BECKER FRIEDMAN INSTITUTE
FOR RESEARCH IN ECONOMICS
THE UNIVERSITY OF CHICAGO

Working Paper Series
No. 2017-06

Cash Burns: An Inventory Model with a Cash-Credit Choice

Fernando Alvarez and Francesco Lippi

April 3, 2017

JEL Codes: E41, G2

Becker Friedman Institute
for Research in Economics

Contact:
773.702.5599
bfi@uchicago.edu
bfi.uchicago.edu

Cash burns: An inventory model with a cash-credit choice*

Fernando Alvarez
University of Chicago, NBER

Francesco Lippi
University of Sassari, EIEF

April 3, 2017

Abstract

We present a dynamic cash-management model where agents choose whether to pay with cash or credit at every point in time. In the model credit usage depends on the current stock of cash, a novel result that matches recent micro evidence on household's payment choices. The optimality of such decision rule is novel and cannot be obtained by models where cash-credit decisions are made at the "beginning" of each period. We discuss how to use the model to account for cross country-evidence on the intensity of credit usage and for several statistics on the size and frequency of cash withdrawals. We use the model to assess the household's welfare cost of phasing out cash.

JEL Classification Numbers: E41, G2

*An earlier version of the paper circulated under the name: "A dynamic cash management and payment choice model". We benefited from the comments of the participants to "Consumer Finances and Payment Diaries: Theory and Empirics", Ottawa, Canada, 18-19 October 2012. Philip Barrett and David Argente provided excellent research assistance.

1 Introduction and overview

We propose a dynamic model of cash-management and means-of-payment choice in which optimizing households use both cash and credit. Credit is modeled as a payment instrument that involves a cost but requires no inventory at hand, such as a debit or a credit card, while use of cash is modeled as a standard inventory problem. A key feature of the model is that at each moment the agent can choose to pay with either cash or credit. This natural and realistic feature is an essential novelty of our model which implies that the preferred payment instrument depends on the stock of cash holdings at the time of the purchase: agents use cash for a purchase as long as they have cash with them, and use credit otherwise. Agents behave as if "cash burns" in their hands, a pattern that has been noticed by the empirical literature (see Arango, Huynh, and Sabetti (2011); Huynh, Schmidt-Dengler, and Stix (2014)). Yet agents find it optimal to intermittently replenish their cash holdings, and so use both cash and credit. Our model is the first in the literature to have both simultaneous use of cash and credit by households, as well as credit use that depends on the level of cash holdings. This feature is novel and cannot be obtained by models where cash-credit decisions are made at the "beginning" of each period since, by assumption, those models do not keep track of the dynamics of cash balances and so the use of credit cannot be conditioned on cash holdings. Our model provides a simple analytical mapping between the fundamental parameters and several observable statistics on cash holding behavior, concerning the size and frequency of cash withdrawals, as well as means of payment choices such as the share of purchases paid in cash. To illustrate its applicability we use our structural model to assess the household's welfare cost of phasing out cash, a policy endorsed by Rogoff (2016) as a measure against several illegal practices.

We consider a version of the Baumol-Tobin cash inventory model where each cash withdrawal is subject to a fixed cost augmented with two special features. First, we assume the agent randomly receives some *free* withdrawal opportunities, as in Alvarez and Lippi (2009). This assumption gives rise to a precautionary motive in the demand for cash holdings, such as an above zero cash-balance at the time of cash withdrawals, a featured extensively found in household diary and survey evidence. Second, we allow agents to pay for their (exogenous) constant flow of expenditure using either cash or credit. Our choice of modeling expenditures as a constant flow has the interpretation of small size purchases. Paying with cash requires to have positive cash inventory at the time of the purchase, which is costly to accumulate and which has an opportunity cost, both standard features of an inventory model. Paying with credit entails a cost per transaction.[1] In principle it is possible for the agent to pay all of its

[1] The choice of cash vs credit based on the size (i.e. dollar value) of purchases has been addressed in both theoretical and empirical literature, see Arango, Bouhdaoui, and Bounie (2012) and Bouhdaoui and

consumption using credit but, as long as there are some free cash withdrawal opportunities, it will be optimal for the agent to also use cash.

The paper has two main predictions for the use of payment instruments. First, as long as the cost of cash withdrawals is low enough relative to the cost of credit, the agents never use credit. They take advantage of their free trips to the bank, possibly some costly trips, and use cash only. Second, if the cost of cash withdrawals is sufficiently high, then cash is withdrawn only upon a free cash withdrawal opportunity. In this case the agent uses credit when she runs out of cash to finance her expenditure, while waiting for the next free cash withdrawal opportunity. Intuitively, for an agent with cash at hand, the cost of obtaining it is sunk. As a result, using cash is optimal since the agent pays only the opportunity cost. Thus, the agent prefers to pay in cash when she has it.

We analytically derive several model predictions and compare them with the cross-country evidence gathered from households surveys and diary data (as summarized in e.g. Bagnall et al. (2014)). As mentioned the model predicts that credit is more likely to be used when the agent is short of cash, a pattern that is robustly documented in micro studies.[2] Moreover, the model has predictions concerning the average share of expenditures paid with cash, the frequency of cash withdrawals and their average size, the average cash holdings (both unconditional as well as at the time of cash withdrawals). We show that for low inflation rates (a realistic assumptions for developed economies), the 5 moments listed above are functions of 2 parameters only: the normalized cost of credit, namely the ratio between the cost of credit and the cost of cash withdrawals, and the frequency of the free withdrawal opportunities. Using data for the US, we discuss how these two parameters can be identified from observations on e.g. the number of cash withdrawals and the share of cash purchases. We then compare the model predictions on moments such as the size of cash withdrawals and the average cash holdings with the data. We show that, in spite of its simplicity, the model gets the appropriate magnitudes for the observed moments.

Finally, we use our structural model to quantify the cost of a policy that limits household's cash usage, a policy endorsed by Rogoff (2016) to fight several illegal activities which are known to be cash intensive. Our objective is to quantify the welfare cost for households who are forced to move from their optimal cash-credit share to one where the cash share is zero.

Bounie (2012). Empirically, smaller transactions are more likely to be paid with cash than with credit, which motivates the assumption of a fixed cost per transaction in the literature. Yet, there are many small transactions paid with both cash and credit. Our assumption of a constant flow of expenditure thus focus on transactions that are all of the same size and small. It is thus complementary to the explanation in the literature based on size and it is able to address the choice of means of payments for small size transactions.

[2]See, for example, Stix (2004), Mooslechner, Stix, and Wagner (2006), Arango, Huynh, and Sabetti (2011), Arango, Hogg, and Lee (2013) and Huynh, Schmidt-Dengler, and Stix (2014). See Appendix A for a brief review of this evidence.

The analysis shows that such cost is small.

Related literature. Many papers in the literature incorporate alternative means of payments as in the seminal work of Lucas and Stokey (1987) and Prescott (1987). However such models do not have an explicit inventory theoretical model of money, so they cannot simultaneously speak to observations such as the fraction of purchases made in cash as well as cash-management statistics, such the frequency and size of cash withdrawals. Technically, in this type of models, the cash-in-advance constraint, which is exogenously determined, binds in every period. As a result, in every period withdrawals occur and all cash is spent. Hence, statistics such as frequency of withdrawals, size of withdrawals, cash at withdrawals are all exogenously determined by the choice of the model's time period. Other models incorporate both cash management and the choice of means of payments, which ends up being dictated by the size of the purchases. Examples of such models are Whitesell (1989) or Freeman and Kydland (2000). Yet while these models introduce cash-management, those choices are all "within" the period, so that agents cannot choose at every moment whether to use cash or credit. Hence in these models the optimal use of credit cannot depend on cash holdings, as the data strongly suggest.[3]

The closest related models in the literature are Sastry (1970) and Bar-Ilan (1990). Sastry (1970) is one of the earliest inventory model featuring a sequential cash versus credit choice. In his deterministic Baumol-Tobin model with no discounting the agent is allowed to use credit, namely an overdraft ("negative cash"), so that when cash holding reach zero the agent may continue to consume and postpone the payment of the fixed withdrawal cost. Bar-Ilan (1990) extends this setup to a dynamic stochastic inventory model. The main difference with our setup is that the cost of credit in both of these models is assumed to be proportional to the average *stock* of credit over the holding period, completely analogue to the opportunity cost of cash. This implies that the agent using credit will periodically decide to pay the fix rebalancing cost to keep the average stock of credit under control. Notice that under this assumption it is infinitely costly not to pay the fixed transanction cost since this implies a diverging stock of credit. In our setup, in contrast, the cost of credit is proportional to the expenditure *flow*, e.g. it is a fixed fraction of the purchase value. Using credit does not require any fixed cost to "rebalance" the cash credit stock. Credit purchases are immediately debited on to the agent's checking account. Our assumption explicitly distinguishes between the credit technology from the cash technology, and thus makes a credit-only strategy of purchases feasible for the agent.

[3] A close analogy between our sequential formulation and these papers' simultaneous cash-credit choice is found in the difference between sequential search, as in McCall's model, versus simultaneous search, as in Stigler's search model. See Sargent's (1987) chapter 2 for a description of the two types of search models.

Organization of the paper. The structure of the paper is as follows. We illustrate the model's key idea in Section 2 with a simple deterministic steady-state model. Section 3 introduces uncertainty and a proper dynamic treatment of the inventory problem with payment-choice. This section characterizes the conditions under which both cash and credit are used by the agents. Section 4 derives the model implications for the frequency and size of cash withdrawals and the intensity of credit usage. We discuss some cross-country evidence to illustrate how the model can be calibrated to actual economies. In Section 5, we use our structural model to quantify the cost of a policy that imposes a zero-cash usage restriction. Section 6 extends our model by allowing for a random cost of cash withdrawals, a feature which appears desirable for empirical applications. Section 7 concludes.

2 A deterministic model with means of payment choice

This section presents a steady-state deterministic model that highlights the main mechanism of the dynamic stochastic model of Section 3. Indeed some key formulas from this simple model coincide with, or are close to, the more complex decision rules of the stochastic model. The main counterfactual prediction of the deterministic model is the lack of a precautionary motive, so that real balances are always zero at the time of a withdrawal.[4]

Consider an agent who consumes e per unit of time and can pay for this using cash or credit. If she pays with credit she incurs a direct cost γ per unit bought. The cost γ can be understood as the time cost of using credit for small value transactions. The technology to withdraw cash (from an interest bearing checking account) is as follows: at any time the agent can pay a fixed cost b and replenish her cash balances which, as in the canonical Baumol-Tobin model, are subject to an opportunity cost R (e.g. forgone interest on deposits). Moreover, the agent has $p \geq 0$ withdrawals per period that come for free. The latter assumption is a simple parametrization of the technology for the cash withdrawals, proxying e.g. for the number of ATMs (cheap withdrawals) available to the agent.

To understand the nature of the optimal policies considered below notice that in a deterministic setup an agent with positive cash balances will not pay the fixed cost b to withdraw cash unless cash balances are zero. Consider now the decision of whether to purchase goods using cash or credit. For an agent with positive cash balances $m > 0$ it is not optimal to pay the cost $\gamma\, e$ to use credit, since the cost of acquiring the cash is sunk at this time.[5] For an agent with zero cash balances $m = 0$, there are two possible choices: the first one is to pay

[4] Readers familiar with continuous time impulse control problems may move directly to Section 3.

[5] Another possibility is to use credit and deposit the cash to earn a higher interest. With a fixed cost for depositing this is not optimal unless the cash balances are very large, a situation that will not occur along an optimal path.

the cost $\gamma\ e$ and finance consumption using credit, waiting until the next free withdrawal opportunity to replenish cash balances. The second choice is to pay the fixed cost b and withdraw cash. We first separately describe the solution of these two cases and next analyze the best choice among the two.

2.1 The deterministic cash-credit model

Consider an agent who finances her expenditures using cash, and who pays with credit once cash balances are depleted. Assume further that no costly withdrawal ever takes place so that b is never paid and the number of withdrawals n equals the number of withdrawals that come for free p. After a cash withdrawal of size $W = m^*$, she spends $\tau_a = m^*/e$ units of time paying for consumption with cash, incurring an opportunity cost $Rm^*/2$, where $m^*/2$ is the average cash balance conditional on $m > 0$ and R is the opportunity cost of cash– which includes the nominal interest rate as well as the probability of cash theft. After cash balances hit zero, the remaining time until a free withdrawal opportunity, denoted by τ_r, is given by $\tau_r = 1/p - m^*/e$. Notice that $\tau_r + \tau_a = 1/p$. The steady state cost in every cycle of duration $1/p$ can be written as: $\tau_r\ \gamma\ e + \tau_a\ R\ m^*/2$. The cost per unit of time is thus $p\,\tau_r\ \gamma\ e + p\,\tau_a\ R\ m^*/2$. Thus the minimized cost of the strategy that uses both cash and credit is:

$$v_r(R, \gamma, p, e) = \min_{0 \le m^* \le e/p} p \left[(1/p - m^*/e)\ \gamma\ e\ + (m^*/e)\ R\ e\ \frac{(m^*/e)}{2} \right],$$

subject to the constraint that the time spent using credit is non-negative, i.e. $m^*/e \le 1/p$. We denote by s the "cash share", namely the ratio of the expenditure paid with cash to total expenditure per unit of time, given by

$$s = \frac{\tau_a}{\tau_a + \tau_r} = \min \left\{ p \frac{m^*}{e}\ , 1\ \right\}.$$

Denoting the average real balances by M and using the cash share s we can write $M = s\ m^*/2$. The cost minimizing policy yields $\frac{m^*}{e} = \frac{W}{e} = \min\left\{ \frac{1}{p}\ ,\ \frac{\gamma}{R} \right\}$, $s = \min\left\{ 1\ ,\ \frac{\gamma\ p}{R}\ \right\}$, $\frac{M}{e} = \min\left\{ \frac{1}{2\ p}\ ,\ \frac{p}{2} \left(\frac{\gamma}{R}\right)^2 \right\}$ which imply

$$v_r(R, \gamma, p, e) \ = \ \begin{cases} \left(1 - \frac{\gamma\ p}{2\ R}\right) \gamma\ e & \text{if } R \ge \gamma p \\ \\ \left(\frac{R}{2\ p}\right) e & \text{if } R < \gamma p \end{cases} \tag{1}$$

When $R \geq \gamma p$ credit is "cheap", so that both cash and credit are used. When $R < \gamma p$ credit is expensive and it is not used.

2.2 Deterministic Baumol-Tobin model with p free withdrawals

Let us consider a modified Baumol-Tobin model in which the agent pays only for the withdrawals in excess of the p free adjustments per period.[6] The agent chooses a withdrawal of size m^* when cash balances are exhausted ($m = 0$). The policy implies an average cash balance $M = m^*/2$ and a number $n = e/m^*$ of cash withdrawals. The agent's choice of m^* gives the minimized cost function

$$v_a(R, b, p, e) \equiv \min_{m^*} \left[R\, \frac{m^*}{2} \; + b \max\left(\frac{e}{m^*} - p \; , \; 0 \right) \right] \quad .$$

where the cost is given by the sum of the opportunity cost of cash holdings and the cost associated with cash withdrawals in excess of p. The optimal policy for a technology with $p \geq 0$ is $\frac{m^*}{e} = \frac{1}{p}\sqrt{\min\left(2 \frac{b}{e} \frac{p^2}{R}, \; 1\right)}$. For $p > 0$ there is no reason to have less than p withdrawals, since these are free. Hence, for $R < 2p^2 b/e$ the agent will choose a constant level money holdings: $m^* = e/p$. Note that the interest elasticity of money is zero over this range, while it is equal to $1/2$ if $R > 2p^2 b/e$. The average withdrawal size W and the average cash balances satisfy: $W = m^* \quad , \quad M = 2 \; W = 2 \; m^*$. Replacing the optimal m^* choice in the cost function yields

$$v_a(R, b, p, e) = \begin{cases} \left(\sqrt{2 \; R \; \frac{b}{e}} - p \frac{b}{e} \right) e & \text{if } R \geq 2 \; p^2 \; \frac{b}{e} \quad \text{and } n > p \\ \\ \left(\frac{R}{2 \; p} \right) e & \text{if } R < 2 \; p^2 \; \frac{b}{e} \quad \text{and } n = p \end{cases} \tag{2}$$

where the top branch gives the cost for the case in which the number of withdrawals exceeds p. Note that in this deterministic setup an agent with positive cash balances will not pay the fixed cost b to withdraw cash unless cash balances are zero.

2.3 The full deterministic problem

We now analyze the conditions under which it is optimal to use credit instead of withdrawing fresh cash when $m = 0$. To do so we compare the steady-state cost of the two policies computed above. The value function for the problem is then $v(R, b, \gamma, p, e) =$

[6] See Alvarez and Lippi (2009) for a more detailed analysis of this model and Appendix C for estimates of cash theft probabilities, which are a component of R, in Italy and the US.

$\min\left\{v_a\left(R,b,p,e\right)\ ,\ v_r\left(R,\gamma,p,e\right)\right\}$. We define the threshold function $\underline{b}$, as the value of b that equates the two minimized costs: $v_a(R,\underline{b},p,e) = v_r(R,\gamma,p,e)$. We have that

$$\underline{b}(R,\gamma,p,e) = \frac{\gamma^2}{2\ R}\ e\ . \tag{3}$$

which implies that credit is used when $b \geq \underline{b}$ provided that $\gamma p \leq R$.[7] If $b \geq \underline{b}$ and $\gamma p > R$ then credit is not used and $n = p$. Finally, for $b < \underline{b}$ credit is not used and $n > p$ since some costly withdrawals in access of the p free withdrawals are now optimal .

The next proposition summarizes the behavior of the deterministic model. It considers two cases depending on whether $\gamma \gtreqless 2p\,b/e$, and for each case it analyzes optimal policy as a function of R.

PROPOSITION 1. Let $p > 0$ and $\gamma > 0$. Then $W/M = 2/s$ and

If $\gamma > 2\,p\,\frac{b}{e}$, then

if $R \in \left(0\ ,\ 2\,p^2\,\frac{b}{e}\right]$	$-\frac{\partial \log M/e}{\partial \log R} = 0$	only cash used	$n = p$	$s = 1$
if $R \in \left(2\,p^2\,\frac{b}{e}\ ,\ \frac{\gamma^2}{2\,b/e}\right]$	$-\frac{\partial \log M/e}{\partial \log R} = 1/2$	only cash used	$n > p$	$s = 1$
if $R \in \left(\frac{\gamma^2}{2\,b/e}\ ,\ \infty\right)$	$-\frac{\partial \log M/e}{\partial \log R} = 2$	cash & credit used	$n = p$	$s = \gamma p/R$

Otherwise, i.e. if $\gamma \leq 2\,p\,\frac{b}{e}$, then

if $R \in (0, \gamma p]$	$-\frac{\partial \log M/e}{\partial \log R} = 0$	only cash used	$n = p$	$s = 1$
if $R \in (\gamma p\ ,\ \infty)$	$-\frac{\partial \log M/e}{\partial \log R} = 2$	cash & credit used	$n = p$	$s = \gamma p/R$

The proposition illustrates three robust properties of the model. First, the model has only two parameters, γp and $p^2\,b/e$, as the alert reader will have noticed. In the modified Baumol-Tobin model the shape of the money demand depends only on $\hat{b} \equiv p^2\,b/e$. For a given value of $\hat{b}$, the cash-credit aspect of the model depends only on γp. We will see that this property continues to hold in the stochastic model analyzed below. Second, for credit to be used it is necessary that the cost of a withdrawal is above the threshold defined in equation (3). If $b > \underline{b}$ then the agent uses both cash and credit to finance her consumption, and costly withdrawals are never used. The condition for the optimality of credit depends on a combination of the

[7] The expression for $\underline{b}$ comes from equating: $v_r = \gamma e[1 - \gamma p/(2R)]$ with $v_a = e\sqrt{2\ R\ \frac{b}{e}} - pb$.

fundamental parameters R, b and γp which, intuitively, imply that the cost of cash (which is increasing in R and b) must be high relative to the cost of credit (which is increasing in γ and p). Third, the interest rate elasticity of money demand is increasing in the interest rate. There are two cases: the first corresponds to a large cost of credit ($\gamma p > 2\,p^2\,b/e$), in which case there are three qualitatively different behavior depending on the level of interest rates. If interest rates are very low, credit is not used and $n = p$, resulting in an elasticity of zero. For intermediate level of interest rates, credit is not used, but $n > p$, so the local behavior is identical to Baumol-Tobin, producing an interest rate elasticity of 1/2. For higher interest rates, both cash and credit are used. The interest rate elasticity is higher here because both the cash share as well as the size of the withdrawals react to interest rates. If instead the cost of credit is low ($\gamma \leq 2\,p\,b/e$) there is no intermediate case, since credit always dominates the Baumol-Tobin type of behavior.

3 A dynamic stochastic model with means of payment choice

In this section we solve a discounted, stochastic dynamic problem which joins the optimal cash management problem with the optimal choice of means for payment. As in the deterministic problem the agent faces a total consumption per unit of time $e > 0$ which must be paid with either cash or credit: at each instant the agent can choose to pay in cash $c \in [0,\ e]$ and to pay the remaining $e - c$ using credit. If the payment is made by credit, the agent pays a flow cost γ per dollar.[8] The quantity γe can be interpreted as the "handling" and "verification and authorization" costs estimated in Klee (2008) using grocery receipts data. The state of the agent's problem is given by her real cash balances $m \geq 0$. If $m = 0$ either cash must be withdrawn or credit has to be used. If $m > 0$ the agent faces a cash/credit choice. The law of motion of real balances is then $dm = -\,(c + m\,\pi)\,dt$ provided that no adjustment takes place, where π is the constant inflation rate. The agent can adjust her cash balances paying the fixed cost $b \geq 0$. Additionally, there is a Poisson process with constant arrival rate $p \geq 0$, which describes the arrival of a free adjustment opportunity. When such an opportunity occurs the agent can adjust her cash balances at no cost. As standard in monetary models, we assume that holding cash m entails an opportunity cost $R\,m$ per unit of time, where R can be interpreted as the sum of the nominal interest rate plus a probability that cash is lost or stolen. We assume that the agent minimizes the expected discounted cost, using a constant discount rate $r \geq 0$. There are three substantive differences between the model

[8] It turns out that, given that all expenditures are of the same size, it is equivalent to assume that there is a fixed cost, since the optimal policy will be of the bang-bang type.

analyzed in this section and the steady-state deterministic model of Section 2. First, we take into account explicitly the role of inflation, as can be seen in the law of motion. Second, the free adjustment opportunities arrive stochastically. Third, real costs are discounted by an appropriate rate r.

Formally we denote by $V(m)$ the minimum expected discounted cost of supporting a constant flow of expenditure e when the current real cash at hand is $m \geq 0$. The function V, defined in $V : \mathbb{R}_+ \to \mathbb{R}_+$ must solve the following functional equation:

$$0 = \min \Bigg\{ \min_{0 \leq c \leq e} R\, m + \gamma[e-c] + p \min_{z \geq 0} [V(z) - V(m)] - V'(m)(c + \pi m) - r\, V(m) \ , \\ b + \min_{z \geq 0} V(z) - V(m) \Bigg\} \quad \text{for all} \ \ m \geq 0 \ . \tag{4}$$

The outer $\min\{\cdot, \cdot\}$ in the functional equation (4) chooses between two strategies. The term in the first line represents the case where no costly withdrawals (that involve paying the fixed cost b) occur, although random free withdrawal opportunities may arise, and the agent chooses what fraction of her consumption to pay in cash versus credit. This is a standard continuous time Bellman equation, with flow cost $Rm + \gamma(e - c)$ and with expected changes due to either the arrival of the free adjustment opportunity or to the depletion of cash. The minimization with respect to c describes the agent's choice of the optimal means of payment. The term in the second line corresponds to the strategy of exercising control, i.e. paying the fixed cost b and adjusting cash holdings. For each m the value function is equal to the value of the strategy that yields the minimum cost. Whenever an adjustment is made, either paying the cost b or when a free adjustment opportunity arrives, the post-adjustment quantity of cash is chosen optimally. The optimal policy for the problem in equation (4) consists of deciding for each $m \geq 0$ whether a costly withdrawal is made or not and, if no adjustment is made, which payment instrument (cash or credit) to use. Notice that this formulation does not impose any restriction concerning when the adjustments takes place or when different payment instruments are used. We maintain the following assumption throughout this section.

Assumption 1. We let $b \geq 0$, $\gamma \geq 0$, $\pi \geq 0$, $p \geq 0$, and $r + p + \pi > 0$, $e > 0$, $R > 0$.

If $e = 0$ the problem becomes uninteresting since there is no expenditure to finance. The parameters b and γ are costs, and p a probability rate, so they must be non-negative. The requirement that $r + p + \pi > 0$ and $R > 0$ are important. For instance, if $r + p + \pi = 0$ there is no intertemporal incentives to use cash.[9]

[9] While here we treat R and r, p, π as independent parameters the value of R and $r + p + \pi$ can indeed

3.1 Two candidate policies

It will turn out that the optimal policy of the problem depicted in equation (4) is one of two types, depending on parameters. We refer to one of the policies as a cash-burning policy, defined as follows:

Definition 1. We define a *m^*-cash burning policy* as a threshold $m^* \geq 0$ for which:

1. Credit is only used when $m = 0$, and cash is used for every $m \in (0, m^*)$.
2. Cash is only adjusted when a free adjustment opportunity arrives.
3. Immediately after a cash adjustment, cash holdings m take the value m^*.

Note that the value of following a m^*-cash burning policy is the function $V : [0, m^*] \to \mathbb{R}_+$ that satisfies:

$$(r+p)V(m) = mR + pV(m^*) - V'(m)[\pi m + e] \quad \text{for all } m \in (0\,,\,m^*] \quad \text{and} \tag{5}$$

$$(r+p)V(0) = \gamma e + pV(m^*) \tag{6}$$

The first one is the standard Bellman equation, where we assume everything is paid in cash. The second equation says that at $m = 0$ agents use credit and wait for a free withdrawal opportunity. For completeness we define an alternative policy, which we refer to as a modified Baumol-Tobin (BT) policy:

Definition 2. We define a *m^*-Baumol-Tobin policy* as a threshold $m^* \geq 0$ for which:

1. Credit is never used.
2. Cash is adjusted when either a free adjustment opportunity arrives or when $m = 0$.
3. Immediately after a cash adjustment, cash holdings m take the value m^*.

We briefly comment on the differences between the two polices. There is a sense in which cash burns in the agent's hands under both policies, since in both cases, as long as it is available ($m > 0$), cash is the preferred means of payment. Note that when a m^*-Baumol-Tobin policy is followed the o.d.e. in equation (5) holds in the range of inaction $(0, m^*]$. However, under this policy the boundary condition at $m = 0$ is given by:

$$V(0) = b + V(m^*) \ . \tag{7}$$

be related –for instance $r + \pi$ should be the shadow nominal interest rates. We return to this relationship in the next section.

For a cash burning policy to be optimal, i.e. to solve the problem in equation (4), one needs to establish that it is optimal to pay with cash at $m \in (0, m^*]$ and with credit at $m = 0$, where the optimal withdrawal m^* must be determined. Finally, it has to be shown that at $m = 0$ it is optimal to wait for a free adjustment opportunity (instead of paying b to withdraw). Likewise, for a BT policy to be optimal, i.e. to solve the problem in equation (4), one needs to establish that it is never optimal to pay with credit and that at $m = 0$ it is optimal to pay b and choose the optimal withdrawal level m^* (see Appendix B for a formal proof that all these properties are verified under the optimal policy).

Note that the feasible policies consistent with equation (4) are much broader than the two candidate policies defined above. For instance, one could consider a policy in which credit is used for some time at $m = 0$ and a costly withdrawal occurs after T periods unless a free withdrawal arrives. Such a policy is the optimal one in the models of Sastry (1970) and Bar-Ilan (1990). Interestingly, as we explain below, the optimal policy for the problem in equation (4) will either take the form of a modified Baumol-Tobin one (where credit is not used) or of a cash-burning policy.

3.2 Characterizing the optimal cash-credit choice

The next proposition characterizes the optimal cash vs credit choice. Appendix B provides the proof as well as a detailed analytic characterization of decision the rules, including approximate solutions for m^* in the case of zero inflation.

Proposition 2. A cash burning policy with m^* given by equation (9) is optimal provided that $b \geq \underline{b}$ where the lower bound for the fixed cost of adjustment is given by

$$\underline{b} = \frac{e}{r+p}\left[\gamma - R\,\frac{m^*}{e}\right] \ , \tag{8}$$

and m^*/e solves:

$$0 \leq \frac{m^*}{e} = \frac{\left(1 + (r+p+\pi)\frac{\gamma}{R}\right)^{\frac{\pi}{\pi+r+p}} - 1}{\pi} \leq \frac{\gamma}{R} \ . \tag{9}$$

Instead if $b \leq \underline{b}$ a Baumol-Tobin policy is optimal and m^* solves:

$$\left(1 + \frac{m^*}{e}\pi\right)^{1+(r+p)/\pi} = \frac{m^*}{e}(r+p+\pi) + 1 + (r+p)(r+p+\pi)\frac{b}{eR} \ . \tag{10}$$

The proposition shows that there is a threshold $\underline{b}$ for the fixed cost of adjustment b above which the cash-burning policy is optimal and below which the Baumol-Tobin policy is optimal. When $b > \underline{b}$ the optimal policy consists of using both cash credit. A withdrawal of size m^*, as determined by equation (9), occurs every time a free withdrawal opportunity arises. When cash eventually hits $m = 0$, then it is optimal to finance consumption using credit until a free opportunity for a cash withdrawal arises. Thus, under cash-burning $n = p$ and the fixed cost b is never incurred by the agent. Conversely, when $b < \underline{b}$ credit is not used and the optimal policy at $m = 0$ consists in paying the fixed cost b to make a withdrawal of size m^*, as determined by equation (10). Notice that the optimal threshold $\underline{b}$ defined in equation (8) summarizes the effect of all fundamental parameters $(\gamma, R, \pi, r, p, e)$ into one single function that determines the nature of the optimal policy and, in particular, the optimality of using credit. Intuitively, equation (8) implies that the use of credit is optimal whenever the cost of using cash (which is increasing in R and b, and decreasing in p) is high relative to the cost of credit usage (which is increasing in γ).

Finally we notice that the optimal policy is of the bang-bang type, in the sense that using credit strictly dominates, or is dominated by, the use of costly withdrawals. This result differs from the ones of Sastry (1970) and Bar-Ilan (1990) where the use of some costly credit as well as some costly withdrawals is optimal. As mentioned in the introduction, this difference originates on the way the cost of credit is modeled in these papers compared to ours. We assume that the cost of credit is proportional to the expenditure *flow*: γe. They instead assume the cost of credit to be proportional to the *stock* of accumulated credit, the exact analogue of "negative cash", which periodically requires the agent to pay the fixed cost b to rebalance the credit cost which would otherwise diverge. We see our assumption as a reasonable description of the case of revolving credit (credit that is automatically debited on the agent's checking account at the end of the holding period).

The next proposition analyzes how the threshold $\underline{b}$ changes as a function of the parameters:

Proposition 3. The function $\underline{b} \geq 0$ is bounded above by $e\gamma/(r+p)$ and it is homogenous of degree one in (γ, R). Moreover

$$
\begin{aligned}
\frac{\partial \underline{b}}{\partial \gamma} &> 0 \ \text{ with } \ \lim_{\gamma \to 0} \underline{b} = 0 \ \text{ and } \ \lim_{\gamma \to \infty} \underline{b} = \infty \ , \\
\frac{\partial \underline{b}}{\partial R} &< 0 \ \text{ with } \ \lim_{R \to 0} \underline{b} = \frac{e\,\gamma}{(r+p)} \ \text{and} \ \lim_{R \to \infty} \underline{b} = 0 \ , \\
\lim_{r+p \to 0} \underline{b} &= \frac{e\,R}{\pi^2} \left[\left(1 + \frac{\gamma \pi}{R}\right) \log\left(1 + \frac{\gamma \pi}{R}\right) - \frac{\gamma \pi}{R} \right] = \frac{\gamma^2}{2\,R}\, e \ + \ \pi\, o\left(\frac{\gamma^2}{R}\right) \geq 0 \ , \\
\lim_{\pi \to 0} \underline{b} &= \frac{e\gamma}{r+p} \left[1 - \frac{\log\left(1 + (r+p)\frac{\gamma}{R}\right)}{(r+p)\frac{\gamma}{R}} \right] = \frac{\gamma^2}{2\,R}\, e \ + \ (r+p)\, o\left(\frac{\gamma^2}{R}\right) \geq 0 \ , \\
\frac{\partial \underline{b}}{\partial \pi} &< 0 \ , \ \lim_{\pi \to \infty} \underline{b} = 0 \ , \ \frac{\partial \underline{b}}{\partial (r+p)} < 0 \ , \ \text{and} \ \lim_{r+p \to \infty} \underline{b} = 0 \ .
\end{aligned}
$$

The proposition shows that the critical threshold $\underline{b}$ is increasing in the credit cost γ and decreasing in the opportunity cost of using cash R. In addition, by varying γ/R the threshold $\underline{b}$ ranges from zero to infinity. The approximations in the second to last line shows that if γ^2/R is small then $\underline{b}$ coincides with the one of the deterministic model. Moreover, the threshold $\underline{b}$ is decreasing in inflation: higher inflation increases the range of parameters for which the cash burning policy is optimal. Notice however that for finite values of γ and R, $\underline{b} > 0$ which implies that there exists a sufficiently small value of $b > 0$ that makes the use of credit dominated by a cash-only policy. Also notice that in our model a credit-only policy, i.e. one where cash is not used, does not occur as long as $p > 0$. The last two results suggest that the use of cash is very resilient: technical innovations that reduce the cost of credit are also likely to reduce the cost of cash withdrawals (increase p and or lower b) so that cash usage remains convenient for agents.

4 Model predictions about observable moments

Next, we consider several statistics of interest generated by a household who follows the optimal policy described above. We denote by $s \equiv c/e$ the cash share, namely the long run average fraction of purchases paid with cash. We denote by M the average cash holdings of the household. This is the expected value of real balances under the invariant distribution of real balances (m) implied by the optimal decision rules. We let n be the expected number of withdrawals per unit of time and W the expected size of withdrawals under the invariant distribution. Finally, we let $\underline{M}$ be the expected value of cash at the time of a withdrawal. Table 1 reports sample means for each of these moments for a few OECD economies, taken from household surveys and diary data by Bagnall et al. (2014) and Alvarez and Lippi (2013).[10] One main difference is a markedly smaller fraction of expenditures that is paid in cash in the US (or France, around 20% of total expenditure) relative to e.g. Germany (or Italy, where it is around 50% of the total expenditure). The higher cash use in Germany and Italy is also reflected in higher values of cash holdings (M/e).

Next, we illustrate how those statistics map into the fundamental parameters of the dynamic model considering two cases: first a household who follows the cash burning policy, which is optimal when $b > \underline{b}$, so that credit is used. Second, the case when $b < \underline{b}$ and credit is not used. We stress that the main empirical appeal of the cash burning policy is twofold.

[10] Bagnall et al. (2014) analyze data from large-scale payment diary surveys conducted between 2009 and 2012 in Australia, Austria, Canada, France, Germany, the Netherlands and the United States that enable international comparisons. Alvarez and Lippi (2013) focus on Austria and Italy.

Table 1: Selected moments on cash holding patterns

	Fra	Ger	Ita	US
Cash balances (*median*), M/e^d	1.3	2.6	6.5	0.6
Number of cash withdrawals, n	96	57	55	64
Withdrawal size (*median*), W/M	1.7	2.1	1.3	2.3
Cash at withdrawals (*median*), $\underline{M}/M$	-	0.3	0.4	0.7
Cash share of expenditures, $s = c/e$	0.15	0.53	0.52	0.23

The data source is Bagnall et al. (2014) Tables 1 to 4. Entries are sample means (unless otherwise indicated). The Italian data are from Alvarez and Lippi (2013) for household who posses a ATM card. Cash balances M/e^d are measured relative to total expenditures per day, $e^d = e/365$. The number of cash withdrawals is per year.

First, it is consistent with the data as it rationalizes the use of both cash and credit. And, second, the policy aligns with the empirical observation that households are more likely to use credit when their cash balances are running low, a fact that is amply documented in Arango, Huynh, and Sabetti (2011); Kosse and Jansen (2012); Huynh, Schmidt-Dengler, and Stix (2014); Arango, Bouhdaoui, and Bounie (2012). The next proposition summarizes the main results and closed form expressions for the observables focusing on the simple case of zero inflation $\pi = 0$.

Proposition 4. Let $s \equiv c/e$ denote the cash share., i.e. the share of purchases paid with cash. For the cash-burning and the Baumol-Tobin policy we have:
(i) Cash-burning policy. Let $r + p > 0$, $\pi = 0$, $R > 0$ and $b \geq \underline{b}$. Then

$$n = p \ , \ s = 1 - \left(1 + \frac{\gamma(r+p)}{R}\right)^{-\frac{p}{r+p}} \ , \ \frac{M}{e} = \frac{m^*}{e} - \frac{s}{p} \ , \ W = \frac{e}{p} \, s \ ,$$

thus, using equation (9) with $\pi \downarrow 0$ gives $\frac{m^*}{e} = \frac{1}{r+p} \log\left(1 + \frac{\gamma(r+p)}{R}\right) \geq 0$. For $r \downarrow 0$ these expressions are simple functions of $\gamma p/R$:

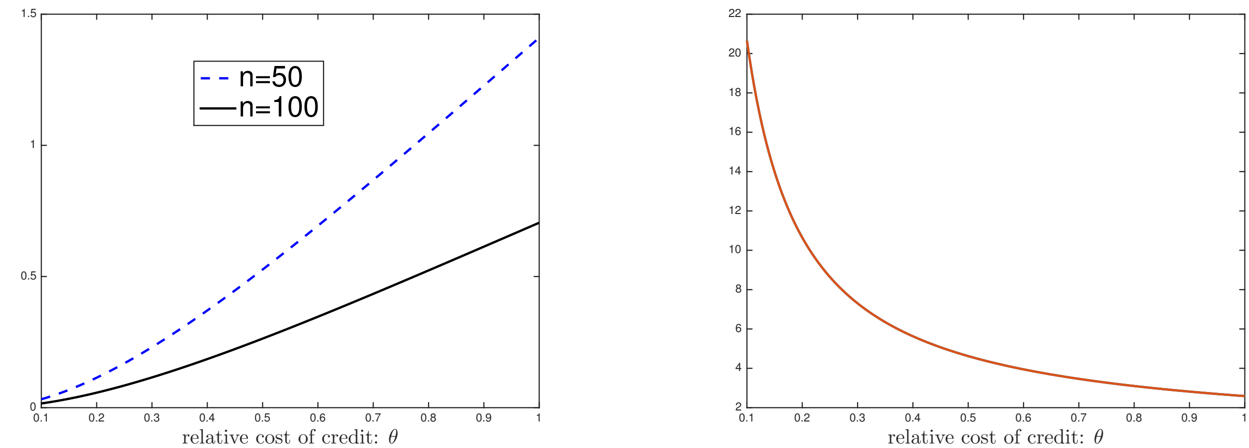
$$s = \frac{\frac{\gamma p}{R}}{1 + \frac{\gamma p}{R}} \ , \ \frac{M}{e} = \frac{1}{p}\left[\log\left(1 + \frac{\gamma \, p}{R}\right) - 1 + \left(1 + \frac{\gamma \, p}{R}\right)^{-1}\right] \text{ and}$$

$$\frac{W}{M} = \frac{1}{\left(1 + \frac{R}{\gamma \, p}\right) \log\left(1 + \frac{\gamma \, p}{R}\right) - 1} \text{ and } \underline{M}/M = 1 \ .$$

(ii) Baumol-Tobin policy. Let $R > 0$, $p > 0$, $\pi = r = 0$ and $b < \underline{b}$. Then $s = 1$ and

Figure 1: Moments under cash-credit policy

Cash relative to daily expenditure $M/(e/365)$ | Withdrawal to average cash: W/M



$$
\begin{aligned}
n &= \frac{p}{1-e^{-m^* \frac{p}{e}}} > p \ , \ \frac{W}{M} = \frac{m^*}{M} - \frac{p}{n} \in (0,2) \ , \\
\frac{M}{e} &= \left[\frac{1}{1-e^{-\frac{p}{e}m^*}} \frac{m^*}{e} - 1/p \right] , \quad \text{and} \quad \underline{M}/M = \frac{p}{n} \ .
\end{aligned}
$$

The proposition gives closed form expression for n, s, M, W and $\underline{M}$ under either of the two optimal policies for low-inflation ($\pi \to 0$) and low-discounting ($r \to 0$), both reasonable assumptions for developed economies. The model is over identified since it has essentially two parameters, p and $\gamma p/R$ that determine 5 observables. Next we briefly comment on the economics of the proposition.

Define the scalar $\theta \equiv \gamma p/R$ as the normalized cost of credit, namely the cost of credit (γ) relative to the cost of cash (R/p). Under cash-burning, the cash share is $s = \frac{\theta}{1+\theta}$ which is monotone increasing in the normalized cost of credit, with $s \to 1$ as $\theta \to \infty$ and with $s \to 0$ as $\theta \to 0$. The equation shows that for the credit share to increase, due to e.g. the new availability of cheaper credit cards, it is necessary that the cost of credit falls faster than the cost of cash. This simple observation may account for the remarkable resilience of cash usage in several developed countries (as documented by e.g. Bagnall et al. (2014)). If technical progress in payment instruments reduces the cost of credit as well as the cost of cash, such that θ is constant, then cash usage is unaffected.

The money demand M/e under the cash-burning policy is a function of two parameters: $\theta \equiv \gamma p/R$ and p. The left panel of Figure 1 plots the average money holdings M relative to daily total expenditures $e/365$ for $p = 50$ and $p = 100$. Recall that p also equals the number of cash withdrawals n in the cash-credit model. The right panel of the figure plots the average withdrawal size relative to the average money holding, W/M, under a cash-burning policy. The ratio W/M is decreasing in θ, ranging from $W/M \to 0$ as $\theta \to \infty$, to $W/M \to \infty$ as $\theta \to 0$, as shown in Figure 1. For comparisons, when $b < \underline{b}$ and the modified BT policy is optimal, we have that $W/M \leq 2$.

Since the cash-credit model has essentially two parameters it is simple to calibrate it, as well as to test it on data. For instance, a credit-intensive country like the US where $s \approx 0.25$ implies a $\theta \approx 0.4$, which in turn suggests a normalized withdrawal size $W/M \approx 5$. The latter value is between the median and the mean estimates for W/M reported by Bagnall et al. (2014). Also, using that $n = p$, the model identifies a money demand curve for M/e, which for $\theta \approx 0.4$ predicts average money holdings near 0.5; a value that is not far from the 0.6 observed in the data as shown inTable 1.

Elasticities of observables with respect to θ. We conclude with a discussion of the model predictions concerning the interest rate elasticity of some key statistics which have been considered in the empirical literature, such as the interest elasticity of money demand which is essential to measure the welfare costs of inflation, see e.g. Lucas (2000). The elasticity of s with respect to $\theta \equiv \gamma p/R$ is $1/(1+\theta)$, which implies that the interest elasticity of s is given by $\frac{\partial \log s}{\partial \log \theta} = 1 - s$, or one minus the cash share, so that the elasticity is decreasing in θ, and smaller than one. We use the expression for the cash share to illustrate the tension between the use of cash vs. credit and the role of p. If $b < \underline{b}$ *only* cash is used, i.e. $s = 1$. If $b > \underline{b}$ both cash and credit are used. Yet, if p tends to zero, then $\underline{b}$ is finite, so there are values for which $b > \underline{b}$ and *only* credit is used, i.e. $s = 0$. Thus, we find that the interesting case is the one where $b > \underline{b}$ and $p > 0$, so that $0 < s < 1$.

The elasticity of Mp/e is: $0 \leq \frac{\partial \log Mp/e}{\partial \log \theta} = \frac{\left(\frac{\theta}{1+\theta}\right)^2}{\log(1+\theta) - \theta/(1+\theta)} \leq 2$ and is decreasing in θ. Thus money demand is decreasing in the opportunity cost R, with an interest rate elasticity *increasing* in R satisfying:

$$0 \leq -\frac{\partial \log M/e}{\partial \log R} = \frac{\partial \log Mp/e}{\partial \log \theta} \leq 2$$

For comparison, the model with cash purchases only ($c = e$) summarized in part (ii) of Proposition 4, has an (absolute value) of the interest rate elasticity of money to *cash* consumption, M/c, that is increasing in the level R, but bounded above by $1/2$. This difference reflects the elasticity of the cash share s, which is between 0 and 1, and the elasticity of the money demand relative to *cash* consumption M/c, also between 0 and 1:

$$0 \leq -\frac{\partial \log M/c}{\partial \log R} = \frac{\gamma p/R}{1+\gamma p/R} \; \frac{1}{\log\left(1+\gamma p/R\right)} \leq 1 \; .$$

The higher interest elasticity of money demand produced by the model with a cash credit margin illustrates the importance of jointly modeling the cash-inventory problem and the cash-credit choice. By doing so the agent chooses both the extensive margin (how much cash vs credit to use) as well as the intensive margin (how many cash withdrawals to make). Failure to account for the interaction between these two margins, as done for instance in Alvarez and Lippi (2009) where the cash expenditure is taken as exogenous, leads to underestimate the interest rate elasticity of money demand.

The simple model has the ability to qualitatively capture several empirical facts, such as the use of both cash and credit and the fact that credit is used when cash is low. Moreover the basic parametrization of the model, discussed above using US data on the cash share and the number of cash withdrawals, delivers magnitudes for the average cash holdings and cash withdrawals that are in the ballpark of observed statistics. Our parsimonious 2 parameter model, nonetheless, also has some clear shortcomings to match the data. For instance, notice that the average cash at the time of withdrawal $\underline{M}$ equals the mean cash holdings M under a cash-burning policy. This somewhat surprising and stark result is an immediate consequence of the fact that withdrawal times under the cash-burning policy are uncorrelated with money holdings (their arrival rate p is exogenous and independent of m). Since a withdrawal is equally likely to happen with any money balance $m \in (0, m^*)$, then the two statistics are the same. This prediction of the model is in contrast with the data where $\underline{M} < M$, an issue that motivates the extension of Section 6.

5 An application: How costly is it to ban cash usage?

In this section we use our model to quantify the household's cost of a policy that limits cash usage. The motivation for such a policy is that, because of its anonymous nature, cash is heavily used for illegal activities. Rogoff (2016) argues that phasing out paper money would help fighting some big problems as corruption, tax evasion, drug trade and others. Because of this argument some countries, like Sweden, have been gradually pursuing the objective of a cashless economy.

In this section we use our model to study the cost forcing all agents to be cashless, i.e. to make all payment using credit only. We use the agent's value function for a cash credit policy to quantify the welfare cost of moving from the optimal policy, where the agent chooses a certain cash-credit mix $s^* \in (0,1)$, to a policy where cash is phased out and the agent is forced to finance all consumption using credit, i.e. using $s = 0$. Our objective is to quantify the welfare cost for agents who are forced to move from their optimal policy s^* to the mandated one where the cash share is zero ($s = 0$). For the case of zero (i.e. small) inflation and low discount rates this welfare cost has an accurate, and extremely simple, analytic approximation which depends *only* on the optimal cash share s^* and on the cost of credit γ.

Let us use the flow value function $rV(m)$ to measure the agent's minimized flow cost to finance a total annual expenditure (e), using both cash and credit. Simple analysis of the closed form expression for $V(m)$ shows that for $\pi = 0$ and $r \to 0$ gives[11]

$$v^* \equiv \lim_{\pi=0, r\to 0} rV(m) = e\frac{R}{p}\ \log\left(1 + \frac{\gamma p}{R}\right) \quad . \tag{11}$$

Next, let v_0 be the flow cost of financing the expenditure stream e using credit only, so that $s = 0$. It is straightforward to see that $v_0 = \gamma e$. Thus, we define the cost of implementing Rogoff's ban on cash as $\ell = (v_0 - v^*)/e$, where the normalization by e allows us to read ℓ in units of yearly consumption. Using Proposition 4 to write the normalized cost of credit $\theta \equiv \gamma p/R$ as a function of the cash share $s = \theta/(1+\theta)$, we obtain

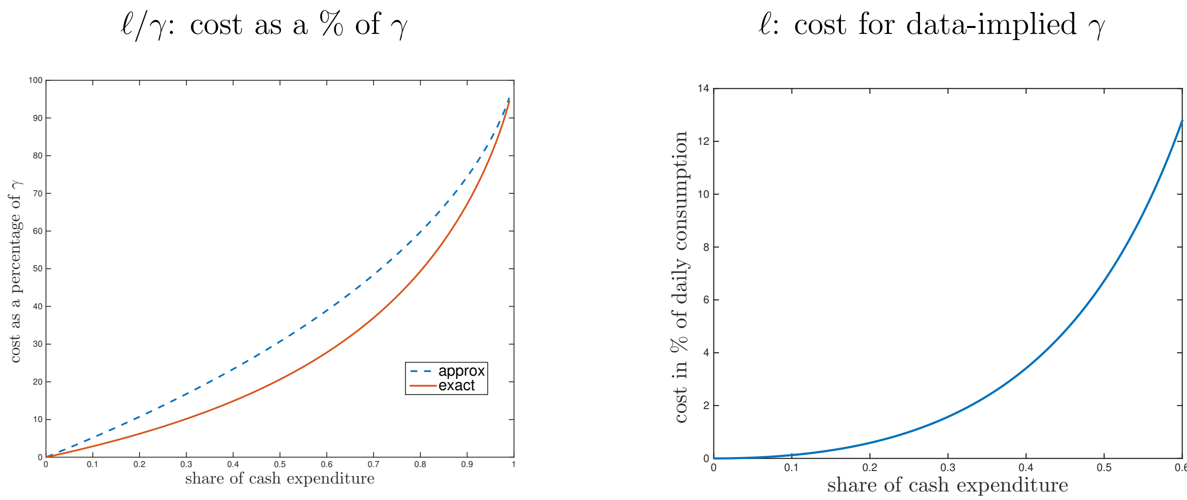
$$\ell(s,\gamma) \equiv \frac{v^* - v_0}{e} = \gamma\left(1 + \frac{1-s}{s}\log\left(1-s\right)\right) \quad . \tag{12}$$

Notice that this is a flow cost expressed as a fraction of the per-period (e.g. per year) consumption. The cost of implementing a zero-cash policy depends *only* on two parameters: the cost of credit γ and the share of cash purchases s. Simple analysis of $\ell(s,\gamma)$ shows that it equals zero at $s = 0$ and it is monotone increasing for $s > 0$. As credit becomes more convenient than cash (lower θ), the cash share s falls and the cost of moving to a no-cash system ($s = 0$) decreases. Naturally, the cost ℓ is scaled by the cost of credit, γ.

The left panel of Figure 2 plots the cost of implementing a zero cash policy relative to the cost of credit, namely ℓ/γ. The dashed line is the approximation given in equation (12), which is only a function of the cash share s, the solid line is the exact cost computed assuming inflation equal to 2 percent and a time discount $r = 0.02$. The convexity of the ℓ function implies that at intermediate values of s the cost (per year) of the zero-cash restriction amounts to a small fraction of the cost of credit use. For a country like Germany, where $s \approx 0.5$, the cost is approximately 30% of the yearly credit cost, γe. For the US, where $s \approx 0.25$, the value is around 15% of the yearly credit cost.

[11] See Appendix B for a closed form solution for this value function.

Figure 2: Flow cost of imposing zero-cash restriction



The "exact" cost function in the left panel assumes inflation equal to 2 percent and a time discount $r = 0.02$. The right panel uses, for each value of s on the x-axis, the corresponding value of γ implied by Proposition 4 under the assumption that $n = 50$ and $R = 0.02$.

Next we quantify the cost of credit, γ, using the relation between s and $\gamma p/R$ from Proposition 4 and that $p = n$. Assuming a nominal interest rate of 2 percent (the opportunity cost of cash R), and $n \approx 50$ per year as measured in the US and Germany, the model gives that $\gamma = \frac{sR}{(1-s)n}$. Using this equation shows the cost of credit γ to be a tiny number, about 4 basis point for $s = 0.5$ and 1.3 basis point for $s = 0.25$. The right panel of Figure 2 plots the implied cost per year of imposing the zero cash restriction assuming $n = 50, R = 0.02$ and computing the cost of credit γ implied by each level of s. As suggested by the estimates discussed before, the cost of implementing Rogoff's restriction appears tiny over the range of cash share values observed in the data: for a household with an annual consumption of 40K the cost is approximately 10 dollars per year in Germany, where the cash share is around 50%, it is about 2 dollars in the US where the cash share is around 25%.

Two remarks are useful to put the above figures in perspective. First, the magnitude of the cost scales proportionally with the value of γ as equation (12) shows. It is thus useful to note that the small estimates for γ discussed above, in the order of a few basis points, are likely affected by our specific modeling assumptions, in particular the assumption that all cash withdrawals are free under the cash-burning policy. Modifying this assumption, as done in the extension of Section 6, will likely increase the cost of credit. For arbitrary values of the cost of credit the left panel of Figure 2 can be used to gauge the cost of the zero-cash policy at any given level of the cash share. Second, our estimated costs of phasing out cash is borne by households who already possess the credit technology. A more encompassing measure of the social cost would also include the costs borne by the currently unbanked (about 7 percent of households in the US, see FDIC (2015)).

6 An extension: random variation in fixed cost

The model of Section 3 is very stark in that credit is used at $m = 0$ provided that the fixed transaction costs is sufficiently high ($b > \underline{b}$) in which case we have that $\underline{M} = M$ and $n = p$. Instead, if the transaction cost is sufficiently low ($b < \underline{b}$) credit is never used, i.e. $s = 1$ and the model becomes the Baumol-Tobin with random free withdrawal opportunities discussed in Alvarez and Lippi (2009). The prediction that the cash share is either zero or 1 seem too stark against the data. In particular, there is substantive evidence, based on micro data, that the amount of cash at the time of withdrawals is smaller than the average cash balance, i.e. $\underline{M} < M$. We show in this section that allowing the fixed cost b to be random and persistent allows to account for this fact while retaining the other features of the model. The variation in b implies that agent follows a cash-burning policy when she faces a high cash withdrawal cost, while she follows a Baumol-Tobin policy when the cost of cash withdrawals is below a critical threshold. The analysis can equivalently be conducted assuming the cost of credit to be random.

In particular we assume that there is a Poisson process with constant intensity λ whose occurrence indicates that a new value of b has been drawn. This Poisson process is independent of the one for the arrival of the free adjustment opportunities. The new values for the cost $\tilde{b}$ are drawn from the cumulative distribution function $F : \mathbb{R} \to [0, 1]$. Conditional on a change in the value of the fixed cost, the new value $\tilde{b}$ is assumed to be independent of the current value b. In this case the value function has two arguments, (m, b), so we write it $V : \mathbb{R}^2_+ \to \mathbb{R}_+$. We denote by primes the derivative of V with respect to m. The value function solves the following functional equation:

$$0 = \min \Bigg\{ \min_{0 \le c \le e} R\, m + \gamma[e - c] + p \min_{z \ge 0} [V(z, b) - V(m, b)] + \lambda \left[\int V\left(m, \tilde{b}\right) dF\left(\tilde{b}\right) - V(m, b) \right] \\ - V'(m, b)(c + \pi m) - r\, V(m, b) \quad , \quad b + \min_{z \ge 0} V(z, b) - V(m, b) \Bigg\} \quad \forall\ (m, b) \in \mathbb{R}^2_+ \qquad (13)$$

The interpretation of the terms in this functional equation is analogue to the one in equation (4): the outer min operator compares the value of using credit with the value of paying the fixed cost and replenishing cash balances. There are two differences. First, as mentioned above, b is also part of the state. Second, in the first term there is an extra expression given by the contribution to the expected change of the value function due to the change in the cost from b to a value drawn from the distribution with c.d.f. F.

In what follows we proceed, based on the analysis of the special cases analyzed in the previous section, by constructing a solution for a type of cash-burning policy which combines the two cases analyzed above.

DEFINITION 3. A threshold cash-burning policy is defined by a cost threshold $\underline{b}$ and a cash-target function m^*. For all $m > 0$ and all $b \geq 0$ the agent uses only cash. If $m = 0$ the agent withdraws cash when $0 \leq b \leq \underline{b}$, and uses credit when $b > \underline{b}$. For all (m, b) cash balances are set equal to $m^*(b)$ every time that a free adjustment opportunity arrives. Additionally, cash balances are set to $m^*(b)$ if $m = 0$ and $b \leq \underline{b}$.

Hence $\underline{b}$ is the critical threshold so that at $m = 0$ the agent uses credit if $b < \underline{b}$ and uses cash otherwise. Appendix C gives a detailed characterization of the value function for this problem under a threshold cash-burning policy. We use this extended model in a calibration that illustrates how it can produce cash management behavior featuring both cash and credit usage and where the amount of cash at the time of a withdrawal is smaller than the average cash holdings, i.e. $M < \underline{M}$. As mentioned, the latter feature is seen in the data but is not produced by the model with a constant cost of withdrawal studied in Section 3.

A quantitative assessment. As a benchmark, we use the observable statistics for Italian households which were summarized in Table 1. We focus on households that own an ATM card, a group for which the cash-credit margin is feasible.[12] For this group the data shows that share of cash expenditures is close to 50 percent of total (non-durable) expenditures and that the average currency holdings is about 6 days of expenditures. Moreover, it shows that the amount of cash at the time of withdrawal relative to the average money balances $\underline{M}/M$ is around 0.4, that the ratio of the average withdrawal to the average money balances is about W/M is 1.3 and that the households with ATM withdraw cash about 50 times per year.

Table 2 reports the comparable moments predicted by 3 cash management models. For comparison, all three models are calibrated to reproduce the same value of average currency holdings ($M/e = 6$) and the same number of cash withdrawals ($n = 53$). Column [1] uses the cash-only model described in part (ii) of Proposition 4. This model corresponds to the best cash management strategy when the (fixed and deterministic) withdrawal costs are low ($b < \underline{b}$), so that credit is not used. As discussed in Alvarez and Lippi (2009) this model has essentially 2 independent parameters, p and $\frac{bp^2}{eR}$, which we use to target the mean level of cash holdings and the number of withdrawals (first two lines of the table). The predictions

[12] In Italy the vast majority of ATM cards also serve as debit cards.

Table 2: Selected moments on cash holding patterns: different models

		Model		
		[1] cash-only	[2] cash-credit	[3] mixed
Average Currency	M/e^d	6	6	6
Number of cash withdrawals	n	53	53	53
Withdrawal size	W/M	1.1	0.9	1.0
Cash at withdrawals	$\underline{M}/M$	0.6	1.0	0.8
Cash share of expenditures	s	1.0	0.8	0.9

Cash balances M/e^d are measured relative to total expenditures per day, $e^d = e/365$. The number of cash withdrawals is per year. The parameters for model [3] are chosen to broadly match frequency of withdrawals and size of money holdings. The low value of b is 1% percent of daily consumption, the high value is 30%. The rate at which b changes is $\lambda = 250$ times per year.

for the other moments can be seen as an over-identifying test of the model. The share of cash expenditure is equal to 1 in this model since credit is not used.

The second model in column [2] is the cash-credit model described in part (i) of Proposition 4. This model corresponds to the best cash management strategy when the (fixed and deterministic) withdrawal costs are sufficiently high ($b > \underline{b}$) so that both cash and credit are used. The model has only 2 independent parameters: p and $\gamma p/R$ which are used to target the sample moments for M/c and n. The cash share predicted by the model is 80% of expenditures, above the one observed in the data. Likewise, the model predicts a high level of cash at the time of withdrawals, namely $\underline{M}/M = 1$.

The third model, summarized in column [3], is the one where the cost of withdrawal b is random. We parametrize the model assuming that b can take either a low value (1% of daily expenditures) or a high value (30% of daily expenditures) with equal probability. In the former case the agent finds it optimal not to use credit when cash is exhausted (since $b < \underline{b}$), while in the second the agent uses credit at $m = 0$ (since $b > \underline{b}$) waiting for a free withdrawal opportunity or a change in b. We assume that the rate at which the withdrawal cost changes is 200 per year (on average every working day).[13] Intuitively, the behavior produced by this model is close to a weighted average of the behavior of the cash-only and the cash-credit model characterized in columns [1] and [2].[14] The model's is able to account for a cash share below 100% and for a smaller level of cash at the time of withdrawals, $\underline{M}/M < 1$.

[13] The other structural parameters are taken from, or are close to, the structural estimates in Alvarez and Lippi (2009): the number of free withdrawal per year is $p = 40$, the opportunity cost of cash is $R = 2\%$ (this includes the nominal interest rate and the probability of cash theft, as discussed in Appendix D).

[14] The option value motives which might cause the outcomes to differ from a weighted average of the two polar models are small when the value of λ is high.

The main point of this analysis is to illustrate the tractability of the model and its potential for empirical analysis. Further extensions could be introduced to improve e.g. the fit of the number of cash transactions per year. For instance, we could introduce unexpected large cash-purchases, as in Alvarez and Lippi (2013), which may increase the number of transactions without first-order effects on the other steady state statistics. We leave this exploration for future work.

7 Conclusions

We presented a model that combines the ingredients of the dynamic cash inventory problem with the ingredients of the cash-credit choice. The key novelty compared to the previous literature is that we allow agents to use either cash or credit at each moment. This natural and realistic assumption implies an optimal rule for credit usage by the agent which turns out to depend on the amount of cash at hand. We find this feature interesting because it makes contact with a body of recent evidence showing that the likelihood of using cash increases with the level of cash holdings, as documented in e.g. Arango, Huynh, and Sabetti (2011); Arango, Bouhdaoui, and Bounie (2012) and Huynh, Schmidt-Dengler, and Stix (2014) using diary data for Canada and Austria. We showed that, in spite of its simplicity, the model predictions' on credit usage, the size of cash withdrawals and the average cash holdings are aligned with the magnitudes observed in the data. We used our model to quantify the cost of phasing out cash, a policy endorsed by Rogoff (2016) to fight several cash-intensive illegal activities. We estimate that the households' cost of moving from their optimal cash-credit share to one where the cash share is zero is a small fraction of the daily consumption, for a household with a 40K the cost is approximately 10 dollars per year in Germany, where the cash share is around 50%, it is about 2 dollars in the US where the cash share is around 25%. This estimated costs assume agents already have access to both cash and credit. A more encompassing estimate of the total costs of phasing out cash must include an estimate of the cost of banking the unbanked households.

Our model abstracts from aspects of the cash credit choice that have been emphasized before: the size of purchases (e.g. Whitesell (1989)) and the acceptability of credit at the points of sale (as in e.g. Huynh, Schmidt-Dengler, and Stix (2014)). Future models might benefit by unifying those aspects into a single model and quantify the relative importance of each of these frictions by using the relevant micro data.

References

Alvarez, Fernando E. and Francesco Lippi. 2009. "Financial Innovation and the Transactions Demand for Cash." *Econometrica* 77 (2):363–402.

———. 2013. "The demand of liquid assets with uncertain lumpy expenditures." *Journal of Monetary Economics* 60:753–770.

Arango, Carlos, Yassine Bouhdaoui, and David Bounie. 2012. "Modeling the Share of Cash Payments in the Economy." Manuscript, Bank of Canada.

Arango, Carlos, Dylan Hogg, and Alyssa Lee. 2013. "Why Is Cash (Still) So Entrenched? Insights from from Canadian Shopping Diaries." *Contemporary Economic Policy* forthcoming.

Arango, Carlos, Kim Huynh, and Leonard Sabetti. 2011. "How Do You Pay? The Role of Incentives at the Point-of-Sale." Working Papers 11-23, Bank of Canada.

Bagnall, John, David Bounie, Kim P. Huynh, Anneke Kosse, Tobias Schmidt, Scott Schuh, and Helmut Stix. 2014. "Consumer cash usage: a cross-country comparison with payment diary survey data." Working Paper Series 1685, European Central Bank.

Bar-Ilan, A. 1990. "Overdrafts and the Demand for Money." *The American Economic Review* 80 (5):1201–1216.

Bouhdaoui, Yassine and David Bounie. 2012. "Modeling the Share of Cash Payments in the Economy: An Application to France." *International Journal of Central Banking* 8 (4):175–195.

FDIC . 2015. "FDIC National Survey of Unbanked and Underbanked Households ." Executive summary, Federal Deposit Insurance Corporation.

Freeman, Scott and Finn E. Kydland. 2000. "Monetary Aggregates and Output." *The American Economic Review* 90 (5):pp. 1125–1135.

Huynh, Kim, Philipp Schmidt-Dengler, and Helmut Stix. 2014. "Whenever and Wherever: The Role of Card Acceptance in the Transaction Demand for Money." Tech. rep., Free University of Berlin, Humboldt University of Berlin, University of Bonn, University of Mannheim, University of Munich.

Klee, Elizabeth. 2008. "How people pay: Evidence from grocery store data." *Journal of Monetary Economics* 55 (3):526–541.

Kosse, Anneke and David-Jan Jansen. 2012. "Choosing how to pay: the influence of home country habits." Dnb working papers, Netherlands Central Bank, Research Department.

Lucas, Jr, Robert E and Nancy L Stokey. 1987. "Money and Interest in a Cash-in-Advance Economy." *Econometrica* 55 (3):491–513.

Lucas, Robert E. Jr. 2000. "Inflation and Welfare." *Econometrica* 68 (2):247–274.

Mooslechner, Peter, Helmut Stix, and Karin 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 & the Economy* 2006 (Q2):111–134.

Prescott, Edward C. 1987. "A multiple means-of-payment model." *New Approaches to Monetary Economics* :42–51.

Rogoff, Kenneth S. 2016. *The Curse of Cash*. No. 10798 in Economics Books. Princeton University Press.

Sargent, Thomas J. 1987. *Dynamic Macroeconomic Theory*. Harvard University Press.

Sastry, A. S. Rama. 1970. "The Effect of Credit on Trasactions Demand for Cash." *Journal of Finance* 25 (4):777–781.

Stix, Helmut. 2004. "How do debit cards affect cash demand? Survey data evidence." *Empirica* 31:93–115.

Whitesell, William C. 1989. "The Demand for Currency versus Debitable Accounts: Note." *Journal of Money, Credit and Banking* 21 (2):246–251.

APPENDICES – FOR ONLINE PUBLICATION

Cash burns: An inventory model with a cash-credit choice

F. Alvarez and F. Lippi

April 3, 2017

A Some direct evidence on cash credit usage

There is now a growing body of evidence that used both diaries for means of payments simultaneously with statistics about cash management. We briefly mention here some of the contributions which connect with the effects highlighted by our paper.

Arango, Huynh, and Sabetti (2011) analyze the pattern of the means used for purchases using diaries where 2350 individuals in Canada are asked to record all purchases they made for three days, how they paid for them (cash, credit card, debit card, check, etc), what type of good they were, perceptions of the means of payments available on the POS, demographics such as family size, income, education, gender, information about the type of credit and debit card held, the amount of cash balances held at the beginning of the three day period, as well as other variables of interest. In particular they fit multinomial logit(s) to the means of payment chosen and they find that, controlling for other variables, the amount of cash a the beginning of the diary has a negative impact on the probability that credit or debit is used as means of payments, especially for purchases of small value. They state that "higher initial cash holdings leads to higher probability of paying with cash. The result is especially pronounced for transactions below 25 dollars. The probability of paying with cash for an individual carrying 150 dollars could be twice as large compared with that of someone with only 5 dollars. However, as transaction value increases the marginal cost of paying with cash goes up reducing the difference in probabilities between high and low cash holders." Kosse and Jansen (2012) also report significant positive effect of cash holdings at the beginning of the diary in a Tobit regression on whether a purchase was paid in cash. They use a diary with purchases for one day for 2200 individuals in the Netherlands containing similar information as the one used for Canadian consumers.

In Arango, Bouhdaoui, and Bounie (2012), the authors compare two simple statistical models of means of payments. Each model has one free parameter per individual. One model assumes that for each individual payments above a threshold size are made with cash and otherwise with other means of payments. They refer to this as the TS model, as it is meant to capture the cash-credit models in which, due to a fixed cost of "credit" transactions, cash is used only for small-size transactions. This threshold is estimated for each individual and a goodness of fit statistic is estimated. The second model assumes that payments are made with cash as long as cash is available, and otherwise they are paid with credit. They refer to this model as CH (for cash holdings) and it is meant to capture precisely what the model in our paper describes: cash "burns" in the hands of the household, and hence its availability determines whether it is used or not. Interestingly the authors report that "We find that the CH model outperforms the TS model, and does a good job replicating the distribution of cash shares in both Canada and France." While the CH rule is assumed by these authors, our model provides an explanation in terms of primitives (the cost of credit, the cost of cash withdrawals, etc) of why and when the behavior in the CH model is optimal.

In Huynh, Schmidt-Dengler, and Stix (2014), Table 1, the authors report evidence taken from diary data from both Austria and Canada to show that (i) currency holdings are larger than zero at the time of a cash withdrawal (between 1/3 and 1/2 of the mean or median currency holdings) and that (ii) agents predominantly use cash rather than credit when they have enough cash at hand. The latter fact is particular relevant for our paper: most consumers (81% in Austria, 65% in Canada) with the possibility to choose between cash and credit will use cash as long as they have enough of it at hand. It is exactly this choice that our theoretical model will focus on.

B Characterization of optimal policy in the dynamic stochastic model

We present a few lemmas to characterize the best cash-burning policy as well as the best BT-policy, without analyzing which of the two policies is better, see Appendix B.1 for the proofs. For a cash burning policy to be optimal, i.e. to solve the problem in equation (4), one needs to establish that

(i) It is optimal to pay with cash at $m \in (0, m^*]$ and with credit at $m = 0$.

(ii) The size of the withdrawal m^* is optimal.

(iii) At $m = 0$ it is optimal to wait for a free adjustment opportunity (instead of paying b to withdraw).

The first lemma establishes the existence and uniqueness of cash burning policies. Moreover, it also establishes that if such a policy is followed, the first order conditions for the use of cash is verified, i.e. Item i is satisfied.

Lemma 1. For each $m^* > 0$ there is a unique value function V for the cash burning policy:

$$V(m) = \left(pV(m^*) - \frac{eR}{r+p+\pi}\right)\frac{1}{r+p} + m\frac{R}{r+p+\pi} + A\left(1 + \frac{\pi}{e}m\right)^{-\frac{r+p}{\pi}} \quad (14)$$

for all $m \in [0\,,\, m^*]$, where

$$A = \frac{e}{r+p}\left[\gamma + \frac{R}{r+p+\pi}\right]\,. \quad (15)$$

The function V is strictly decreasing and convex in an interval $[0, \bar{m}]$ for some $\bar{m} > 0$. Thus the first order conditions for the optimal use of cash are verified, i.e.

$$-\gamma - V'(0) = 0 \quad \text{and} \quad -\gamma - V'(m) < 0 \quad \text{for} \quad m \in (0\,,\, m^*]\;. \tag{16}$$

The next lemma characterizes the optimal cash target m^* for a cash-burning policy. For a BT policy to be optimal, i.e. to solve the problem in equation (4), one needs to establish that

(a) It is never optimal to pay with credit.

(b) The size of the withdrawal m^* is optimal.

(c) At $m = 0$ it is optimal to pay b to withdraw.

LEMMA 2. Let V be the value of following a cash-burning policy. If m^* is chosen optimally, i.e. if

$$V'(m^*) = 0\,, \tag{17}$$

holds, then m^* is given by equation (9). Moreover m^*/e is increasing in π.

Inspection of equation (9) shows that the optimal cash replenishment level is increasing in γ/R, i.e. the cost of credit relative to the opportunity cost of cash. The next lemma characterizes the value of a Baumol-Tobin policy. In this case the value function and cash target m^* are constructed so that Item b is satisfied (this lemma is straight from parts of Propositions 2 and Proposition 3 in Alvarez and Lippi (2009) where there is no possibility of paying with credit).

LEMMA 3. For each $m^* > 0$ there is a unique value function V for the Baumol-Tobin policy. This has the functional form as in equation (14), as for the cash-burning policy, except that now A is given by:

$$A = \frac{e}{r+p}\left[R\,\frac{m^*}{e} + (r+p)b + \frac{R}{r+p+\pi}\right]\;. \tag{18}$$

The function V is strictly convex and decreasing in an interval $[0, \bar{m}]$ for some $\bar{m} > 0$. Moreover, if m^* is chosen optimally so that equation (17) holds, then m^* solves equation (10).

The next remark develops the expressions for the special case where $\pi = 0$, by using L'Hopital rule in all the relevant formulas:

Remark 1. If $\pi = 0$, then the optimal target m^* for a cash-burning policy is given by:

$$\frac{m^*}{e} = \frac{1}{r+p} \log\left(1 + \frac{\gamma(r+p)}{R}\right) \geq 0 \,. \tag{19}$$

and the optimal target m^* for a Baumol-Tobin policy cash-burning is given implicitly by:

$$\exp\left(\frac{m^*}{e}(r+p)\right) = 1 + \frac{m^*}{e}(r+p) \;+ (r+p)^2 \frac{b}{eR} \,. \tag{20}$$

The analysis of the expression for m^* in equation (19) for a cash-burning policy shows that m^*/e is decreasing in $r+p$ and decreasing in R. Moreover $m^*(r+p)$ is increasing in $r+p$. A first order expansion of m^*/e on γ/R evaluated at $(r+p) = 0$ or at $\gamma/R = 0$, gives $m^*/e \approx \gamma/R$ which is the expression for the deterministic steady state cash-credit model. The elasticity of m^*/e with respect to R is

$$0 \leq -\frac{\partial \log(m^*/e)}{\partial \log R} = \frac{\frac{\gamma(r+p)}{R}}{1+\frac{\gamma(r+p)}{R}} \;/\; \log\left(1 + \frac{\gamma(r+p)}{R}\right) \leq 1.$$

And so the elasticity is, in absolute value, a decreasing function of $\frac{\gamma(r+p)}{R}$.

In the case of the BT policy, the optimal return point m^* in equation (20) shows that $\frac{m^*}{e}$ is increasing in $\frac{b}{eR}$, $\frac{m^*}{e} = 0$ as $\frac{b}{eR} = 0$ and $\frac{m^*}{e} \to \infty$ as $\frac{b}{eR} \to \infty$. Moreover, for small $\frac{b}{eR}$, we can approximate $\frac{m^*}{e}$ by the the solution to the Baumol-Tobin model, or $\frac{m^*}{e} = \sqrt{2\,\frac{b}{eR}} + o\left(\sqrt{\frac{b}{eR}}\right)$. Finally, the interest rate elasticity of m^*/e is smaller than $1/2$ and it is decreasing in $(r+p)/R$.

B.1 Proofs for the dynamic stochastic model

Proof. (of Lemma 1.) Inserting equation (14) for an arbitrary A one readily verifies that it solves the o.d.e. in equation (5) for $m \in (0, m^*)$. The value of A is obtained by imposing equation (6). We can take $\bar{m}$ to be the minimum between m^* and the point where $V' = 0$. The existence of the initial decreasing and convex segment follows by inspection. $V'(0) = -\gamma$ is obtained by differentiating equation (14) using equation (15) for A and evaluating at $m = 0$. Using the convexity and the value of $V'(0)$ the f.o.c. for c are directly verified for all $m \in [0, m^*]$. □

Proof. (of Lemma 2.) Solving for $V'(m^*) = 0$ using equation (14) and equation (15). That

$m^*/e \leq \gamma/R$ is equivalent to:

$$\left(1+(r+p+\pi)\frac{\gamma}{R}\right)^{\frac{1}{\pi+r+p}} \leq \left(1+\pi\,\frac{\gamma}{R}\right)^{\frac{1}{\pi}}$$

which is equivalent to $(1+x)^{1/x}$ is decreasing in x for $x \geq 0$. To simplify the notation of the derivative, consider without loss of generality that $r=0$ and that $R=1$ so that:

$$\frac{\partial m^*/e}{\partial \pi} = \frac{(1+(p+\pi)\gamma)^{\frac{\pi}{\pi+p}}\left[\frac{\pi p}{(p+\pi)^2}\log\left(1+(p+\pi)\gamma\right)+\frac{\pi}{\pi+p}\frac{\gamma\pi}{1+(p+\pi)\gamma}-1\right]+1}{\pi^2}$$

We first show that this derivative is strictly positive at $\pi=0$ provided that $p+r>0$ and $\gamma/R>0$. For $\pi=0$ we can write:

$$\begin{aligned}&(1+(p+\pi)\gamma)^{\frac{\pi}{\pi+p}} = \\ &1+\frac{\log\left(1+\gamma p\right)}{p}\pi+\frac{1}{2}\left[\left(\frac{\log\left(1+\gamma p\right)}{p}\right)^2-\frac{2\log\left(1+\gamma p\right)}{p^2}+2\frac{\gamma}{p(1+p\gamma)}\right]\pi^2+o(\pi^2)\end{aligned}$$

and

$$\frac{\pi p}{(p+\pi)^2}\log\left(1+(p+\pi)\gamma\right) \quad = \quad \frac{\pi}{p}\log\left(1+p\gamma\right)+\pi^2\left(\frac{\gamma}{p\left(1+\gamma p\right)}-\frac{2}{p^2}\log\left(1+p\gamma\right)\right)$$

Replacing this into the expression for $\partial(m^*/e)/\partial\pi$ and taking the limit as $\pi \downarrow 0$ we get:

$$\begin{aligned}&\frac{\partial m^*/e}{\partial \pi} = \left(\frac{\gamma}{p\left(1+\gamma p\right)}-\frac{2}{p^2}\log\left(1+p\gamma\right)\right) \\ &\frac{1}{p}\frac{\gamma}{1+p\gamma}-\frac{1}{2}\left[\left(\frac{\log\left(1+\gamma p\right)}{p}\right)^2-\frac{2\log\left(1+\gamma p\right)}{p^2}+2\frac{\gamma}{p(1+p\gamma)}\right]+\left(\frac{\log\left(1+\gamma p\right)}{p}\right)^2 \\ = \quad &\frac{1}{p^2}\left[\frac{p\gamma}{(1+\gamma p)}-\log\left(1+\gamma p\right)+\frac{1}{2}\left(\log\left(1+\gamma p\right)\right)^2\right] \geq 0 \ ,\end{aligned}$$

so this derivative is positive when $\pi>0$. To finish the proof we let $f(\pi,\gamma,p) \equiv \partial(m^*/e)/\partial\pi$. Note that $f(\pi,\gamma,p)\to 0$ as $\gamma \to 0$ and for $\pi>0$. Since $f(\pi,\gamma,p)$ is increasing in γ, so that $f(\pi,\gamma,p)>0$ for $p>0$, $\pi>0$, and $\gamma>0$. Thus we have that, in general, $\partial(m^*/e)/\partial\pi>0$ for $p+r>0$, $\pi>0$, and $\gamma/R>0$. □

Proof. (of Proposition 2.) The value of $\underline{b}$ equates $V(0)+\underline{b}=V(m^*)$ so the agent with $m=0$ is indifferent between waiting for a free withdrawal while paying with credit and incurring the fixed cost and adjusting. It remains to show that when $b \leq \underline{b}$ it is optimal to use cash for all $m \in [0,m^*]$. For this take the limit as $m \downarrow 0$ on the o.d.e. given by equation (5) obtaining:

$-V'(0)e = (r+p)V(0) - pV(m^*)$ where V and m^* are the value function and target cash from the Baumol-Tobin policy. Using the boundary condition equation (7), the definition of $\underline{b}$ and the boundary condition for the cash-burning policy equation (6), we get:

$$V(0) = V(m^*) + b \leq \frac{\gamma e + pV(m^*)}{r+p}$$

Thus

$$-V'(0)e = (r+p)V(0) - pV(m^*) \leq (r+p)\frac{\gamma e + pV(m^*)}{r+p} - pV(m^*) = \gamma e \ ,$$

and thus $-V'(0) \leq \gamma$. Hence, using the convexity of V established in Lemma 3, we have:

$$-V'(m) - \gamma \leq 0 \text{ for all } m \in [0\,,\ m^*]$$

and hence it is optimal to use cash for all $m \in [0\,,\ m^*]$. □

Proof. (of Proposition 3) Note that $\underline{b} \geq 0$ is equivalent to $m^*/e \leq \gamma/R$. The homogeneity of degree one w.r.t. (γ, R) follows from the homogeneity of degree one of the expression for $\underline{b}$ and the homogeneity of degree zero of m^*/e. The rest of the expressions follow directly from computations, and using L'Hopital on the relevant places. □

Proof. (of Proposition 4). That $n = p$ is an immediate implication of the cash-burning policy. Now we derive the expression for s. Since agents use credit only when they ran out of cash and until they get a free withdrawal opportunity, we characterize the average value of the cash share by computing two expected times. One is the expected time that an agent that has just hit $m = 0$ will keep zero cash, which is simply $1/p$. The other, is the expected time that an agent that has just hit cash balances m^* will take to first hit cash balances $m = 0$. We denote this quantity as $\tau(m^*)$. For the case of zero inflation $\pi = 0$, which we focus throughout, this equals

$$p\tau(m) \quad = \quad 1 - \tau'(m)e + p\tau(m^*), \quad \text{and } \tau(0) = 0,$$

which has solution: $\tau(m) = \frac{1}{p}\exp\left(p\ \frac{m^*}{e}\right)\left(1 - \exp\left(-p\ \frac{m}{e}\right)\right)$. Evaluated at m^*:

$$\tau(m^*) \quad = \quad \frac{1}{p}\left(\exp\left(p\ \frac{m^*}{e}\right) - 1\right),$$

The expected time $\tau(m^*)$ is increasing in m^*/e and in p. This leads to:

$$s \equiv \frac{\tau(m^*)}{1/p + \tau(m^*)} = 1 - \frac{1}{\exp\left(p\ \frac{m^*}{e}\right)} \ .$$

Replacing the expression for m^* we have:

$$s = 1 - \left(1 + \frac{\gamma(r+p)}{R}\right)^{-\frac{p}{r+p}} \in (0,1) \ ,$$

Now we turn to the average withdrawal size W. To write an expression for W we split the withdrawals between those occurring when $m = 0$, which happens a fraction $(1-s)$ of the time, and those occurring when $m = 0$, which happens a fraction s of the time. This gives

$$W = (1-s) \cdot m^* + s \int_0^{m^*} (m^* - m) h(m) dm \ ,$$

where $h(\cdot)$ is the density of the invariant distribution of cash holdings conditional on $m > 0$, which satisfies

$$h(m) = h'(m)\ \frac{e}{p} \quad , \quad 1 = \int_0^{m^*} h(m) dm \ .$$

W is obtained from the accounting identity $Wn = c$ with $n = p$ and $c = s\,e$:

$$W = \frac{e}{p}\ s = \frac{e}{p}\left(1 - \left(1 + \frac{\gamma(r+p)}{R}\right)^{-\frac{p}{r+p}}\right)$$

The average cash holdings are defined as:

$$M = (1-s) \cdot 0 + s \int_0^{m^*} m h(m) dm$$

The formula for M takes into account that the distribution of cash holdings has a mass point at $m = 0$ of size $(1-s)$. Combining the expression for M with the expression for W gives $M = m^* - W$, which replacing the expressions for m^* and W yields the following money demand:

$$\frac{M}{e} = \frac{m^*}{e} - \frac{s}{p} = \frac{1}{r+p} \log\left(1 + \frac{\gamma(r+p)}{R}\right) - \frac{1}{p}\left(1 - \left(1 + \frac{\gamma(r+p)}{R}\right)^{-\frac{p}{r+p}}\right)$$

When $r \to 0$ the expression simplifies to:

$$\frac{M}{e} = \frac{1}{p}\left[\log\left(1 + \frac{\gamma\ p}{R}\right) - 1 + \left(1 + \frac{\gamma\ p}{R}\right)^{-1}\right]$$

The ratio W/M equals $s/(p\frac{m^*}{e} - s)$, which in terms of parameters gives:

$$\frac{W}{M} = \left[\frac{p}{p+r} \, \log\left(1 + \frac{\gamma(r+p)}{R}\right) \left(1 - \left(1 + \frac{\gamma(r+p)}{R}\right)^{-\frac{p}{r+p}}\right)^{-1} - 1 \right]^{-1} .$$

Also as $r \to 0$, the expression simplifies further to

$$\frac{W}{M} = \frac{1}{\left(1 + \frac{R}{\gamma\, p}\right) \log\left(1 + \frac{\gamma\, p}{R}\right) - 1} ,$$

The following two identities hold regardless of the average cash share s, and including the extreme cases)

$$W + \underline{M} = m^* \;\; \text{and} \;\; \underline{M} = \frac{p}{n} M ,$$

but that since $n = p$ then $\underline{M} = M$. $\square$

C Solution for the problem with random withdrawal cost

We state without proof the straightforward, yet useful result, that V must be weakly increasing in b. In particular:

Lemma 4. For all $m \geq 0$, if $b \geq \tilde{b} \geq 0$, then $V(m, b) \geq V\left(m, \tilde{b}\right)$.

Moreover we state the following lemma:

Lemma 5. Let V be the value of an optimal threshold cash-burning policy. Define V^* as

$$V^*(b) \;\; = \;\; \min_{z \geq 0} V(z, b) \;\; \text{and} \;\; \bar{V}^* = \int_0^\infty V^*(b) \, dF(b) \, , \tag{21}$$

for each $b \geq 0$. The value function V satisfies the following:

$$\begin{aligned} V(m, b) = & \left(p\bar{V}^* - \frac{eR}{r+p+\pi}\right) \frac{1}{r+p} + m \frac{R}{r+p+\pi} + \bar{A} \left(1 + \frac{\pi}{e} m\right)^{-\frac{r+p}{\pi}} \\ & + \frac{p}{r+\lambda+p} \left(V^*(b) - \bar{V}^*\right) + A(b) \left(1 + \frac{\pi}{e} m\right)^{-\frac{r+p+\lambda}{\pi}} \;\; \text{for all } (m, b) \in \mathbb{R}_+^2 \, , \end{aligned} \tag{22}$$

for a constant $\bar{A} \geq 0$ and a function $A(b) : \mathbb{R}_+ \to \mathbb{R}$ which are the unique solution of two equations, defined using the function V^* given in (21), the parameters $(r, \lambda, p, \pi, R, \gamma, e)$, and the function F (see Appendix C.1 for details).

Lemma 5 yields a recursion from a function $V^*(b)$ to another implied $V^*(b)$. In particular, given $V^*(b)$ one can define the corresponding value function $V(m,b)$ using the expression in Lemma 5, and use it to construct the implied minimized function $V^*(b)$. The fixed point of such mapping gives the solution for the value function of the best threshold cash-burning policy. The last issue to be established is that the threshold cash burning policy is optimal. We discuss this formally in Appendix C.3, where we provide conditions for the optimality of using credit at $m=0$ if $b<\underline{b}$ and the optimality of using cash when $m>0$.

Using $\bar{V}(0)=\int_0^\infty V(0,b)\,dF(b)$ to denote the expected value at $m=0$, we have

$$V(0,b)=\begin{cases} b+V^*(b) & \text{if } b\le \underline{b} \\ \frac{\gamma e+pV^*(b)+\lambda\bar{V}(0)}{r+p+\lambda} & \text{if } b\ge \underline{b} \end{cases} \tag{23}$$

which gives the condition for the optimality of using cash vs credit at $m=0$.

C.1 Proofs for the model with random b.

Proof. (of Lemma 5) The Bellman equation following a threshold cash-burning policy is:

$$(r+p+\lambda)\,V(m,b)\;=m\,R\,+\,pV^*(b)+\lambda\,\bar{V}(m)\,-\,[e+\pi m]\,V'(m,b)\quad\text{for all } m\ge 0$$

$$\bar{V}(m)\;\;=\;\;\int_0^\infty V(m,b)\,dF(b) \tag{24}$$

By integrating V w.r.t. b in equation (C.1) the function $\bar{V}$ solves the following ode on m:

$$(r+p)\,\bar{V}(m)\;=m\,R\,+\,p\bar{V}^*-\,[e+\pi m]\,\bar{V}'(m)\quad\text{for all } m\ge 0 \tag{25}$$

Note that the o.d.e. (25) does not depend on λ and has the same solution as the model with no variation on b for an arbitrary constant $\bar{A}$ given in equation (14) in Lemma 1. So we have:

$$\bar{V}\,(m)=\left(p\bar{V}^*-\frac{eR}{r+p+\pi}\right)\frac{1}{r+p}+m\frac{R}{r+p+\pi}+\;\bar{A}\,\left(1+\frac{\pi}{e}m\right)^{-\frac{r+p}{\pi}}$$

The function

$$V_p(m,b)\equiv\bar{V}(m)+p(V^*(b)-\bar{V}^*)/(r+\lambda+p)$$

is a particular solution of the o.d.e. in equation (C.1). Note that it depends on two constants $(\bar{A},\bar{V}^*)$ as well as the function $V^*(b)$. We also have that $V_h(m,b)$ is a solution of the homogenous equation:

$$V_h(m,b) = A(b)\,\left(1+\frac{\pi}{e}m\right)^{-\frac{r+p+\lambda}{\pi}}$$

where we let A be a function of b. Note that we have for all $m \geq 0$, integrating $V_p(m,b) + V_h(m,b)$ with respect to F we must obtain $\bar{V}(m)$. So

$$\bar{V}(m) = \bar{V}(m) \,+\, \frac{p}{r+\lambda+p}\int (V^*(b) - \bar{V}^*) dF(b) \,+\, \left(1+\frac{\pi}{e}m\right)^{-\frac{r+p+\lambda}{\pi}} \int A(b)\, dF(b)$$

and thus, using the definition of $\bar{V}^*$ we have that A must satisfy:

$$\int A(b)\, dF(b) = 0\ .$$

Summarizing, the solution of equation (C.1) is given by $V(m,b) = V_p(m,b) + V_h(m,b)$, which depend on the functions $V^*(b)$, $A(b)$ and the constant $\bar{A}$ and $\bar{V}^*$. We then have:

$$\begin{aligned} V(m,b) \;&=\; \left(p\bar{V}^* - \frac{eR}{r+p+\pi}\right)\frac{1}{r+p} + m\frac{R}{r+p+\pi} + \bar{A}\,\left(1+\frac{\pi}{e}m\right)^{-\frac{r+p}{\pi}} \\ &+\; \frac{p}{r+\lambda+p}\left(V^*(b) - \bar{V}^*\right) + A(b)\,\left(1+\frac{\pi}{e}m\right)^{-\frac{r+p+\lambda}{\pi}} \end{aligned}$$

Now we turn to the behavior at $m = 0$. Note from equation (23) that at $b = \underline{b}$ we have:

$$(r+p+\lambda)\,\underline{b} = \gamma e - (\lambda + r)V^*\left(\underline{b}\right) + \lambda\bar{V}(0)\ .$$

Since V^* is a (weakly) increasing function of b, then there exist a unique solution to $\underline{b}$, although it may be 0 or $+\infty$. We can restate these properties using e form of the value function derived above:

$$A(b) \geq V^*(b)\left(1 - \frac{p}{p+r+\lambda}\right) + b - \frac{p\,\lambda\,\bar{V}^*}{(p+r+\lambda)(r+p)} + \frac{e\,R}{(r+p+\lambda)(r+p)} - \bar{A}$$

with $=$ for $b \leq \underline{b}$. Also, using that $V(0,b) = \bar{V}(0) + p/(p+r+\lambda)(V^*(b) - \bar{V}^*)$,

$$A(b) \geq \frac{\gamma e}{p+r+\lambda} - \left(1 - \frac{\lambda}{p+r+\lambda}\right)\bar{V}(0) + \frac{p\,\bar{V}^*}{(p+r+\lambda)}$$

with equality for $b \geq \underline{b}$. Note that for $b \geq \underline{b}$, the function $A(b)$ does not depend on b. We can also write this condition, by using the form of $\bar{V}$ as:

$$\begin{aligned} A(b) &\geq \frac{\gamma e}{p+r+\lambda} - \frac{p+r}{r+p+\lambda}\left[\left(p\bar{V}^* - \frac{eR}{r+p+\pi}\right)\frac{1}{r+p} + \bar{A}\right] + \frac{p\,\bar{V}^*}{p+r+\lambda} \\ &= \frac{\gamma e}{p+r+\lambda} + \frac{eR}{(r+p+\pi)(r+p+\lambda)} - \frac{p+r}{r+p+\lambda}\bar{A} \end{aligned}$$

with equality for $b \geq \underline{b}$. Equating the expressions for $A(b)$ at $b = \underline{b}$ we have:

$$V^*(\underline{b})(r+\lambda) + \underline{b}\,(p+r+\lambda) - \frac{p\,\lambda\,\bar{V}^*}{(r+p)} + \frac{e\,R}{(r+p)} = \gamma e + \frac{eR}{(r+p+\pi)} + \lambda\,\bar{A} \tag{26}$$

Using the form of $A(b)$ and that it integrates to zero, we obtain one more equation:

$$\begin{aligned} 0 &= \int_0^{\underline{b}} \left[V^*(b)\frac{r+\lambda}{p+r+\lambda} + b\right] dF(b) + F(\underline{b})\left[-\frac{p\,\lambda\,\bar{V}^*}{(p+r+\lambda)(r+p)} + \frac{e\,R}{(r+p+\lambda)(r+p)} - \bar{A}\right] \\ &+ (1 - F(\underline{b}))\left[\frac{\gamma e}{p+r+\lambda} + \frac{eR}{(r+p+\pi)(r+p+\lambda)} - \frac{p+r}{r+p+\lambda}\bar{A}\right] \end{aligned} \tag{27}$$

Note that given the function V^* we can regard equation (26) and equation (27) as a system of two equations in two unknowns $(\bar{A}, \underline{b})$. We can write equation (27) as:

$$\begin{aligned} (p+r)\bar{A} &= \int_0^{\underline{b}} \left[V^*(b)(r+\lambda) + b(p+r+\lambda)\right] dF(b) + F(\underline{b})\left[-\frac{p\,\lambda\,\bar{V}^*}{(r+p)} + \frac{e\,R}{(r+p)} - \lambda\bar{A}\right] \\ &+ (1 - F(\underline{b}))\left[\gamma e + \frac{eR}{(r+p+\pi)}\right] \end{aligned}$$

and replacing $\lambda\bar{A}$ from equation (26) we write an equation for $\bar{A}$:

$$\begin{aligned} (p+r)\bar{A} &= \int_0^{\underline{b}} \left[V^*(b)(r+\lambda) + b(p+r+\lambda)\right] dF(b) - F(\underline{b})\left[V^*(\underline{b})(r+\lambda) + \underline{b}(p+r+\lambda)\right] \\ &+ \gamma e + \frac{eR}{(r+p+\pi)} \end{aligned}$$

or the following equation which gives a unique solution for $\underline{b}$:

$$
\begin{aligned}
(p+r)&\left[V^*(\underline{b})(r+\lambda)+\underline{b}\,(p+r+\lambda)-\frac{p\,\lambda\,\bar{V}^*}{(r+p)}+\frac{e\,R}{(r+p)}-\gamma e-\frac{eR}{(r+p+\pi)}\right] \\
&=\lambda\int_0^{\underline{b}}\left[V^*(b)(r+\lambda)+b(p+r+\lambda)\right]dF(b)-\lambda F(\underline{b})\left[V^*(\underline{b})(r+\lambda)+\underline{b}(p+r+\lambda)\right] \\
&\quad+\lambda\gamma e+\lambda\frac{eR}{(r+p+\pi)}
\end{aligned}
$$

□

C.2 Analytic expressions for $\bar{A}$, $\underline{b}$ and $A(\cdot)$.

The threshold $\underline{b}$ is the unique solution of:

$$
\begin{aligned}
&(p+r)\left[V^*(\underline{b})(r+\lambda)+\underline{b}\,(p+r+\lambda)\right]-p\,\lambda\,\bar{V}^*+e\,R-(p+r+\lambda)\gamma e-(p+r+\lambda)\frac{eR}{(r+p+\pi)} \\
&=\lambda\int_0^{\underline{b}}\left[V^*(b)(r+\lambda)+b(p+r+\lambda)\right]dF(b)-\lambda F(\underline{b})\left[V^*(\underline{b})(r+\lambda)+\underline{b}(p+r+\lambda)\right]
\end{aligned}
$$

The constant A is defined using the threshold $\underline{b}$ by

$$
\bar{A}=V^*(\underline{b})\left(\frac{r+\lambda}{\lambda}\right)+\underline{b}\left(\frac{p+r+\lambda}{\lambda}\right)-\frac{p\,\bar{V}^*}{(r+p)}+\frac{e\,R}{\lambda(r+p)}-\frac{\gamma e}{\lambda}-\frac{eR}{\lambda(r+p+\pi)}
$$

The function A is defined using $\bar{A}$ and the threshold $\underline{b}$ by:

$$
\begin{aligned}
A(b)&=\frac{\gamma e}{p+r+\lambda}-\frac{p+r}{r+p+\lambda}\left[\left(p\bar{V}^*-\frac{eR}{r+p+\pi}\right)\frac{1}{r+p}+\bar{A}\right]+\frac{p\,\bar{V}^*}{p+r+\lambda}\,,\ \text{ if } b\geq\underline{b}\,, \\
A(b)&=V^*(b)\frac{r+\lambda}{p+r+\lambda}+b-\frac{p\,\lambda\,\bar{V}^*}{(p+r+\lambda)(r+p)}+\frac{e\,R}{(r+p+\lambda)(r+p)}-\bar{A}\,,\ \text{ if } b\leq\underline{b}\,.
\end{aligned}
$$

C.3 On the optimality of using cash when $m>0$

Next we show that under the proposed policy we have the following property for the derivative of V at $m=0$.

$$
V'(0,b)=-\gamma \quad \text{if } \ b\geq\underline{b} \quad , \quad V'(0,b)\geq-\gamma \quad \text{if } \ b\leq\underline{b}
$$

We use this property to establish the optimality of using cash for $m>0$ and credit at $m=0$ and $b\geq\underline{b}$. To see why this property has to hold take the limit as $m\downarrow 0$ on the ode given by

(C.1) together with equation (23) to obtain:

$$e\,V'(0,b) = \begin{cases} pV^*(b) + \lambda\,\bar{V}(0) \ - (r+p+\lambda)\,V(0,b) = -\gamma e & \text{if } b \geq \underline{b} \\ pV^*(b) + \lambda\,\bar{V}(0) \ - (r+p+\lambda)\,V(0,b) \geq -\gamma e & \text{if } b \leq \underline{b}\ . \end{cases}$$

It remains to be shown that it is optimal to use cash for $m > 0$. A sufficient condition for this is that $V(\cdot, b)$ has a convex and decreasing segment starting at $m = 0$.

D A foundation for the inventory problem

In this appendix we formulate a more basic problem which gives rise to the inventory problem of Section 3. This allows to better interpret the parameters of the model, especially p and R, as well as to clarify why cash burns in the model of Section 3. Recall that the model of that section *assumes* that households that hold cash balances face an opportunity cost R per unit of time.

As in the previous sections we assume that the agent has a constant consumption flow equal to $e > 0$ per unit of time. This consumption expenditure can be paid for using a cash flow c or a credit (or debit) card. Differently from the previous sections, we allow two sources or randomness, each described by an independent Poisson process with intensities p_1 and p_2. The first describes the arrivals of "free adjustment opportunities" for cash balances. The second describes the arrivals of times where the agent's wallet is stolen, so that neither her cash balances nor her credit-debit can be used to purchase consumption. As in the previous section we assume that a fixed cost b is paid for each adjustment unless it happens at the time of a free adjustment opportunity. We assume that in the event in which the agent's cash is stolen, it must necessarily pay b, at which time it can adjust her cash with no further cost.

The problem of the agent is:

$$G\,(m) = \min_{\{m(t),c(t),\tau_j\}} \mathbb{E}_0 \left\{ \sum_{j=0}^{\infty} \exp\left(-r\,\tau_j\right) \left[I_{\tau_j}\ b + \left(m\left(\tau_j^+\right) - m\left(\tau_j^-\right)\right) \right] \right. \\ \left. + \int_0^{\infty} \exp\left(-r\,t\right)\left(e - c(t)\right)(1+\gamma)dt \right\} \tag{28}$$

subject to $dm = -(c + \pi m)dt - m\,dN_2$, where $m\,(t) \geq 0$, where N_2 is the Poisson counter of times where the cash was lost or stolen up to t, and where $m\,(0) = m$ is given. In this equation τ_j denote the stopping times at which an adjustment of m takes place (a cash withdrawal), $c(t) \in [0, e]$ denotes the fraction of purchases made in cash, and $1 + \gamma$ the cost of buying one unit of goods using credit. The indicator I_{τ_j} is zero −so the cost is not paid− if the adjustment occurs upon a free opportunity, otherwise it is one. The expectation is taken with respect to the two Poisson processes. The parameters of this problem are $r, \pi, p_1, p_2, b, \gamma$ and e. Importantly, R is *not* a parameter of this problem.

Note the differences between this problem and the one analyzed in the previous section. In this problem we write explicitly the cost incurred by the agent, i.e. the amount withdrawn at each adjustment and the fixed costs. Instead in the problem of Section 3 we assume that the agent has an opportunity cost R for each cash holdings, and at the time of a withdrawal we only include explicitly the fixed cost paid.

D.1 Optimal policy

We now write the Bellman equation and characterize the optimal policy. We will show that, provided that b is above some critical value, the optimal policy involves using no credit as long as a positive amount of cash is held, namely that $c(t) = e$ if $m > 0$, and that credit is used only once $m = 0$, in which case $c(t) = 0$. Thus, the agent follows a cash-burning policy. Assuming that for $m \in [0, m^*]$ the value function G is differentiable it must solve the following functional equation:

$$
\begin{aligned}
0 \;=\; \min \Big\{ & \min_{0 \le c \le e} (1+\gamma)[e-c] + p_1 \min_{z \ge 0} [z - m + G(z) - G(m)] \\
& + p_2 \min_{z \ge 0} [b + z + G(z) - G(m)] - G'(m)(c + \pi m) - r\, G(m) \quad , \\
& b + \min_{z \ge 0} [z - m + G(z) - G(m)] \Big\} \quad \text{for all} \;\; m \ge 0 \, .
\end{aligned}
$$

This equation accounts for the different choices the agent can make about using cash vs credit and deciding whether to replenish cash balances. In particular, the outer min operator compares the value of paying the fixed cost b to replenish cash balances (the last line) versus not doing it (the first two lines). The equation also considers the optimal choice of z, i.e. the replenishment level to choose conditional on adjusting, as well as the choice on c about whether to use cash or credit.

D.2 An equivalent "shadow-cost" representation

We now define a related problem that is closer to the standard inventory theoretical problem where the agent minimizes the shadow cost, as we used in the previous section.

$$V(m) = \min_{\{m(t),c(t),\tau_j\}} \quad \mathbb{E}_0 \left\{ \sum_{j=0}^{\infty} \exp(-r\,\tau_j) \left[I_{\tau_j} b + \int_0^{\tau_{j+1}-\tau_j} \exp(-r\,t)\, R\, m(t+\tau_j)\, dt \right] + \int_0^{\infty} \exp(-rt)(e-c(t))\gamma\, dt \right\} \tag{29}$$

subject to $dm = -(c + \pi m)dt$, $m(t) \geq 0$, where τ_j denote the stopping times at which an adjustment (jump) of m takes place, and $m(0) = m$ is given. The indicator I_{τ_j} equals zero if the adjustment takes place at the time of a free adjustment, otherwise it is one. In this formulation R is the opportunity cost of holding cash and there is *only one* Poisson process with intensity p describing the arrival of a free opportunity to adjust. The parameters of the problem are r, R, π, p, b, γ and e. Note the presence of R as well as the fact that there is only one Poisson process.

We are now ready to show conditions under which (28) and (29) are equivalent and to characterize the solution.

Proposition 5. Either a cash-burning or Baumol-Tobin policy with cash target m^* is optimal for the shadow cost problem (29) defined by parameters (e, r, π, b, p, R) if and only if it is optimal for the total cost problem (28) defined by parameters (e, r, π, b, p_1, p_2) with the same cash-target value m^*, provided that $R = r + \pi + p_2$ and that $p = p_1 + p_2$. Moreover the functions $G(\cdot)$ and $V(\cdot)$ satisfy

$$G(m) = V(m) - m + e/r + p_2 b/r$$

for all $m \geq 0$.

Few remarks are in order. First, note that the parameters (r, π, b, e) are common across the two problems. Second, the value of $R = r + \pi + p_2$ has a natural interpretation as an opportunity cost, it is the sum of the (implied) nominal interest rate and the probability (per unit of time) of losing all the cash. Third, the relation between the value functions is very natural: the total cost includes the present discounted value of the expenditure e/r, which equals all the future withdrawals that have to be made, net of the original cash at hand. The fact that cash is not used immediately is accounted for the opportunity cost $R = \pi + r + p_2$. An extra adjustment is the inclusion of the cost that must be paid when the cash is lost, the term bp_2/r. Fourth, the importance of this result for our purposes is that the optimality of cash-burning does not follow from an arbitrary assumption on the opportunity cost of holding cash. Indeed, it makes clear that when $\pi + r + p = 0$, then it must be the case that $R = 0$, and hence cash does not burn.

Proof. (of Proposition 5.) To show the equivalence we characterize the value function for the total and shadow cost. Once this is done, the proof follows immediately by using equation (5).

Characterization of Total Cost. The first order condition with respect to c gives

$$\begin{aligned} -(1+\gamma) - G'(m) &\leq 0 \text{ if } c = e \text{ , } -(1+\gamma) - G'(m) = 0 \text{ if } 0 < c < e \text{ and} \\ -(1+\gamma) - G'(m) &\geq 0 \text{ if } c = 0 \end{aligned} \tag{30}$$

Under a cash-burning policy there is threshold m^*, the value of cash that the agent chooses after a contact with a financial intermediary (the optimal "replenishment"). The best replenishment level for the agent, $m = m^*$, solves

$$m^* = \arg\min_{z \geq 0} \; z + G(z) \quad . \tag{31}$$

Under a cash-burning policy and using the value of the target m^* we can rewrite the bellman equation as

$$\begin{aligned} rG(m) &= G'(m)(-e - \pi m) + p_1 [m^* - m + G(m^*) - G(m)] + \\ &\quad + p_2 [b + m^* + G(m^*) - G(m)] \quad . \end{aligned} \tag{32}$$

In a cash-burning policy the boundary condition when $m = 0$ is that the agent will use credit until a withdrawal occurs, either because of a free opportunity or because of a theft, so that

$$rG(0) = e(1+\gamma) + (p_1 + p_2)(m^* + G(m^*) - G(0)) + p_2 \, b$$

Combining this boundary condition with (32) we have:

$$G(m) = \begin{cases} \dfrac{e(1+\gamma) + p_2 b + m^* + (p_1 + p_2) \; (m^* + G(m^*))}{r + p_1 + p_2} & \text{if } m = 0 \\ \dfrac{-G'(m)(e + \pi m) + (p_1 + p_2)[m^* + G(m^*)] + p_2 b - p_1 m}{r + p_1 + p_2} & \text{if } m \in (0, m^*) \end{cases}$$

A cash-burning policy is optimal if the agent prefers not to pay the cost b and adjust money balances in the relevant range:

$$m + G(m) \leq b + m^* + G(m^*) \quad \text{for } m \in [0, m^*] \ .$$

Instead following a Baumol-Tobin policy the value function satisfies:

$$G(m) = \begin{cases} b + m^* + G(m^*) & \text{if } m = 0 \\ \dfrac{-G'(m)(e + \pi m) + (p_1 + p_2)[m^* + G(m^*)] + p_2 b - p_1 m}{r + p_1 + p_2} & \text{if } m \in (0, m^*) \end{cases} \tag{33}$$

with boundary:

$$b + m^* + G(m^*) \leq \frac{e(1+\gamma) + p_2 b + m^* + (p_1 + p_2)\ (m^* + G(m^*))}{r + p_1 + p_2} \tag{34}$$

Summarizing, a m^* cash-burning policy is optimal if and only if G and m^* satisfy equation (5) and its first order condition in equation (30) for all $m \in [0, m^*]$, the target m^* satisfies equation (31), and the boundary condition equation (5) is satisfied. A m^* Baumol-Tobin policy is optimal if and only if G and m^* satisfy equation (33) and its first order condition in equation (30) for all $m \in [0, m^*]$, the target m^* satisfies equation (31), and the boundary condition equation (34) is satisfied.

Characterization of Shadow Cost. The derivation of the Bellman equation follows by the same logic used to derive equation (32). As in problem (28) we consider a cash-burning policy where the agents uses no credit as long as cash is available, i.e. that $c(t) = e$ as long as $m(t) > 0$, and to use credit when $m = 0$ until a withdrawal opportunity arises. This gives the following necessary and sufficient first order condition for the choice of c:

$$\begin{aligned} -\gamma - V'(m) &\leq 0 \text{ if } c = e \ , -\gamma - V'(m) = 0 \text{ if } 0 < c < e \quad \text{and} \\ -(1+\gamma) - V'(m) &\geq 0 \text{ if } c = 0 \end{aligned} \tag{35}$$

Denoting by $V'(m)$ the derivative of $V(m)$ with respect to m the Bellman equation for $m > 0$ satisfies

$$rV(m) = Rm + p \min_{z \geq 0} (V(z) - V(m)) + V'(m)(-e - m\pi) \quad .$$

Upon being matched with a financial intermediary the agent chooses the optimal adjustment setting $m = m^*$, or

$$V^* \equiv V(m^*) = \min_{z \geq 0} V(z) \ . \tag{36}$$

At $m = 0$ the agent uses credit and waits for a free withdrawal to arrive i.e.

$$rV(0) = \gamma e + p(V(m^*) - V(0))$$

Combining these equations we have:

$$V(m) = \begin{cases} \dfrac{\gamma e + pV^*}{r+p} & \text{if } m = 0 \\ \dfrac{Rm + pV^* - V'(m)(e + m\pi)}{r+p} & \text{if } m \in (0, m^*) \end{cases} \tag{37}$$

To ensure that it is optimal not to pay the cost and contact the intermediary in the relevant range we require:

$$V(m) \leq V^* + b \quad \text{for } m \in [0, m^*] \ . \tag{38}$$

Instead, if a Baumol Tobin policy is optimal, we have:

$$V(m) = \begin{cases} b + V^* & \text{if } m = 0 \\ \dfrac{Rm + pV^* - V'(m)(e + m\pi)}{r+p} & \text{if } m \in (0, m^*) \end{cases} \tag{39}$$

with

$$(r+p)(V^* + b) \leq \gamma + pV^* \,. \tag{40}$$

Summarizing, a m^* cash-burning policy is optimal for the shadow cost problem with if and only if V and m^* satisfy equation (37) and its first order condition in equation (35) for all $m \in [0, m^*]$, the target m^* satisfies equation (36), and the boundary condition equation (38) is satisfied. A m^* Baumol-Tobin policy is optimal for the shadow cost problem with if and only if V and m^* satisfy equation (39) and its first order condition in equation (35) for all $m \in [0, m^*]$, the target m^* satisfies equation (36), and the boundary condition equation (40) is satisfied. □