m}$ depends only on the realization of the preference shock in the current period, but does not depend on the portfolio entering into the current settlement market, $(m, a)$, or preference realizations in previous periods. Note also that the marginal values of $m$ and $a$ are the same for all agents.

### 3.1.2 Type II Agents

The value function for a type II agent in the settlement market is

$$\begin{aligned} W_2(m) &= \max_{\hat{m}_2} x_2 + U_2(\hat{m}_2) \\ \text{s.t. } x_2 &= \phi(m - \hat{m}_2). \end{aligned}$$

Remember that type II agents do not enter the cash-credit market and go from a settlement market directly to a cash-only market. It can be shown that in the settlement market, type II agents spend all money earned in last period's cash-only market and take no money to the cash-only market if $\gamma > \beta$, which implies a positive nominal interest rate. Type II agents are sellers in the cash-only market and do not need to use money in that market, so they choose not to take money to that market if the nominal interest rate is positive. Following the above argument, we can write the value function for a type II agent in the settlement market as

$$W_2\left(m\right) = \phi m + U_2(0), \tag{3}$$

where $U_2$ is the value for a type II agent in the cash-only market. The marginal value of money for a type II agent in the settlement market is

$$\frac{dW_2(m)}{dm} = \phi,$$

which is the same as for a type I agent.

## 3.2 Cash-Only Market

### 3.2.1 Type II Agents

We first analyze type II agents' problem, which is simpler. In the cash-only market, type II agents sell good $y$ to type I agents for cash. Let $\psi$ be the value of money in terms of $y$ in the cash-only market. The value function for type II agents is

$$\begin{aligned} U_2(0) &= \max_{y_2, m_+} -y_2 + \beta W_2(m_+) \\ \text{s.t. } y_2 &= \psi m_+, \end{aligned}$$

where $y_2$ is the production of good $y$ by type II agents, and $m_+$ is the money balance that type II agents carry into the next period's settlement market. The first-order condition is

$$\psi = \beta\phi_+. \tag{4}$$

While deciding how much money to earn on the cash-only market, type II agents weigh the marginal benefit and marginal cost of earning one additional unit of cash. The marginal cost to produce one unit of good $y$ entails a cost of $1$. The marginal benefit is to earn $1/\psi$ units of money and consume $\phi_+/\psi$ units of good $x$ in the settlement market in the next period, which generates $\beta\phi_+/\psi$ units of utility. Type II agents will choose $(y_2, m_+)$ by equating the marginal cost and marginal benefit, which implies equation (4).

### 3.2.2 Type I Agents

Type I agents decide how much good $y$ to consume and how much money to carry into the next period's settlement market (credit balances can only be settled in the settlement market). Their value function is

$$\begin{aligned} U_j\left(\tilde{m}_j, a_+^j\right) &= \max_{y_j, m_+} u(y_j) + \beta[W_b(m_+, a_+^j) + W_s(m_+, a_+^j)]/2 \\ &= \max_{y_j, m_+} u(y_j) + \beta\{\phi_+ m_+ + a_+^j + [W_b(0,0) + W_s(0,0)]/2\} \\ & \quad \text{s.t.} \begin{cases} y_j = \psi(\tilde{m}_j - m_+), \\ m_+ \geq 0, \end{cases} \end{aligned}$$

where $j \in \{b, s\}$ is the realization of the preference shock in the current period's settlement market. If the cash constraint binds, then

$$y_j = \psi\tilde{m}_j = \beta\phi_+\tilde{m}_j < y^* \text{ and } u'(y_s) > 1.$$

If the cash constraint is loose, then

$$u'(y_j) = 1 \text{ or } y_j = y^*.$$

The cash constraint binds if and only if $\tilde{m}_j < y^*/(\beta\phi_+)$. The value function for type I agents can be rewritten as

$$U_j(\tilde{m}_j, a_+^j) = \begin{cases} u(y^*) - y^* & \\ +\beta\{\phi_+ m_+^j + a_+^j + [W_b(0,0) + W_s(0,0)]/2\}, & \text{if } \tilde{m}_j \geq y^*/(\beta\phi_+), \\ u(\beta\phi_+\tilde{m}_j) & \\ +\beta\{a_+^j + [W_b(0,0) + W_s(0,0)]/2\}, & \text{if } \tilde{m}_j \leq y^*/(\beta\phi_+). \end{cases} \tag{5}$$

The marginal value of money is

$$\frac{\partial U_j(\tilde{m}_j, a_+^j)}{\partial \tilde{m}_j} = \beta\phi_+ u'(y_j) \begin{cases} = \beta\phi_+, \text{ if } \tilde{m}_j \geq y^*/(\beta\phi_+), \\ > \beta\phi_+, \text{ if } \tilde{m}_j < y^*/(\beta\phi_+). \end{cases} \tag{6}$$

The marginal value of credit extended to others is

$$\frac{\partial U_j(\tilde{m}_j, a_+^j)}{\partial a_+^j} = \beta. \tag{7}$$

## 3.3 Cash-Credit Market

In the cash-credit market, type I buyers and sellers trade good $q$. Let $\lambda$ be the value of settlement market goods in the next period, $x_+$, in terms of cash-credit market goods in the current period, $q$. Denote $\omega$ as the value of money in terms of $q$. Taking prices as given, type I buyers and sellers decide how much good $q$ (and how much money) to consume and produce (spend and earn), respectively.

### 3.3.1 Sellers

The seller's value function is

$$V_s(\hat{m}_s) = \max_{q_s, d_s, a_+^s} -q_s + U_s\left(\hat{m}_s + d_s, a_+^s\right) \tag{8}$$

$$\text{s.t. } q_s = \omega d_s + \lambda a_+^s, \tag{9}$$

where $d_s$ is the amount of cash that the seller receives, and $a_+^s$ is the amount of credit extended to others that will be repaid in the next period's settlement market. Let $\tilde{m}_s = \hat{m}_s + d_s$ be the amount of money carried to the cash-only market. The first-order conditions satisfy

$$d_s \; : \; \omega = \frac{\partial U_s\left(\tilde{m}_s, a_+^s\right)}{\partial \tilde{m}_s} = \beta\phi_+ u'(y_s), \tag{10}$$

$$a_+^s \; : \; \lambda = \frac{\partial U_s\left(\tilde{m}_s, a_+^s\right)}{\partial a_+^s} = \beta, \tag{11}$$

where the second equalities in equations (10) and (11) use equations (6) and (7), respectively. The value function $V_s(\hat{m}_s)$ has the following envelope result:

$$\frac{dV_s(\hat{m}_s)}{d\hat{m}_s} = \omega = \frac{\partial U_s\left(\tilde{m}_s, a_+^s\right)}{\partial \tilde{m}_s} = \beta\phi_+ u'(y_s). \tag{12}$$

### 3.3.2 Buyers

Let $d_b$ denote the amount of cash that the buyer spends in the cash-credit market, $\tilde{m}_b = \hat{m}_b - d_b$ the remaining cash that the buyer carries to the cash-only market and $-a_+^b$ the amount of debt incurred by the buyer. We can express the buyer's value function in the cash-credit market as

$$V_b(\hat{m}_b) = \max_{q_b, d_b, a_+^b} v(q_b) + U_b\left(\hat{m}_b - d_b, a_+^b\right) \tag{13}$$

$$\text{s.t.} \begin{cases} q_b = \omega d_b - \lambda a_+^b, \\ 0 \leq d_b \leq \hat{m}_b, \\ -a_+^b \leq \ell, \end{cases} \tag{14}$$

where the three constraints are the budget constraint, the cash constraint and the credit constraint, respectively.

The first-order condition with respect to $d_b$ is

$$v'(q_b)\omega = \frac{\partial U_b\left(\tilde{m}_b, a_+^b\right)}{\partial \tilde{m}_b} = \beta\phi_+ u'(y_b). \tag{15}$$

Note that the solution to $d_b$ is interior under the assumptions $v'(0) = \infty$ and $u'(0) = \infty$.

If the credit constraint is tight, then we have $a_+^b = -\ell$ and

$$q_b = \omega d_b + \lambda\ell = \omega d_b + \beta\ell.$$

If the credit constraint is loose, then $v'(q_b)\lambda = v'(q_b)\beta = \partial U_b\left(\tilde{m}_b, a_+^b\right)/\partial a_+^b = \beta$ or $q_b = q^*$.

The value function for the buyer, $V_b(\hat{m}_b)$, has the following envelope result:

$$\frac{dV_b(\hat{m}_b)}{d\tilde{m}_b} = v'(q_b)\omega = \frac{\partial U_b\left(\tilde{m}_b, a_+^b\right)}{\partial \tilde{m}_b} = \beta\phi_+ u'(y_b). \tag{16}$$

The market clearing conditions are

$$\begin{aligned} q_s &= q_b = q, \\ d_s &= d_b = d, \\ a_s &= -a_b. \end{aligned} \tag{17}$$

## 3.4 Equilibrium

We focus on the steady-state equilibrium where real variables are constant over time. The (gross) equilibrium inflation rate is equal to the money supply growth rate, $\gamma$. The nominal interest rate is $\gamma/\beta - 1 > 0$ under the assumption $\gamma > \beta$. We measure the demand for cash as the real cash balances in terms of the settlement market good $x$ that type I agents acquire at the end of the settlement market (note that type II agents are sellers in the cash-only market and choose not to carry cash at the end of the settlement market if the nominal interest rate is positive). Use $z_b = \phi \hat{m}_b$ and $z_s = \phi \hat{m}_s$ to denote the cash demand by buyers and sellers, respectively. The total money demand is calculated as $z = z_b + z_s$. Let $z_V = \phi d$ denote the value of cash transactions (measured in terms of good $x$) in the cash-credit market. We define the velocity of cash, $\upsilon$, as the ratio of the value of cash transactions in the cash-credit market and the cash-only market over the total demand for cash (all measured in terms of $x$), i.e.,

$$\upsilon = \frac{z_V + \phi \frac{y_b + y_s}{\psi}}{z} = \frac{z_V + \frac{\gamma}{\beta}(y_b + y_s)}{z}, \tag{18}$$

where the second equality uses (4). The value of GDP (measured in units of $x$) is

$$Y = \phi \left( \frac{q}{\omega} + \frac{y_b + y_s}{\psi} \right).$$

We define the ratio of the value of notes in circulation over GDP as

$$\rho = \frac{z_b + z_s}{\phi \left( \frac{q}{\omega} + \frac{y_b + y_s}{\psi} \right)}.$$

The demand for cash at the end of the settlement market is characterized by

$$\text{buyers:} \qquad \phi = \omega v'(q), \tag{19}$$

$$\text{sellers:} \qquad \phi \geq \omega, \text{ with equality if and only if } \hat{m}_s > 0, \tag{20}$$

which are derived by plugging (12) and (16) into (2). Type I buyers' demand for cash is always positive: they must acquire cash in the settlement market to be spent on the cash-credit market and the cash-only market. Type I sellers' cash demand, however, may be either zero or positive. Sellers need to use cash in the cash-only market, and they have two opportunities to acquire it, in the settlement market at a (utility) cost of $\phi$, or in the cash-credit market at a (utility) cost of $\omega$ (remember that $\phi$ is the value of money in terms of the settlement market good $x$, and that $\omega$ is the value of money in terms of the cash-credit market good $q$).

The decision to carry money from the cash-credit market to the cash-only market can be derived from (6) and the Euler equations (10) and (15):

$$\text{buyers:} \qquad \omega v'(q) = \beta\phi_+ u'(y_b), \tag{21}$$

$$\text{sellers:} \qquad \omega = \beta\phi_+ u'(y_s). \tag{22}$$

Under the assumption that $u'(0) = \infty$, both buyers and sellers choose to enter the cash-only market with positive money balances.

In the cash-only market, type I buyers always choose to spend all their money: if $\gamma > \beta$, which implies a positive nominal interest rate, agents will never hold idle cash balances across two settlement markets. Depending on the amount of cash receipts in the cash-credit market, type I sellers may or may not spend all their cash receipts in the cash-only market: their cash constraints may be loose in the cash-only market if they receive large amounts of cash in the cash-credit market, in which case, $y_s = y^*$.

In general, the economy is likely to go through four regimes as $\ell$ increases and drives out cash transactions in the cash-credit market. When the credit limit is small, the buyer chooses a large money balance in the settlement market to prepare for trading in the cash-credit market and cash-only market. Given the positive nominal interest rate, the trading in the cash-credit market and the cash-only market lies below the efficient level, i.e., $q < q^*$ and $y_b < y^*$. For the seller, $q < q^*$ implies that $\phi > \omega$, or that it is cheaper to acquire cash in the cash-credit market than in the settlement market. As a result, sellers choose not to hold cash while entering the cash-only market and rely solely on cash receipts in the cash-credit market to finance purchases of $y$ in the cash-only market. When the credit limit is small, a large amount of trading is carried out in cash in the cash-credit market, which means that the seller may receive more than enough cash revenue in the cash-credit market to have a loose cash constraint in the cash only market, or $y_s = y^*$. We call this situation I.

As the credit limit rises and relaxes the buyer's cash constraint in the cash-credit market, buyers are able to purchase more $q$ in the cash-credit market. At the same time, expecting an increasing fraction of transactions to be conducted through a credit arrangement, buyers spend less cash in the cash-credit market, causing the seller's cash revenue to decrease. At some point, the seller has a binding cash constraint in the cash-only market, and $y_s < y^*$. When this happens, the economy transitions into situation II.

If the credit limit continues to rise, then it is possible that the buyer is able to purchase $q^*$ in the cash-credit market (but still pays part of the consumption of $q$ in cash). Once $q = q^*$, the

| Situation | $q$ | $z_s$ | $y_s$ | Credit constraint | $z_V$ |
|---|---|---|---|---|---|
| I | $< q^*$ | $0$ | $= y^*$ | bind | $> 0$ |
| II | $< q^*$ | $0$ | $< y^*$ | bind | $> 0$ |
| III | $= q^*$ | $> 0$ | $< y^*$ | bind | $> 0$ |
| IV | $= q^*$ | $> 0$ | $< y^*$ | loose | $= 0$ |

Table 1: The four situations

economy enters into situation III. In this situation, $\phi = \omega$, and the seller takes a positive amount of cash from the settlement market to the cash-credit market to compensate for the lower cash revenue in the cash-credit market.

Finally, the credit limit may be so high that the credit constraint becomes loose in the cash-credit market. In this situation (situation IV), transactions in the cash-credit market are entirely conducted in credit (and the amount of trading is at the efficient level $q^*$), and cash is only used in the cash-only market.

To summarize, as $\ell$ gradually increases, the economy goes through four situations depending on whether $q$ is equal to or less than $q^*$, whether $y_s$ is equal to or less than $y^*$, and whether the credit constraint binds in the cash-credit market (see Table 1). Next, we characterize the equilibrium allocation and discuss the effect of credit expansions in each of the four situations.

### 3.4.1 Situation I: Low Credit Regime

In the first situation, buyers consume less than $q^*$ in the cash-credit market, which implies that the buyer's valuation of money is higher than the seller's in the cash-credit market. For the seller, it is cheaper to acquire cash in the cash-credit market than in the settlement market. The seller, therefore, chooses to hold zero money balances while entering the cash-credit market ($z_s = 0$), and finances spending on the cash-only market solely by cash receipts in the cash-credit market. With a very low credit limit, the seller receives enough cash revenue in the cash-credit market to support $y^*$ in the cash-only market. In this situation, the equilibrium allocation $(q, y_b, y_s, y_2)$ is given by

$$y_s = y^*, \tag{23}$$

$$q = q_\gamma, \tag{24}$$

$$y_b = y_\gamma, \tag{25}$$

$$y_2 = \frac{y_b + y_s}{2}, \tag{26}$$

where $y_\gamma$ solves

$$u'(y_\gamma) = \frac{\gamma}{\beta}, \tag{27}$$

and $q_\gamma$ solves

$$v'(q_\gamma) = \frac{\gamma}{\beta}. \tag{28}$$

Equations (24) and (25) follow from (19), (21), (22) and (23). The seller's choice in the cash-credit market can be used to derive the relative price of good $q$ and good $x$. Accumulating one more unit of money in the cash-credit market entails a utility cost of $\omega$. The marginal benefit is to enjoy $\psi$ units of good $y$ in the cash-only market, or $\psi u'(y^*) = \beta\phi_+$ in terms of utility gain using (4). In equilibrium, the seller equates the marginal cost and the marginal benefit of acquiring one additional unit of cash, which implies the relationship

$$\omega = \beta\phi_+.$$

As the buyer decides how much money to prepare in the settlement market, the cost of each additional unit of money is $\phi$. The benefit is to enjoy $\omega$ $(= \beta\phi_+)$ units of good $q$ in the cash-credit market, or $\psi$ $(= \beta\phi_+)$ units of good $y$ in the cash-only market, which brings $\beta\phi_+ v'(q)$ or $\beta\phi_+ u'(y_b)$ units of additional utility, respectively. Equating the marginal benefit and marginal cost gives equations (24) and (25). Equation (26) is the market clearing condition in the cash-only market.

In situation I, money demand and velocity are described by

$$\begin{aligned}
z_V &= \phi d = \phi \frac{q_\gamma - \lambda\ell}{\omega} = \phi\frac{q_\gamma - \beta\ell}{\beta\phi_+} = \frac{\gamma}{\beta}(q_\gamma - \beta\ell), && (29)\\
z_b &= z_V + \phi\frac{y_b}{\psi} = \frac{\gamma}{\beta}(q_\gamma - \beta\ell + y_\gamma), && (30)\\
z_s &= 0,\\
z &= z_b.
\end{aligned}$$

The velocity of cash is given by

$$\upsilon = \frac{z_V + \frac{\gamma}{\beta}(y_b + y_s)}{z} = \frac{z_b + \frac{\gamma}{\beta}y^*}{z_b} = 1 + \frac{\gamma}{\beta}\frac{y^*}{z_b}.$$

The amount of cash with a higher velocity (of two) is $(\gamma/\beta)y^*$. GDP (measured in units of $x$) is

$$Y = \phi\left(\frac{q}{\omega} + \frac{y_b + y_s}{\psi}\right) = \frac{\gamma}{\beta}(q_\gamma + y_\gamma + y^*).$$

The value of notes in circulation over GDP is

$$\rho = \frac{z_b + z_s}{Y} = \frac{q_\gamma - \beta\ell + y_\gamma}{q_\gamma + y^* + y_\gamma}.$$

In the cash-credit market, buyers purchase $q_\gamma$ units of good, $\lambda\ell$ of which is paid in credit, and the remaining amount, $q_\gamma - \lambda\ell$, is paid in cash. The value of cash transactions in the cash-credit market, measured in units of good $x$, is therefore $\phi(q_\gamma - \lambda\ell)/\omega$. Using $\omega = \beta\phi_+$, we can derive equation (29). Equation (30) says that buyers prepare cash to spend in the cash-credit market and the cash-only market: the value of cash spending is $z_V$ on the cash-credit market, and $\phi y_\gamma/\psi$ (or $(\gamma/\beta)y_\gamma$ by equation (4)) on the cash-only market. The seller does not take money to the cash-credit market. The total demand for cash is determined by the cash demand by buyers.

**Proposition 1** *Effect of $\ell$ in situation I: $dq/d\ell = dy_b/d\ell = dy_s/d\ell = dy_2/d\ell = 0$, $dz_V/d\ell = dz_b/d\ell = dz/d\ell = -\gamma < 0$, $dv/d\ell > 0$, and $d\rho/d\ell < 0$.*

In situation I, sellers do not take cash into the cash-credit market, but they receive enough cash revenue in the cash-credit market to have a loose cash constraint in the cash-only market. As a result, they purchase $y^*$ units of good $y$ in the cash-only market and take the remaining cash into the next period's settlement market. The buyer's cash-only consumption, $y_b$, is determined by the inflation rate and remains at $y_\gamma$. When the seller's cash constraint is loose in the cash-only market, the marginal benefit of earning extra money in the cash-credit market is determined by the inflation rate, which implies that the amount of trading in the cash-credit market, $q$, depends solely on the inflation rate. The amount of trading in the cash-only sector is fixed at $(y^* + y_\gamma)/2$.

As credit expands, as long as the seller's cash-constraint is still loose in the cash-only market, the economy remains in situation I, and the allocation profile remains the same. More credit usage does crowd out (the value of) cash transactions in the cash-credit market at a constant speed of $\gamma$, and the buyer cuts down on cash demand at the same rate. The total demand for cash decreases at the same speed. As for velocity, note that the part of cash used in both markets has a higher velocity of two, while the rest of cash has a lower velocity of one. When the economy is in situation I, the amount of high-velocity cash is fixed at $\phi y^*/\psi = (\gamma/\beta)y^*$ and accounts for an increasing fraction of the total money demand. However, the total money demand decreases in response to credit expansions as less cash (and more credit) is used to support the same output level. As a result, the value of notes in circulation over GDP decreases.

### 3.4.2 Situation II: Intermediate Credit Regime

In situation II, buyers consume less than $q^*$ in the cash-credit market, which implies that it is still cheaper for sellers to acquire cash in the cash-credit market than in the settlement market. The seller, therefore, chooses to hold zero money balances while entering the cash-credit market (i.e., $z_s = 0$), and finances spending on the cash-only market by cash receipts in the cash-credit market. This is the same as in situation I. However, in situation II, the seller does not receive enough cash revenue in the cash-credit market to purchase $y^*$ units of good in the cash-only market. In this situation, the seller's cash constraint in the cash-only market binds, and the seller spends the entire cash revenue earned in the cash-credit market on good $y$. The equilibrium allocation $(q, y_b, y_s, y_2)$ is given by

$$\begin{aligned} y_s &= \frac{q-\beta\ell}{\omega}\psi = \frac{q-\beta\ell}{u'(y_s)}, \qquad (31) \\ v'(q) &= \frac{u'(y_b)}{u'(y_s)}, \qquad (32) \\ y_b &= y_\gamma, \\ y_2 &= \frac{y_b+y_s}{2}. \end{aligned}$$

Equation (31) says that sellers exhaust cash revenues earned in the cash-credit market to purchase good $y$ on the cash-only market. The seller sells $q-\beta\ell$ units of good in cash and receives $(q-\beta\ell)/\omega$ units of cash revenue, which enables the seller to buy $y_s = [(q-\beta\ell)/\omega]\psi$ in the cash-only market. Equation (32) is derived from (21) and (22); it implies that the marginal rate of substitution between $q$ and $y$ is equalized for type I buyers and sellers. Using (22) and (4), we can derive the price of $q$ in terms of good $y$ as

$$\frac{\psi}{\omega} = \frac{\beta\phi_+}{\beta\phi_+ u'(y_s)} = \frac{1}{u'(y_s)}.$$

In deciding how to allocate cash spending in the cash-credit market and the cash-only market, the buyer equates the marginal benefit and the marginal cost of consuming one more unit of $q$. The benefit is $v'(q)$. The cost is to give up $1/u'(y_s)$ units of $y$, or $u'(y_b)/u'(y_s)$ in terms of utility. The buyer's cash-only consumption and the market clearing condition in the cash-only market remain the same as in situation I.

The money demand is described by

$$\begin{aligned} z_V &= \phi d = \phi \frac{q - \beta \ell}{\omega} = \phi \frac{y_s}{\psi} = \phi \frac{y_s}{\beta \phi_+} = \frac{\gamma}{\beta} y_s, && (33) \\ z_b &= z_V + \phi \frac{y_b}{\psi} = \frac{\gamma}{\beta}(y_\gamma + y_s), && (34) \\ z_s &= 0, \\ z &= z_b. \end{aligned}$$

In situation II, the seller uses the entire cash revenue earned in the cash-credit market to purchase good $y$ in the cash-only market, so the value of cash transactions in the cash-credit market is directly linked to the seller's consumption in the cash-only market through equation (33). Another implication is that cash acquired by the buyer is equal to the value of cash-only market transactions (equation (34)). The buyer acquires cash for spending on the cash-credit market and the cash-only market. The buyer's spending on the cash-credit market becomes the seller's cash revenue, all of which is then spent by the seller to purchase good $y$ in the cash-only market. As long as $q < q^*$ or $\phi > \omega$, the seller does not acquire cash on the settlement market, and the total demand for cash is the same as the buyer's demand for cash.

The velocity of cash is

$$\upsilon = \frac{z_V + \frac{\gamma}{\beta}(y_b + y_s)}{z} = 1 + \frac{y_s}{y_\gamma + y_s} = 1 + \frac{1}{y_\gamma / y_s + 1}.$$

The amount of cash with a velocity of two is the amount of cash used to finance the seller's spending on the cash-only market, which accounts for $y_s/(y_\gamma + y_s)$ of the total cash demand.

The value of GDP measured in units of $x$ is

$$Y = \phi \left( \frac{q}{\omega} + \frac{y_b + y_s}{\psi} \right) = \frac{\gamma}{\beta} \left[ \frac{q}{u'(y_s)} + y_\gamma + y_s \right].$$

The value of notes in circulation over GDP is

$$\rho = \frac{z_b + z_s}{Y} = \frac{y_\gamma + y_s}{\frac{q}{u'(y_s)} + y_\gamma + y_s} = \frac{1}{1 + \frac{q}{u'(y_s)(y_s + y_\gamma)}},$$

where the term $q/\left[u'(y_s)(y_s + y_\gamma)\right]$ is the value of output in the cash-credit market relative to the value of output in the cash-only market. When the seller relies on cash revenue in the cash-credit sector to purchase good $y$ in the cash-only sector and has a binding cash constraint in the cash-only sector, the value of notes in circulation is equal to the value of output in the cash-only sector. The same amount of cash is used to support trading in the cash-credit sector (together with credit transactions). The value of notes in circulation over GDP depends on the output ratio in the two sectors.

The effect of credit expansions is stated in Proposition 2.

**Proposition 2** *Effect of $\ell$ in situation II: $dq/d\ell > 0$, $dy_b/d\ell = 0$, $dy_s/d\ell < 0$, $dy_2/d\ell < 0$, $dz_V/d\ell = dz_b/d\ell = dz/d\ell < 0$, $dv/d\ell < 0$, and $d\rho/d\ell < 0$.*

**Proof.** First, $dy_b/d\ell = dy_\gamma/d\ell = 0$ is obvious. From (25), (32) and (27), we have

$$v'(q)u'(y_s) = \frac{\gamma}{\beta},$$

which implies that

$$v''(q)\frac{dq}{dy_s}u'(y_s) + v'(q)u''(y_s) = 0,$$

or

$$\frac{dq}{dy_s} = -\frac{v'(q)u''(y_s)}{v''(q)u'(y_s)} < 0.$$

Using (22), (4) and (33), we can derive

$$\begin{aligned} y_s &= \frac{q - \beta\ell}{u'(y_s)}, \\ \ell &= [q - y_s u'(y_s)]/\beta, \\ \frac{d\ell}{dq} &= \frac{1}{\beta}\left\{1 - \frac{d[y_s u'(y_s)]}{dy_s}\frac{dy_s}{dq}\right\}. \end{aligned}$$

Following the assumption $-yu''(y)/u'(y) < 1$, we have

$$\frac{d[yu'(y)]}{dy} = u'(y) + yu''(y) > 0.$$

Since $d[y_s u'(y_s)]/dy_s > 0$ and $dq/dy_s < 0$, $d\ell/dq > 0$, or

$$\frac{dq}{d\ell} > 0,$$

it then follows that

$$\begin{aligned}
\frac{dy_s}{d\ell} &= \frac{dy_s}{dq}\frac{dq}{d\ell} < 0, \\
\frac{dy_2}{d\ell} &= -\frac{1}{2}\frac{dy_s}{d\ell} > 0, \\
\frac{dz_V}{d\ell} &= \frac{dz}{d\ell} = \frac{\gamma}{\beta}\frac{dy_s}{d\ell} < 0.
\end{aligned}$$

For the velocity of cash, we have

$$\frac{dv}{d\ell} \propto \frac{dy_s}{d\ell} < 0,$$

where "$\propto$" means "has the same sign with."

Finally, we show that $d\rho/d\ell < 0$. Rewrite $\rho$ as

$$\rho = \frac{1}{1 + \frac{q}{u'(y_s)(y_s+y_\gamma)}}.$$

$$\begin{aligned}
\frac{d\rho}{d\ell} &\propto \frac{d\left[u'\left(y_s\right)\left(y_s+y_\gamma\right)/q\right]}{d\ell} \\
&= \left\{\frac{d[u'(y_s)y_s/q]}{dy_s} + y_\gamma\frac{d[u'(y_s)/q]}{dy_s}\right\}\frac{dy_s}{d\ell} \\
&= \left\{\frac{d[u'(y_s)y_s/q]}{dy_s} + \frac{y_\gamma}{q^2}\left[u''(y_s)q - u'(y_s)\frac{dq}{dy_s}\right]\right\}\frac{dy_s}{d\ell} \\
&= \left\{\frac{d[u'(y_s)y_s/q]}{dy_s} + \frac{y_\gamma}{q^2}\left[u''(y_s)q - u'(y_s)(-)\frac{v'(q)u''(y_s)}{v''(q)u'(y_s)}\right]\right\}\frac{dy_s}{d\ell} \\
&= \left\{\underbrace{\frac{d[u'(y_s)y_s/q]}{dy_s}}_{+} + \underbrace{\frac{y_\gamma}{q^2}\frac{u''(y_s)}{v''(q)}\left[qv''(q)+v'(q)\right]}_{+}\right\}\underbrace{\frac{dy_s}{d\ell}}_{-} \\
&< 0.
\end{aligned}$$

■

The buyer's consumption on the cash-only market is determined by the inflation rate and does not respond to the credit limit. An increase in the credit limit relaxes the buyer's cash constraint in the cash-credit market and allows the buyer to consume more in the cash-credit market ($q$ increases). As buyers use more credit in the cash-credit market, they spend less cash in the market ($z_V$ decreases) and acquire less cash in the settlement market ($z_b$ decreases). For sellers, their cash revenue shrinks, but as long as $q < q^*$, it is cheaper for them to earn cash in the cash-credit market, and they choose not to acquire money in the settlement market. As a result, the reduction in cash revenue results in less consumption in the cash-only market ($y_s$ decreases). The total demand for cash also decreases because buyers carry less cash and sellers have zero demand for cash. As for the velocity of cash, the fraction of cash with a higher velocity (of two) shrinks as $y_s$ decreases (remember that the amount of cash to finance $y_s$ has a velocity of two); the velocity of cash, therefore, decreases as credit expands. A credit expansion increases output in the cash-credit sector ($q$ increases) and reduces output in the cash-only sector ($y_s$ decreases). However, the rise in the value of output in the cash-credit sector dominates the fall in the value of output in the cash-only sector, leading to an increase in GDP.[6] Hence, the value of notes in circulation over GDP decreases with the credit limit.

#### 3.4.3 Situation III: High Credit Regime

As the credit limit continues to rise, buyers will be able to afford more consumption in the cash-credit market. Once $q$ reaches $q^*$, the economy switches to situation III. In situation III, with

$$q = q^*, \tag{35}$$

and

$$\phi = \omega,$$

the seller has the same cost of acquiring cash in the settlement market and in the cash-credit market. In this situation, sellers start to acquire some of their cash in the settlement market. The cash-only market allocation $(y_b, y_s, y_2)$ is given by

$$y_b = y_s = y_2 = y_\gamma. \tag{36}$$

[6]Although the relative price of $q$ in terms of $y$ decreases, the effect is dominated by the expansion of $q$, as shown in the proof of Proposition 2.

The demand for money is given by

$$
\begin{aligned}
z_V &= \phi \frac{q^* - \beta\ell}{\omega} = q^* - \beta\ell, \\
z_b &= z_V + \phi \frac{y_b}{\psi} = q^* - \beta\ell + \frac{\gamma}{\beta} y_\gamma, \\
z_s &= -z_V + \phi \frac{y_s}{\psi} = -(q^* - \beta\ell) + \frac{\gamma}{\beta} y_\gamma, \\
z &= z_b + z_s = 2\frac{\gamma}{\beta} y_\gamma.
\end{aligned} \tag{37}
$$

The velocity of cash is

$$
\upsilon = \frac{z_V + \frac{\gamma}{\beta}(y_b + y_s)}{z} = \frac{q^* - \beta\ell + 2\frac{\gamma}{\beta} y_\gamma}{2\frac{\gamma}{\beta} y_\gamma} = 1 + \frac{q^* - \beta\ell}{2\frac{\gamma}{\beta} y_\gamma}.
$$

The part of cash that has a higher velocity (of two) is the cash used in the cash transactions in the cash-credit sector in the amount of $q^* - \beta\ell$. The value of GDP measured in units of $x$ is

$$
Y = \phi \left( \frac{q}{\omega} + \frac{y_b + y_s}{\psi} \right) = \frac{\gamma}{\beta} \left[ \frac{q^*}{u'(y_\gamma)} + 2y_\gamma \right].
$$

The value of notes in circulation over GDP is

$$
\rho = \frac{z_b + z_s}{Y} = \frac{2\frac{\gamma}{\beta} y_\gamma}{\frac{\gamma}{\beta}\left[ \frac{q^*}{u'(y_\gamma)} + 2y_\gamma \right]} = \frac{2y_\gamma}{\frac{q^*}{u'(y_\gamma)} + 2y_\gamma}.
$$

The effect of credit expansion in situation III is described in Proposition 3.

**Proposition 3** *In situation III,* $dq/d\ell = dy_b/d\ell = dy_s/d\ell = dy_2/d\ell = dz/d\ell = 0$, $dz_V/d\ell = dz_b/d\ell = -\beta < 0$, $dz_s/d\ell = \beta > 0$, $d\upsilon/d\ell = -\beta^2/(2\gamma y_\gamma) < 0$, *and* $d\rho/d\ell = 0$.

Once the level of trading in the cash-credit market reaches the efficient level of $q^*$, an increase in the credit limit has no further impact on the consumption in the cash-credit market. Expanding $q$ beyond $q^*$ means that the cost of acquiring money in the settlement market would fall below the cost of acquiring money in the cash-credit market ($\phi < \omega$), so the seller would never choose to sell more than $q^*$ in the cash-credit market. With the level of cash-credit market transactions being fixed at $q^*$, a credit expansion crowds out cash transactions ($z_V$) one for one (see equation (37)). Consumption in the cash-only market is determined by the inflation rate ($y_b = y_s = y_\gamma$). As more transactions in the cash-credit market are conducted in credit, the buyer reduces cash demand at the speed of $\beta$ to maintain the same consumption profile. However, the fall in cash payments implies that sellers have to make up the shortfall by acquiring more cash at the same speed to sustain their spending in the cash-credit market. The increase in the seller's demand exactly offsets the decrease in the buyer's demand, so that the total demand for cash remains constant despite shrinking cash usage. The overall velocity of cash decreases as the fraction of cash with a higher velocity (of two) decreases with less cash being used in the cash-credit market. The value of notes in circulation over GDP is flat irrespective of credit expansions, which is consistent with the recent trend in cash demand that we documented in the introduction.

### 3.4.4 Situation IV: Non-binding Credit Constraint Regime

If $\ell \geq \ell^* = q^*/\beta$, then the credit constraint is loose in the cash-credit market. In this situation, all cash-credit market transactions are conducted through a credit arrangement, and cash is completely driven out in the cash-credit market. The sole purpose of carrying cash is to finance cash-only market transactions. The allocation $(q, y_b, y_s, y_2)$ is fixed and described by

$$\begin{aligned} q &= q^*, \\ y_b &= y_s = y_2 = y_\gamma. \end{aligned}$$

The demand for cash is fixed and given by

$$\begin{aligned} z_V &= 0, \\ z_b &= z_s = \frac{\gamma}{\beta} y_\gamma, \\ z &= z_b + z_s = 2\frac{\gamma}{\beta} y_\gamma. \end{aligned}$$

The velocity of cash is fixed at

$$v = \frac{z_V + \frac{\gamma}{\beta}(y_b + y_s)}{z} = 1.$$

The value of GDP measured in units of $x$ is fixed at

$$Y = \phi \left( \frac{q}{\omega} + \frac{y_b + y_s}{\psi} \right) = \frac{\gamma}{\beta} \left[ \frac{q^*}{u'(y_\gamma)} + 2y_\gamma \right].$$

The value of notes in circulation over GDP is fixed at

$$\rho = \frac{z_b + z_s}{Y} = \frac{2y_\gamma}{\frac{q^*}{u'(y_\gamma)} + 2y_\gamma}.$$

Once the credit constraint becomes loose, further credit expansions have no impact on the allocation, the demand for cash, the velocity of cash or the value of notes in circulation over GDP. In situation IV, credit expansions no longer have an effect on the economy (see Proposition 4).

**Proposition 4** *In situation IV,* $dq/d\ell = dy_b/d\ell = dy_s/d\ell = dy_2/d\ell = dz_b/d\ell = dz_s/d\ell = dz_V/d\ell = dz/d\ell = dv/d\ell = d\rho/d\ell = 0$.

We calculate the cut-off values of $\ell$ that separate the four situations. The cut-off value of $\ell$ that separates situations I and II (denoted by $\tilde{\ell}$) solves $y_s = y^*$ in situation II, or $q_\gamma - \beta\ell = y^*$, or

$$\tilde{\ell} = \frac{q_\gamma - y^*}{\beta}.$$

The cut-off value of $\ell$ that separates situations II and III (denoted by $\hat{\ell}$) solves $z_s = -(q^* - \ell\beta) + (\gamma/\beta)y_\gamma = 0$ in situation III , i.e.,

$$\hat{\ell} = \frac{q^* - \frac{\gamma}{\beta}y_\gamma}{\beta} = \frac{q^* - y_\gamma u'(y_\gamma)}{\beta}.$$

Under the assumption that $-yu''(y)/u'(y) < 1$, we have $d[yu'(y)]/dy > 0$, or $y^*u'(y^*) = y^* > y_\gamma u'(y_\gamma)$. Together with $q_\gamma < q^*$, we have $\tilde{\ell} < \hat{\ell}$. If $q_\gamma < y^*$, i.e., if the inflation rate is high enough, then situation I disappears. Under the assumption that $q^* > y^*$, we have $\hat{\ell} > 0$, so situation II always exists. The cut-off value of $\ell$ that separates situations III and IV is $\ell^* = q^*/\beta$.

To summarize, our model predicts that when the credit limit is small, i.e., when $\ell < \tilde{\ell}$, credit expansions do not affect the allocation in the economy. The credit expansion beyond $\tilde{\ell}$ may have real effects on the economy. As described in Proposition 2, when $\ell \in [\tilde{\ell}, \hat{\ell}]$, more widespread use of credit expands cash-credit activities ($q$ increases), shrinks cash-only activities ($y_s$ decreases), and crowds out cash transactions in the cash-credit sector and the total demand for cash. However, the trend stops once credit expands beyond $\hat{\ell}$. Once $\ell$ exceeds $\hat{\ell}$, the value of cash transactions in the cash-credit sector continues to decrease owing to credit expansion (until the credit constraint becomes loose, or $\ell \geq \ell^*$), but the real allocation and the total demand for cash become fixed. What happens in our model when $\ell \in [\hat{\ell}, \ell^*]$ is consistent with the recent trend in cash demand that we discussed in the introduction. We depict the effect of credit expansions in Figure 2.

## 3.5 Discussion

As mentioned in the introduction, the main reason why our model can successfully capture the flat cash demand despite diminishing cash usage is that we model the circulation of cash between economic activities featuring non-uniform substitutability between cash and alternative means of payment. Standard versions of popular monetary models, including the overlapping-generations model in Samuelson (1958), the cash-in-advance model in Lucas and Stokey (1987), and the monetary search model in Lagos and Wright (2005), have very different predictions about the effect of credit expansions on cash velocity, demand for cash, and allocation. In the following, we sketch the standard Lagos and Wright (2005) model with credit (refer to Gu et al., 2013 for a more detailed analysis) and compare its predictions with those of our model.

In essence, our model reduces to a standard Lagos and Wright (2005) model with credit if we drop the cash-only market and retain only the settlement market and the cash-credit market.[7] As a result, the model has only type I agents who participate in the settlement market and the cash-credit market, but there are no type II agents. The problems in the settlement market can be formulated in the same way as before (see (1) and (2)). A standard result is that sellers take zero cash balances to the cash-credit market.

Without the cash-only market, agents face different problems in the cash-credit market. The buyer solves the problem

$$\begin{aligned} V_b(\hat{m}_b) &= \max_{q_b, d_b, a^b_+} v\left(q_b\right) + \beta W\left(\hat{m}_b - d_b, a^b_+\right) \\ &\text{s.t.} \begin{cases} q_b = \omega d_b - \lambda a^b_+, \\ \hat{m}^+ \geq d_b \geq 0, \\ -a^b_+ \leq \ell. \end{cases} \end{aligned}$$

If $\gamma > \beta$ and $\ell$ is small (the condition on $\ell$ will be given later), then both the cash and credit

[7] The cash-credit market and the settlement market in our model correspond to the decentralized market and centralized market, respectively, in Lagos and Wright (2005).

constraints bind, and we have

$$\begin{aligned} q_b &= \omega \hat{m}_b + \lambda \ell, \\ \frac{dV_b(\hat{m}_b)}{d\hat{m}_b} &= v'(q_b)\,\omega. \end{aligned}$$

Sellers solve the problem

$$\begin{aligned} V_s(0) &= \max_{q_s, d_s, a^s_+} -q_s + \beta W\left(d_s, a^s_+\right) \\ \text{s.t. } q_s &= \omega d_s + \lambda a^s_+. \end{aligned}$$

The first-order conditions are

$$\begin{aligned} \omega &= \beta \phi_+, \\ \lambda &= \beta. \end{aligned}$$

Finally, the market clearing conditions satisfy equation (17).

Combining the analysis of decisions on the settlement market and cash-credit market, we can characterize the buyer's money demand in the settlement market by

$$\phi = \frac{dV_b(\hat{m}_b)}{d\hat{m}_b} = v'(q)\,\omega = v'(q)\,\beta\phi_+.$$

In a steady-state equilibrium, the amount of goods being traded in the cash-credit market is given by $q = q_\gamma$. The demand for cash is

$$\begin{aligned} z_V &= z_b = z = \phi \frac{q_\gamma - \beta\ell}{\omega} = \frac{\gamma}{\beta}(q_\gamma - \beta\ell), \\ z_s &= 0. \end{aligned}$$

The velocity of cash is fixed at $\upsilon = 1$. The value of GDP measured in units of $x$ is

$$Y = \phi \frac{q}{\omega} = \frac{\gamma}{\beta} q.$$

The value of notes in circulation over GDP is

$$\rho = \frac{z}{Y} = \frac{\frac{\gamma}{\beta}(q_\gamma - \beta\ell)}{\frac{\gamma}{\beta} q_\gamma} = 1 - \frac{\beta\ell}{q_\gamma}.$$

As $\ell$ increases, the economy experiences three regimes marked by two cut-off values of $\ell$: $\ell_\gamma = q_\gamma/\beta$ and $\ell^* = q^*/\beta$. If $\ell \leq \ell_\gamma$ (regime A), buyers hold positive amounts of money and use both money and credit to purchase $q$. Credit expansions do not affect the allocation in this regime: $q$ is fixed at $q_\gamma$. Credit crowds out cash usage and cash demand at a constant rate $\gamma$ until it becomes zero when $\ell = \ell_\gamma$. Once $\ell$ exceeds $\ell_\gamma$, the economy becomes a pure credit economy and money demand remains at zero. If $\ell_\gamma \leq \ell \leq \ell^*$ (regime B), $q$ increases with $\ell$ at rate $\beta$ until it reaches $q^*$ when $\ell = \ell^*$. Finally, if $\ell \geq \ell^*$ (regime C), the credit constraint is loose, and the allocation remains at the efficient level ($q = q^*$). Figure 3 illustrates how money demand, velocity and allocation responds to increasing credit usage in the Lagos and Wright model. The model has very different implications about the effect of credit expansions from our model, which we discuss in more detail below.

In terms of money demand and cash velocity, our model predicts that, once credit usage exceeds a threshold (so that the economy enters into situation III), further crowding out of cash transactions by credit usage induces changes in cash management practices, resulting in lower cash velocity, but has no effect on the demand for cash, a phenomenon observed in many industrialized countries. As credit replaces cash in the cash-credit sector, agents who are buyers in that sector cut down on their cash demand (to be spent in that sector). However, agents who are sellers in that sector make up for the shortfall by acquiring more cash in advance (to spend in the cash-only sector). Because the cash acquired by sellers in advance (which is only used in the cash-only sector) has a lower velocity relative to that acquired by buyers (which is spent in both sectors), the redistribution of money demand from buyers to sellers causes the velocity of cash to decrease. In contrast, in a standard Lagos and Wright model with credit, the velocity of cash is fixed (at one if using the definition in equation (18)). With a fixed velocity, the model predicts that money demand shrinks as credit crowds out cash transactions at a constant rate until the economy transforms into a cashless economy. As a result, the model cannot capture the recent trend in cash demand.

Regarding allocation, the Lagos-Wright model with credit predicts that credit has no effect on allocation in a monetary economy: $q$ is constant at $q_\gamma$ if $z > 0$ (which occurs in regime A). In contrast, our model predicts that credit may affect allocation in a monetary economy. If $\tilde{\ell} < \ell < \hat{\ell}$ (or when the economy is in situation II), the demand for money is positive, and a credit expansion enlarges cash-credit market output (and reduces the seller's cash-only market consumption).[8]

[8]Credit expansions have a similar effect on allocation in situation I in our model and in regime A in Lagos and Wright (2005): the allocation ($q$ and $y$ in our model, and $q$ in Lagos and Wright) depends only on the inflation rate

Before concluding, we make a few additional comments about our model's predictions. First, although our main focus is to explain the *recent* trend in cash demand, our model is also consistent with observations in earlier periods. As shown in Figures 4 and 5 in Fung et al. (2014), the total demand for cash measured as the value of notes in circulation as a percentage of GDP had been decreasing before the 1980s, which is consistent with our model's prediction for the situation with low levels of credit usage ($\ell < \hat{\ell}$). To a certain extent, the earlier observations validate our model.

Second, one may think that with cash playing a diminishing role in the cash-credit sector (cash-credit market), in order to sustain the same level of total demand for cash, activities in the cash-only sector (cash-only market) have to expand to offset the reduced role for cash in the cash-credit sector. Suppose we take a standard cash-good-credit-good model by Lucas and Stokey (1987) and impose credit limits in the cash-credit sector. The set-up resembles our model with one critical difference: cash does not circulate sequentially between the cash-only sector and the cash-credit sector, as in our model. As credit expands in the cash-credit sector, the demand for cash in that sector decreases. Unless the cash-only sector expands, the total demand for cash would decrease. Our model proposes a different mechanism, which does not rely on the expansion of the cash-only sector to support a flat total demand for cash when cash plays a diminishing role in the cash-credit sector: the results listed in Proposition 3 are derived with a fixed size of cash-only market (output is fixed at $y_\gamma$). In our model, cash received in the cash-credit sector re-enters and is reused to finance spending in the cash-only sector. As credit expansion crowds out cash usage in the cash-credit economy, agents who receive cash in the cash-credit sector will face a tighter cash constraint in the cash-only sector and respond by increasing their demand for cash (which has a lower velocity). This force counteracts the diminishing demand for cash to be spent in the cash-credit sector, so that the total demand for cash remains fixed. Given the wide difference in the estimate of the size of the cash-only sector and its trend, it is a desirable feature that our model's prediction does not hinge on the size of the cash-only sector.[9]

---

and is not affected by credit expansions. The reason is that, the seller's marginal benefit of acquiring money in the cash-credit market is determined by the inflation rate if the seller has a loose cash constraint (in the cash-credit market in Lagos and Wright, and in both the cash-credit market and the cash-only market in our model). Another similarity is that both models predict that credit expansions no longer affect the economy when the credit constraint becomes loose (situation IV in our model, and regime C in Lagos and Wright).

[9]For example, the estimate of the size of the underground economy, which tends to be cash sensitive, varies much across different studies. Terefe et al. (2011) estimate that the Canadian underground economy was 2.7% in 1992 and 2.2% in 2008. According to Schneider and Buehn (2012), the Canadian underground economy was 16.3% in 1999 and 15.4% in 2010. Dunbar and Fu (2012) study income under-reporting (which is easier if the income is received in cash) in Canada in 1999 and 2005, and find that the proportion of households under-reporting

Third, in our model we focus on the competition between cash and credit, but our story has relevance for other electronic payments, such as debit cards, general purpose pre-paid cards, mobile payments, virtual currencies, etc. A common feature of these increasingly popular payment instruments is that they are more convenient than cash in some transactions, but they are yet to become the perfect substitute for cash in all activities. Our paper captures these features in a simplistic manner: the substitutability is perfect in the cash-credit sector, and there is zero substitutability in the cash-only sector. We expect that the main forces to generate our results remain relevant in more general settings where the substitutability between cash and other payment instruments is modelled in a less-extreme manner. As long as the replacement of cash by these payment schemes varies across sectors in the economy and some agents use cash receipts from less-cash-intensive sectors to finance more-cash-intensive activities, the mechanism described in our paper would hold.

# 4 Conclusion

This paper builds a model to reconcile a puzzling recent trend in cash demand: the total demand for cash, measured by the value of notes in circulation as a percentage of GDP, remains more or less flat despite diminishing cash usage owing to intensive competition from other means of payment, such as credit cards. Our model emphasizes that the substitutability between cash and other means of payment is uneven across different economic activities, and that it is important to model how cash circulates across these activities. If some agents use cash receipts in less-cash-intensive sectors to finance spending in more-cash-intensive sectors, then the total demand for money may not decrease even if cash plays a diminishing role in transactions. Once the expansion of alternative means of payment exceeds a certain level, agents adjust their cash management practices causing the velocity of cash to fall, so that the demand for cash can remain flat despite diminishing cash transactions.

More generally, our study shows the importance of modelling in greater detail how cash circulates in the economy: compared with models with a more standard set-up, our model has very different implications about how credit affects allocation, cash demand and the velocity of cash. In view of this, we suggest a few directions for future research. First, our model suggests that the overall velocity of cash should decrease as cash plays a diminishing role in

was between 30% and 45% in both years, and under-reported income rose by about 70% between 1999 and 2005 (GDP growth was 16% in the same period).

transactions. To test this implication, we need to gather time-series data on the velocity of cash, measured, for example, by the lifespan of bank notes. Second, we need to deepen our understanding about the cash-intensive sector such as the underground economy. Who are the participants? How do their cash management practices evolve over time? How do buyers in the sector acquire cash? What are the connections with less-cash-intensive sectors? Third, if we interpret some of the sellers in the cash-credit sector as firms, then our model predicts that as credit expands beyond a certain level, firms' cash demand will increase. Existing studies focus more on the consumer side: for example, many central banks have conducted consumer payment choice surveys.[10] Useful (albeit imperfect) aggregate time-series data have also been constructed to monitor consumers' cash demand. However, studies on merchants' cash demand are very limited. Rigorous empirical studies in this area would be valuable.

[10]For example, see Mooslechner et al. (2006) for the Austrian survey, Bagnall et al. (2011) for the Australian survey, Arango and Welte (2012) for the Canadian survey, Deutsche Bundesbank (2013) for the German survey, Jonker et al. (2012) for the Dutch survey, and Foster et al. (2013) for the survey in the United States. Recently, Bagnall et al. (2014) conducted a cross-country comparison with diary survey data to study consumer cash usage and management.

# A Credit Expansion along the Extensive Margin

In this appendix, we discuss the model's implication if we model credit expansion along the extensive margin, i.e., as a higher fraction of buyers using credit cards (with no credit limit). At the beginning of the settlement market, type I agents experience a preference shock and become buyers or sellers with the same probability. At the same time, type I buyers experience a technological shock: $\delta$ fraction of them have access to credit cards, while the rest do not. In the following, we use subscript "$c$" to refer to buyers with credit cards, and "$m$" to refer to buyers who have no credit cards and use cash only.

The settlement market and cash-only market problems remain the same as in the main text. In the cash-credit market, the seller's problem remains the same, but we need to distinguish between buyers who have credit cards and those who do not. The value function for a buyer who has a credit card is

$$V_c\left(\hat{m}_c\right) = \max_{q_c, a} v\left(q_c\right) + U_c\left(\hat{m}_c, a_+\right)$$
$$\text{s.t. } q_c = -\beta a_+.$$

The first-order condition implies that $q_c = q^*$, i.e., buyers with credit cards buy the optimal amount of quantity in the cash-credit market. The marginal value of money to a credit card user is

$$\frac{dV_c\left(\hat{m}_c\right)}{d\tilde{m}_c} = \frac{dU_c\left(\hat{m}_c, a_+\right)}{d\tilde{m}_c} = \beta\phi_+ u'\left(y_c\right).$$

The value of a buyer who uses only money in the cash-credit market is

$$V_m\left(\hat{m}_m\right) = \max_{q_m,, d} v\left(q_m\right) + U_m\left(\hat{m}_m - d\right)$$
$$\text{s.t. } \begin{cases} q_m = \omega d, \\ \hat{m}_m \geq d \geq 0, \end{cases}$$

where the two constraints are the budget constraint and the cash constraint, respectively. The first-order condition with respect to $d$ is

$$v'\left(q_m\right)\omega = \frac{\partial U_m\left(\tilde{m}_m\right)}{\partial \tilde{m}_m} = \beta\phi_+ u'\left(y_m\right).$$

The value function $V_m(\hat{m})$ has the following envelope result:

$$\frac{dV_m(\hat{m}_m)}{d\tilde{m}_m} = \frac{\partial U_m(\tilde{m}_m)}{\partial \tilde{m}_m} = v'(q_m)\,\omega.$$

The market clearing conditions become

$$\begin{aligned} q_s &= \delta q_c + (1-\delta)\,q_m, \\ d_s &= (1-\delta)\,d, \\ a_s &= -\delta a. \end{aligned}$$

Collect the equations that characterize the equilibrium allocations and prices as follows:

| | | |
|---|---|---|
| settlement market type I buyers with credit | : | $\phi = \beta u'(y_c)\,\phi_+$, |
| settlement market type I buyers without credit | : | $\phi = \omega v'(q_m)$, |
| settlement market type I sellers | : | $\phi \geq \omega$, with equality if $\hat{m}_s > 0$, |
| cash-credit market buyers with credit | : | $q_c = q^*$, |
| cash-credit market buyers without credit | : | $\omega v'(q_m) = \beta u'(y_m)\,\phi_+$, |
| cash-credit market sellers | : | $\omega = \beta u'(y_s)\,\phi_+$. |

The equilibrium allocation $(q, y_c, y_m, y_s, y_2)$ is given by

$$v'(q_m)\,u'(y_s) = u'(y_m)\,, \tag{38}$$

$$q_c = q^*, \tag{39}$$

$$y_c = y_\gamma, \tag{40}$$

$$y_m = y_\gamma, \tag{41}$$

$$y_2 = \frac{y_\gamma + y_s}{2}, \tag{42}$$

$$(y_s - y^*)\left[y_s - \frac{(1-\delta)\,q_m}{u'(y_s)}\right](q_m - q^*) = 0. \tag{43}$$

Similar to the model discussed earlier, there are three situations depending on whether type I sellers acquire cash in the settlement market or not, and whether sellers consume $y^*$ in the cash-only market or not (situation IV corresponds to the case where $\delta = 1$).

The velocity of cash is

$$\upsilon = \frac{z_V + \phi \frac{(1-\delta)y_m + \delta y_c + y_s}{\psi}}{z}.$$

The value of GDP measured in units of $x$ is

$$Y = \phi \left[ \frac{(1-\delta)q_m + \delta q_c}{\omega} + \frac{(1-\delta)y_m + \delta y_c + y_s}{\psi} \right].$$

The value of notes in circulation over GDP is

$$\rho = \frac{z}{Y}.$$

## A.1 Situation I: Low Credit Regime

In situation I, sellers do not accumulate cash in the settlement market. They receive enough cash revenue in the cash-credit market to purchase $y^*$ units of consumption in the cash-only market, i.e.,

$$y_s = y^*.$$

Buyers who have no credit cards purchase $q_\gamma$ units of goods in the cash-credit market, or

$$q_m = q_\gamma.$$

In situation I, money demand is given by

$$\begin{aligned}
z_V &= \phi(1-\delta)\frac{q_m}{\omega} = \phi(1-\delta)\frac{q_\gamma}{\beta\phi_+} = (1-\delta)\frac{\gamma}{\beta}q_\gamma, \\
z_c &= \phi\frac{y_c}{\psi} = \phi\frac{y_c}{\beta\phi_+} = \frac{\gamma}{\beta}y_\gamma, \\
z_m &= \phi\frac{q_m}{\omega} + \frac{y_m}{\psi} = \frac{\gamma}{\beta}(q_\gamma + y_\gamma), \\
z_s &= 0, \\
z &= \delta z_c + (1-\delta)z_m = \frac{\gamma}{\beta}[(1-\delta)q_\gamma + y_\gamma].
\end{aligned}$$

The velocity of cash is

$$\upsilon = \frac{z_V + \phi\frac{(1-\delta)y_m + \delta y_c + y_s}{\psi}}{z} = 1 + \frac{y^*}{(1-\delta)q_\gamma + y_\gamma}.$$

The value of GDP measured in units of $x$ is

$$\begin{aligned} Y &= \phi\left[\frac{(1-\delta)q_m+\delta q_c}{\omega}+\frac{(1-\delta)y_m+\delta y_c+y_s}{\psi}\right] \\ &= \frac{\gamma}{\beta}\left[(1-\delta)q_\gamma+\delta q^*+y_\gamma+y^*\right]. \end{aligned}$$

The value of notes in circulation over GDP is

$$\rho=\frac{z}{Y}=\frac{(1-\delta)q_\gamma+y_\gamma}{(1-\delta)q_\gamma+y_\gamma+\delta q^*+y^*}.$$

**Proposition 5** *Effect of $\delta$ in situation I: $dq_m/d\delta=dy_m/d\delta=dy_c/d\delta=dy_s/d\delta=0$, $dz_V/d\delta=dz/d\delta<0$, $d\upsilon/d\delta>0$, and $d\rho/d\delta<0$.*

**Proof.**

$$\begin{aligned} \frac{dz_V}{d\delta} &= -\frac{\gamma}{\beta}q_\gamma<0. \\ \frac{d\upsilon}{d\delta} &= y^*\frac{q_\gamma}{[(1-\delta)q_\gamma+y_\gamma]^2}>0. \end{aligned}$$

Rewrite $\rho$ as

$$\rho=\frac{1}{1+\frac{\delta q^*+y^*}{(1-\delta)q_\gamma+y_\gamma}}.$$

Then,

$$\begin{aligned} \frac{d\rho}{d\delta} &\propto \frac{d\left[\frac{(1-\delta)q_\gamma+y_\gamma}{\delta q^*+y^*}\right]}{d\delta} \\ &\propto -q_\gamma(\delta q^*+y^*)-[(1-\delta)q_\gamma+y_\gamma]q^* \\ &< 0. \end{aligned}$$

■

## A.2 Situation II: Intermediate Credit Regime

In situation II, we still have $q_m<q^*$ (but $q_m>q_\gamma$), and sellers still do not hold cash. However, unlike in situation I, the seller does not have enough cash revenue on the cash-credit market to support consumption of $y^*$ on the cash-only market. In this case, the seller spends all the cash receipts from the cash-credit market to purchase $y$ on the cash-only market. As a result,

$$y_s = \psi \frac{(1-\delta)q_m}{\omega} = \frac{(1-\delta)q_m}{u'(y_s)}.$$

In situation II, money demand is given by

$$\begin{aligned}
z_V &= \phi \frac{y_s}{\psi} = \phi \frac{y_s}{\beta \phi_+} = \frac{\gamma}{\beta} y_s, \\
z_c &= \frac{\gamma}{\beta} y_\gamma \text{ (same as in situation I)} \\
z_m &= \phi \left( \frac{q_m}{\omega} + \frac{y_\gamma}{\psi} \right) = \frac{\gamma}{\beta} \left( \frac{y_s}{1-\delta} + y_\gamma \right), \\
z_s &= 0, \\
z &= \delta z_c + (1-\delta) z_m = \frac{\gamma}{\beta}(y_s + y_\gamma).
\end{aligned}$$

The velocity of cash is

$$\upsilon = \frac{z_V + \phi \frac{(1-\delta)y_m + \delta y_c + y_s}{\psi}}{z} = \frac{y_s + y_\gamma + y_s}{y_s + y_\gamma} = 1 + \frac{y_s}{y_s + y_\gamma}.$$

The value of GDP measured in units of $x$ is

$$\begin{aligned}
Y &= \phi \left[ \frac{(1-\delta)q_m + \delta q_c}{\omega} + \frac{(1-\delta)y_m + \delta y_c + y_s}{\psi} \right] \\
&= \frac{\gamma}{\beta} \left[ y_s + \frac{\delta q^*}{u'(y_s)} + y_\gamma + y_s \right].
\end{aligned}$$

The value of notes in circulation over GDP is

$$\rho = \frac{z}{Y} = \frac{y_s + y_\gamma}{y_s + \frac{\delta q^*}{u'(y_s)} + y_\gamma + y_s}.$$

In situation II, the effect of credit expansions is described in the following proposition.

**Proposition 6** *Effect of $\delta$ in situation II: $dq_m/d\delta > 0$, $dy_m/d\delta = dy_c/d\delta = 0$, $dy_s/d\delta < 0$, $dz_V/d\delta = dz/d\delta < 0$, and $d\upsilon/d\delta < 0$. Moreover, if $\sigma_q < y_\gamma u'(y_\gamma)/q^*$, then $d\rho/d\delta < 0$.*

**Proof.** $dy_m/d\delta = dy_c/d\delta = 0$ follows directly from (40) and (41). From $v'(q_m)\,u'(y_s) =$

$u'(y_m) = \gamma/\beta$, we can derive

$$\upsilon'' \frac{dq_m}{dy_s} = \frac{\gamma}{\beta} \frac{-u''}{(u')^2}$$

and $dq_m/dy_s < 0$ or $dy_s/dq_m < 0$. Equation $y_s = (1-\delta)q_m/u'(y_s)$ implies that $\delta = 1 - y_s u'(y_s)/q_m$ and

$$\frac{d\delta}{dq_m} = -\frac{\frac{d[y_s u'(y_s)]}{dy_s}\frac{dy_s}{dq_m} q_m - y_s u'(y_s)}{(q_m)^2} > 0,$$

or $dq_m/d\delta > 0$. It then follows that $dy_s/d\delta = [dy_s/dq_m][dq_m/d\delta] < 0$. For the money demand, $dz/d\delta = dz_V/d\delta = \gamma dy_s/d\delta < 0$.

Rewrite $\upsilon$ as

$$\upsilon = 1 + \frac{1}{1 + y_\gamma/y_s}.$$

Then,

$$\frac{d\upsilon}{d\delta} \propto \frac{dy_s}{d\delta} < 0.$$

Rewrite $\rho$ as

$$\rho = \frac{1}{1 + \frac{y_s + \frac{\delta q^*}{u'(y_s)}}{y_s + y_\gamma}}.$$

Denote $\rho_Y = (y_s + y_\gamma) / \left(y_s + \frac{\delta q^*}{u'(y_s)}\right)$, then $d\rho/d\delta \propto d\rho_Y/d\delta$. Because

$$\frac{d\rho_Y}{d\delta} = \frac{d\rho_Y}{dy_s}\underbrace{\frac{dy_s}{d\delta}}_{<0} + \underbrace{\frac{(y_s + y_\gamma)\left(-\frac{q^*}{u'(y_s)}\right)}{\left(y_s + \frac{\delta q^*}{u'(y_s)}\right)^2}}_{<0},$$

we have $d\rho_Y/d\delta < 0$ if $d\rho_Y/dy_s > 0$. In the following, we find the condition to ensure $d\rho_Y/dy_s > 0$. First notice that

$$\frac{d\rho_Y}{dy_s} \propto \frac{\delta q^*}{u'(y_s)} - y_\gamma - (y_s + y_\gamma)\left[\frac{d\delta}{dy_s}\frac{q^*}{u'(y_s)} - \frac{\delta q^* u''(y_s)}{(u'(y_s))^2}\right]. \tag{44}$$

Equilibrium conditions (38) and (43) imply that

$$\begin{aligned}
\frac{d\delta}{dy_s} &= -\frac{\left(u'\left(y_s\right)+y_s u''\left(y_s\right)\right) q_m + y_s \frac{v'(q_m)u''(y_s)}{v''(q_m)}}{q_m^2} \\
&= -\left(1-\sigma_y+\frac{\sigma_y}{\sigma_q}\right)\frac{u'\left(y_s\right)}{q_m}.
\end{aligned} \tag{45}$$

Plugging (45) into the right-hand side of (44) yields

$$\begin{aligned}
RHS &= \frac{\delta q^*}{u'\left(y_s\right)} - y_\gamma - \left(y_s+y_\gamma\right)\left[-\left(1-\sigma_y+\frac{\sigma_y}{\sigma_q}\right)\frac{q^*}{q_m}+\frac{\delta q^* \sigma_y}{y_s u'\left(y_s\right)}\right] \\
&= \frac{\delta q^*}{u'\left(y_s\right)} - y_\gamma + \left(y_s+y_\gamma\right)\left(1-\sigma_y+\frac{\sigma_y}{\sigma_q}-\frac{\delta}{1-\delta}\sigma_y\right)\frac{q^*}{q_m}.
\end{aligned}$$

The last equality uses the fact that $y_s u'(y_s) = (1-\delta)q_m$. Since $q_m < q^*$, it suffices to have $1-\sigma_y+\sigma_y/\sigma_q-\delta/(1-\delta)\sigma_y > 1$, which requires $\sigma_q < 1-\delta$. Because in situation II, $\delta \in \left(\tilde{\delta}, \hat{\delta}\right]$, it is sufficient that $\sigma_q < 1-\hat{\delta} = y_\gamma u'(y_\gamma)/q^*$. ∎

## A.3 Situation III: High Credit Regime

As the fraction of credit card users continues to increase, cash becomes more scarce in the cash-credit market relative to the settlement market. At a certain point, the value of cash in the two markets is equalized with $\phi = \omega$. The economy enters into situation III, and sellers start to accumulate some cash to make up for the decreasing cash revenue in the cash-credit market. In this situation, $q_m = q^*$ and $y_s = y_\gamma$. The money demand is given by

$$\begin{aligned}
z_V &= \phi\frac{(1-\delta)q^*}{\omega} = (1-\delta)q^*, \\
z_c &= \frac{\gamma}{\beta}y_\gamma \text{ (same as in situations I and II)} \\
z_m &= \phi\left[\frac{q^*}{\omega}+\frac{y_\gamma}{\beta\phi_+}\right] = q^* + \frac{\gamma}{\beta}y_\gamma, \\
z_s &= \phi\frac{y_s}{\psi} - z_V = \phi\frac{y_\gamma}{\beta\phi_+} - z_V = \frac{\gamma}{\beta}y_\gamma - (1-\delta)q^*, \\
z &= \left[\delta z_c + (1-\delta)z_m\right] + z_s = 2\frac{\gamma}{\beta}y_\gamma.
\end{aligned}$$

The velocity of cash is

$$\upsilon = \frac{z_V + \phi \frac{(1-\delta)y_m + \delta y_c + y_s}{\psi}}{z} = 1 + \frac{(1-\delta)q^*}{2\frac{\gamma}{\beta}y_\gamma}.$$

The value of GDP measured in units of $x$ is

$$\begin{aligned} Y &= \phi \left[ \frac{(1-\delta)q_m + \delta q_c}{\omega} + \frac{(1-\delta)y_m + \delta y_c + y_s}{\psi} \right] \\ &= q^* + 2\frac{\gamma}{\beta}y_\gamma. \end{aligned}$$

The value of notes in circulation over GDP is

$$\rho = \frac{z}{Y} = \frac{2\frac{\gamma}{\beta}y_\gamma}{q^* + 2\frac{\gamma}{\beta}y_\gamma}.$$

The effect of credit expansion in this situation is described in the proposition below.

**Proposition 7** *Effect of $\delta$ in situation III:* $dq_m/d\delta = dy_c/d\delta = dy_m/d\delta = dy_s/d\delta = dz/d\delta = d\upsilon/d\delta = d\rho/d\delta = 0$, $dz_V/d\delta = -q^* < 0$, *and* $d\upsilon/d\delta < 0$.

Once the fraction of buyers with credit cards exceeds the threshold value $\hat{\delta}$, a further credit expansion reduces the weight of cash transactions in the cash-credit market, but the total money demand curve becomes flat. The reduction in the buyer's money demand is exactly offset by the rise in the seller's money demand.

The cut-off value of $\delta$ that separates situations I and II, denoted by $\tilde{\delta}$, solves $y^* = (1-\delta)q_\gamma$, or

$$\tilde{\delta} = 1 - \frac{y^*}{q_\gamma}.$$

The cut-off value of $\delta$ that separates situations II and III, denoted by $\hat{\delta}$, solves $y_\gamma = (1-\delta)q^*/u'(y_\gamma)$, or

$$\hat{\delta} = 1 - \frac{y_\gamma u'(y_\gamma)}{q^*}.$$

# References

[1] Arango, C., K. P. Huynh, B. Fung, and G. Stuber (2012): "The Changing Landscape for Retail Payments in Canada and the Implications for the Demand for Cash," *Bank of Canada Review*, (Autumn): 31-40.

[2] Arango, C., and A. Welte (2012): "The Bank of Canada's 2009 Methods-of-Payment Survey: Methodology and Key Results," Bank of Canada Discussion Paper No. 2012-6.

[3] Bagnall, J., D. Bounie, K. P. Huynh, A. Kosse, T. Schmidt, S. Schuh, and H. Stix (2014): "Consumer Cash Usage: A Cross-Country Comparison with Payment Diary Survey Data," Bank of Canada Working Paper (forthcoming).

[4] Bagnall, J., S. Chong, and K. Smith (2011): "Strategic Review of Innovation in the Payments System: Results of the Reserve Bank of Australia's 2010 Consumer Payments Use Study," available at http://www.rba.gov.au/publications/consultations/201106-strategic-review-innovation/pdf/201106-strategicreview-innovation-results.pdf.

[5] Bailey, A. (2009): "Banknotes in Circulation – Still Rising. What does this Mean for the Future of Cash?" Keynote address, Banknote 2009 Conference, Washington DC.

[6] Berentsen, A., G. Camera, and C. Waller (2005): "The Distribution of Money Balances and the Non-Neutrality of Money," *International Economic Review*, 46, 465-487.

[7] Canadian Federation of Independent Business (2011): "Changing the Way We Pay: Getting the Transition Right for SMEs."

[8] Deutsche Bundesbank (2013): "Payment Behavior in Germany in 2011," Technical Report, Deutsche Bundesbank.

[9] Dunbar, G. R., and C. Fu (2012): "Income Illusion: Tax Evasion Measurement and a Test," Manuscript.

[10] Foster, K., S. Schuh, and H. Zhang (2013): "The 2010 Survey of Consumer Payment Choice," Federal Reserve Bank of Boston Research Data Report No. 13-2.

[11] Fung, B., M. Molico, and G. Stuber (2014): "Electronic Money and Payments: Recent Developments and Issues," Bank of Canada Discussion Paper No. 2014-2.

[12] Gu, C., F. Mattesini, and R. Wright (2013): “Money and Credit Redux,” Manuscript.

[13] Jonker, N., A. Kosse, and L. Hernandez (2012): “Cash Usage in the Netherlands: How Much, Where, When, Who and Whenever One Wants?” De Nederlandsche Bank Occasional Studies.

[14] Lagos, R., and R. Wright (2005): “A Unified Framework for Monetary Theory and Policy Analysis,” *Journal of Political Economy,* 113(3), 463-484.

[15] Lucas, R. E. Jr., and N. L. Stokey (1987): “Money and Interest in a Cash-in-Advance Economy,” *Econometrica,* 55, 491-513.

[16] Mooslechner, P., H. Stix, and K. Wagner (2006): “How are Payments Made in Austria? Results of a Survey on the Structure of Austrian Households’ Use of Payment Means in the Context of Monetary Policy Analysis,” *Monetary Policy and the Economy* (Q2), 111-134.

[17] Reserve Bank of Australia (2013): “Payments System Board Annual Report.”

[18] Rocheteau, G., and R. Wright (2005): “Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium,” *Econometrica,* 73, 175–202.

[19] Samuelson, P. A. (1958): “An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money,” *Journal of Political Economy,* 66, 467–482.

[20] Schneider, F., and A. Buehn (2012): “Shadow Economies in Highly Developed OECD Countries: What are the Driving Forces?” IZA Discussion Paper No. 6891.

[21] Terefe, B., C. Barber-Dueck, and M.-J. Lamontagne (2011): “Estimating the Underground Economy in Canada, 1992-2008,” Technical Report, Statistics Canada.

[22] Wang, Z., and A. L. Wolman (2014): “Payment Choice and the Future of Currency: Insights from Two Billion Retail Transactions,” Federal Reserve Bank of Richmond Working Paper No. 14-09.

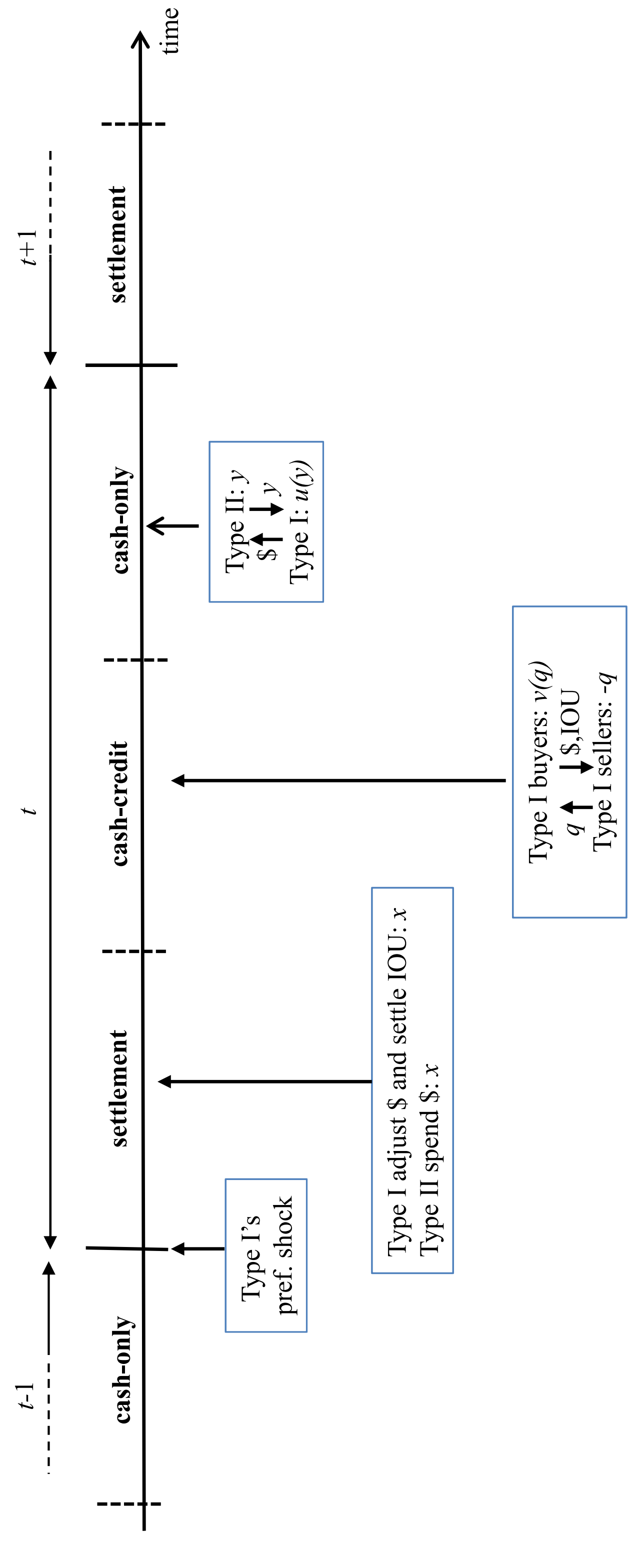


Figure 1: Timing

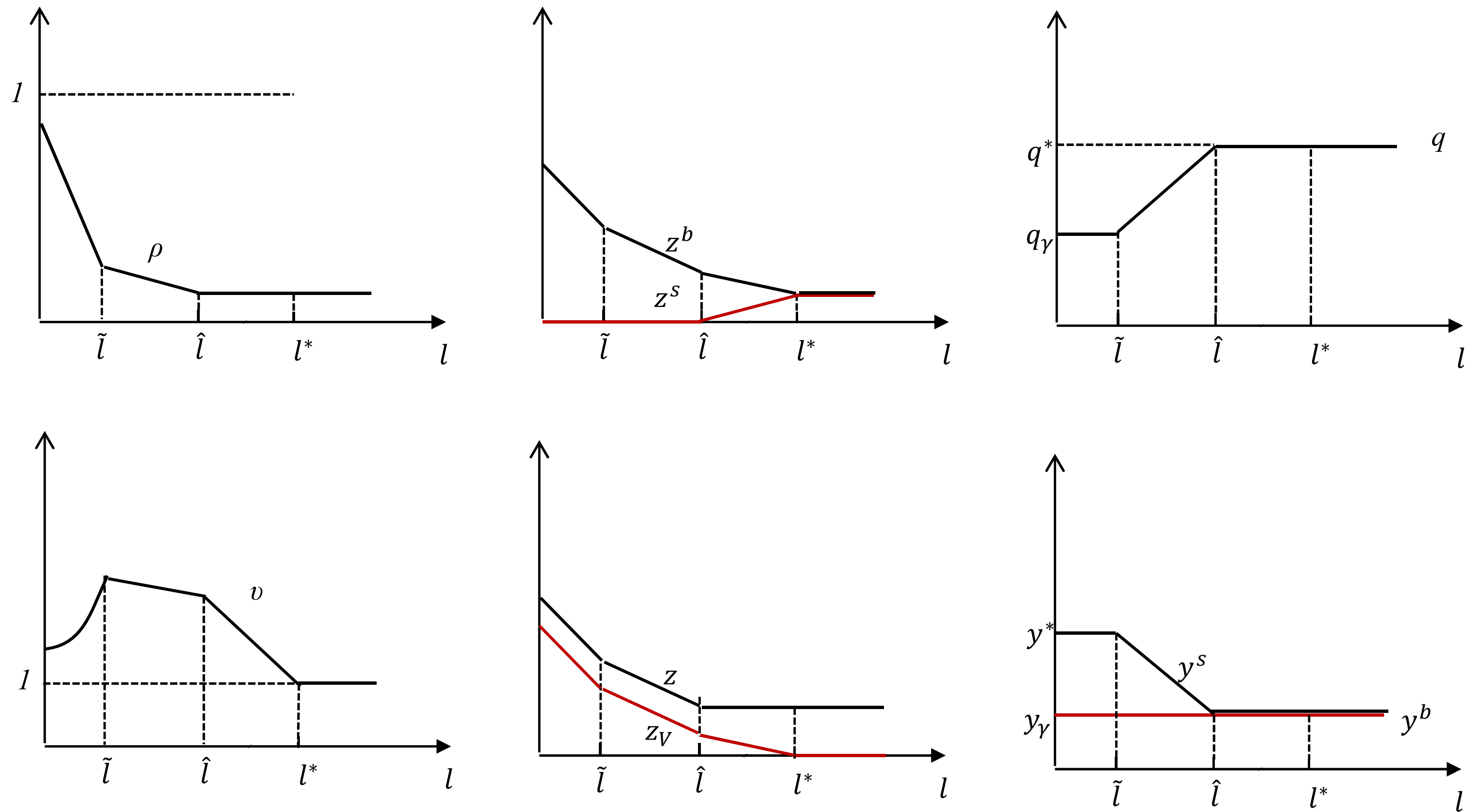


Figure 2: Effect of credit expansion

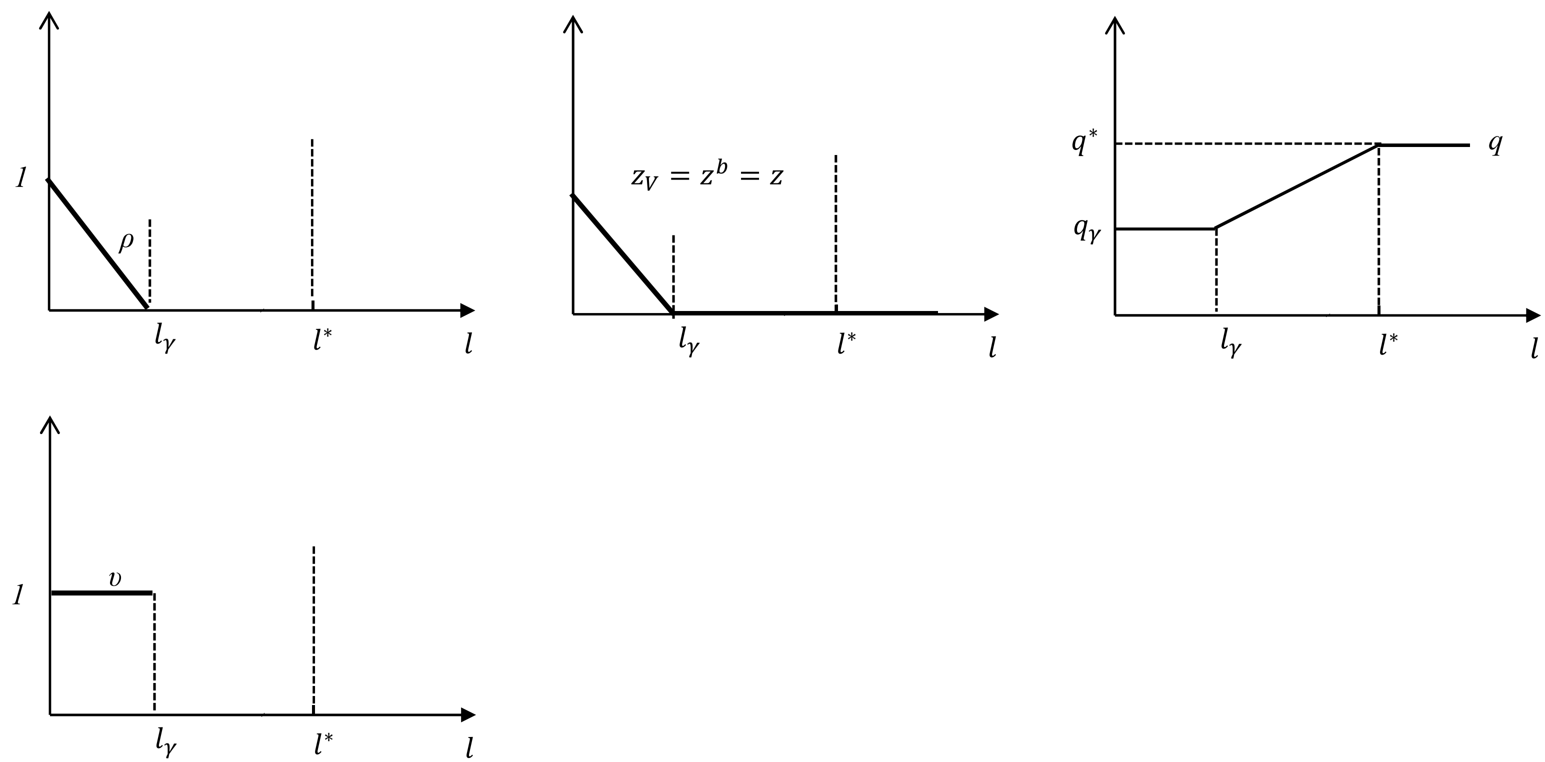


Figure 3: Effect of credit expansion in Lagos and Wright (2005)