IEF

EIEF Working Paper 02/15
February 2015

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

by

Fernando Alvarez

(University of Chicago and NBER)

Francesco Lippi

(University of Sassari and EIEF)

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

Fernando Alvarez

University of Chicago, NBER

Francesco Lippi

University of Sassari, EIEF

February 18, 2015

Abstract

We present a model that characterizes the relationship between optimal dynamic cash management and the choice of the means of payment. The novel feature of the model is the sequential nature of the payments choice: in each instant the agent can choose to pay with either cash or credit. This framework predicts that the current level of the stock of cash determines whether the agent uses cash or credit. Cash is used whenever the agent has enough of it, credit is used when cash holdings are low, a pattern recently documented by households data from several countries. The average level of cash and the average share of expenditures paid in cash depend on the opportunity cost of cash relative to the cost of credit. The model produces a rich set of over-identifying restrictions for consumers' cash-management and payment choices which can be tested using recent households survey and diary data.

JEL Classification Numbers: E41, G2

---

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 requires no inventory of cash at hand so that, for instance, debit cards do not count as "cash" for our purposes. The novel feature of the model is the sequential nature of the payments choice: in each instant the agent can choose to pay with either cash or credit. One novel prediction is that the payment choices of households depend on the level of cash holdings: agents use cash for a purchase as long as they have any cash with them, but use credit otherwise. Still, agents find it optimal to intermittently replenish their cash holdings, and so will use both cash and credit. These results match recent survey and diary evidence such as, 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) who document that (absent other differences) households with larger cash balances are less likely to use credit. More colorfully, the agents in our model find that "cash burns" in their hands.[1] Our model establishes a tight connection between the size and frequency of withdrawals, the average cash balances and the share of payments made in cash. We fully develop these connections which can be used as "over-identifying" restrictions for the estimation and testing of the model. We also characterize the implications for the shape and interest rate sensitivity of the demand for cash. Finally, we find that the model gives a clear example of how some technological innovations are complementary to the use of cash: for instance it implies that the availability and density of ATM networks crowds out the use of credit.

The basic cash management problem that we consider is a version of the Baumol-Tobin inventory model where each withdrawal is subject to a fixed cost b, augmented to have random free withdrawal opportunities, with an expected number of p opportunities per unit of time. This augmented Baumol-Tobin model gives rise to a precautionary motive in the demand for cash holdings that is key to interpret the data as we showed in Alvarez and Lippi (2009). Expenditures are assumed to be a constant flow through time, denoted by e, and cash holdings are assumed to have an opportunity cost of R per dollar per unit of time. Agents discount the real costs incurred at the real rate r, and cash depreciates at the rate of inflation π. The key new element is that, unlike in the augmented Baumol-Tobin's model households can pay with cash or with credit at any time: if they pay with credit they incur

---

[1] Our model abstracts from differences in the size of purchases as well as differences in acceptability of different means of payments. We view our model as complementary of those that emphasize the role of the size of purchases on the choice of the means of payments. There is a large theoretical and empirical literature on models that emphasize those features, such as Arango, Bouhdaoui, and Bounie (2012) and Bouhdaoui and Bounie (2012), among many others.

a flow cost of γe per unit of time. The reason that agents use cash if they can (so "cash burns") is that for an agent with cash at hand the cost of obtaining it is sunk, so that the agent pays only the opportunity cost. Thus, as long as there is either discounting, inflation, or the option value of free cash-adjustments is positive (i.e whenever $p + r + \pi > 0$), agents prefer to pay in cash when they have it, i.e cash burns.

The model establishes conditions under which it is optimal for the agent to use both cash and credit. This means that agents withdraw cash (even though it burns) and to use credit only when their cash is inventory exhausted. We show that this policy is optimal when the fixed cost of adjustment b is higher than some threshold value. In addition to the "cash-burning" behavior, the model implies connections between observable cash-management behavior, and the dynamics of the choice of means of payments. The cash-management statistics that we refer to are: the frequency and average size of withdrawals, the average cash holdings and the average cash holdings at the time of withdrawals. The information about the means of payments concerns whether cash or credit is used (relative to the level of cash holdings), and the average share of expenditures paid with cash. For example, we show that cash-burning behavior implies a negative relationship between the share of payment in cash and the ratio of withdrawal size to average cash holdings.

Many papers in the literature incorporate alternative means of payments, as for example the seminal work of Lucas and Stokey (1987) and Prescott (1987). Those models do not have an explicit inventory theoretical model of money, so they cannot make contact simultaneously with observations such as the fraction of purchases made in cash as well as cash-management statistics such the frequency and size of cash withdrawals, etc. Technically, in this type of models the cash-in-advance constraints binds in every (exogenously determined) model period, so that a withdrawal occurs in every period and all cash is spent during it. Hence, statistics such as frequency of withdrawals, size of withdrawals, cash at withdrawals are all exogenously determined by the choice of the 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 in that the number of cash adjustment per period is modeled, those choices are all "within" the period, so they do not feature a *sequential* choice of how to pay for each purchase as cash is depleted. Hence in these models the optimal use of credit does not depend on cash holdings, as the data strongly suggest.[2]

The structure of the paper is as follows. The rest of the introduction expands on the

---

[2]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.

relation with existing models and on the description of the recent empirical household level evidence regarding cash balances and means of payment choice. We illustrate the model's key idea in Section 2 with a simple deterministic steady-state model. This section is useful mostly for readers that are less familiar with dynamic optimization models. It gives a simpler treatment of most, but not all, the properties of the model. Section 3 introduces uncertainty and a proper dynamic treatment of the inventory with payment-choice. This section gives a sharp characterization of the model parameters for which it is optimal to follow a cash burning policy (where cash balances influence the choice of means of payment) or alternatively where it is optimal to follow the augmented Baumol-Tobin policy (so the level of cash balances are not related to the choice of means of payment). Section 4 extends our theoretical analysis by allowing for a random cost of cash withdrawals, a feature which appears desirable for empirical applications. Section 5 concludes. The randomness of the withdrawal cost implies that an agent follows a cash-burning or an (augmented) Baumol-Tobin policy depending on the current realization of the withdrawal cost. This feature is useful to fit the data, as we illustrate with a simple calibration.

1.1 Literature Review

Next we review a few related money demand models which we think are both important and innovative. Our aim here is to explain the modeling simplifications made by these papers which did not allow them introduce the main feature of our model: the sequential choice of payment method. We start with the classic paper by Lucas and Stokey (1987), and the very large literature that followed this work. In each (discrete) period agents have preferences over two goods, denoted as cash and credit goods, which are imperfect substitutes in consumption. Cash goods are paid with previously accumulated cash but credit goods can be paid with income from the current period. Although this model features a choice of the means of payments, it is only through the choice of consumption or cash and credit goods. For a given good, there is no option to use different payment methods. Cash goods must always be paid for in cash, and the cash-in-advance constraint always binds. Thus in this class of models the number of withdrawals, the size of the withdrawals, and the cash at hand at the time of withdrawals are all functions of the exogenously given length of the time period.

The model in Prescott (1987) considers the problem of an agent who makes (simultaneously) purchases of differentiated goods. Those purchases can be paid with either cash or credit. If they are paid with cash, agents must have previously accumulated cash. If they are paid with credit, agents can pay them with their current income, but for each purchase they incur a fixed cost. As a consequence, agents pay small purchases with cash and large purchases with credit. The purchases of all the goods occur simultaneously, and the cash in

advance constraint for the goods that are chosen to be paid with cash holds with equality in each discrete time period. As a consequence, as was the case for the Lucas-Stokey cash-credit model, the frequency of cash withdrawal and the cash management statistics depend on the exogenously given length of the time period. Moreover the decision of whether to use cash or credit does not depend on the amount of cash at hand, a state variable that is key in our model.

Whitesell (1989) uses a model that combines a cash-credit choice and an inventory model for payments made in cash. He uses a steady state model (or, equivalently, a one period model) in the tradition of Baumol-Tobin and others. In his model agents are faced with a distribution of the size of purchases. There are two ways to pay for goods: cash or credit. Paying with credit has a per-transaction as well as a proportional cost. The total expenditure made in cash has a cost given by the opportunity cost of the average balances and the transactions cost associated with cash withdrawals, as in Baumol-Tobin. Thus the agent's decision consists of a threshold on the minimum size that is paid with credit, and a Baumol-Tobin policy used to finance the cash needed to by the remaining total expenditure, which is only on the smallest purchases. Importantly, all the expenditures of different sizes occur simultaneously. Thus in this model the agents are not faced with the choice of whether to pay for a transaction with cash or credit when they have cash at hand, as they do in our model. It follows that the decision of whether to use cash or credit does not depend on the amount of cash at hand.

The model in Freeman and Kydland (2000) is a full-blown RBC monetary model. In each (discrete time) period there is a combination of a version of Prescott's (1987) model and, for the goods that are not paid in cash, there is a cost as in the Baumol-Tobin model. We highlight that the Baumol-Tobin model applies *within* each model period. Equivalently, the expenditures of different sizes all occur simultaneously within the same period. Again, as in Whitesell (1989), this paper combines both payment choice and an inventory model, but does not expose agents to the choice of what means to payment to choose when they have cash at hand.

1.2   Empirical Evidence

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

5

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.

2 A steady-state deterministic model

This section considers a simple steady-state deterministic model that highlights the main features of the dynamic stochastic model of Section 3. The usefulness of this model is to highlight the main tradeoffs at work in a very simple way. Indeed some key formulas from this simple model coincide or are useful to interpret the more complex decision rules of the stochastic model presented later. The main failure of the deterministic model is the lack of a precautionary motive, so that $m = 0$ always at the time of a withdrawal. The readers familiar with continuous time impulse control problems may skip this section and move directly to Section 3.

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. In this simple set-up the cost γ can be understood as the time cost of using credit for small value transactions. The technology to withdraw cash is as follows. At any time the agent can pay a fixed cost b and obtain cash. In addition, the agent is given p free cash withdrawals per period.

We briefly comment on the main decisions that an agents must make. Note 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. First, 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.[3] Second, for an agent with zero cash balances $m = 0$, there are two possible choices. One is to pay the cost $\gamma\, e$ to consume using credit, and wait until the next free withdrawal opportunity to replenish cash balances. The other possibility is to pay the fixed cost b, withdraw cash, and consume using cash.

Before analyzing the full model, the next subsection describes the solution of the cash inventory problem that arises when the agent does not use credit. Then we consider a model with cash and credit purchases. Finally, we consider the best choice among the two.

---

[3]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.

2.1 Deterministic Baumol-Tobin model with p free withdrawals

The solution of the problem in this case is the same as in Baumol-Tobin model, except that the agent pays only for the withdrawals in excess of the p free adjustments per period, as in Alvarez and Lippi (2009). 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. Let R be the opportunity cost of cash. The agent's choice of m^* solves

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

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

$$\frac{m^*}{e} = \frac{1}{p} \sqrt{\min \left(2 \frac{b \, p^2}{e \, R}, \ 1 \right)} \ .$$

For $p = 0$ we have the Baumol-Tobin setup, so that when R is small the agent decides to economize on withdrawals and choose a large value of m^*. For $p > 0$ there is no reason to have less than p withdrawals, since these are free. Hence, for $b/e > R/(2 \, p^2)$ the agent will choose the same 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 $b/e \leq R/(2 \, p^2)$. In this deterministic set up, the average withdrawal size W and the average cash balances satisfy:

$$W = m^* \ , \quad M = 2 \, W = 2 \, m^*$$

Replacing the optimal m^* choice in (1) yields the value function for this problem:

$$v_a(R, b, p, e) = \begin{cases} \left(\sqrt{2 \, R \frac{b}{e}} - p \frac{b}{e} \right) e & \text{if } R \geq 2 \, p^2 \frac{b}{e} \quad \text{and } n > p \\[2ex] \left(\frac{R}{2 \, p} \right) e & \text{if } R < 2 \, p^2 \frac{b}{e} \quad \text{and } n = p \end{cases} \tag{1}$$

where the top branch gives the cost for the case in which the optimal number of withdrawals per unit of time exceeds the free rate p. The money demand interest rate elasticities is increasing in R:

$$-\frac{\partial \log M/e}{\partial \log R} = \begin{cases} \frac{1}{2} & \text{if } R \geq 2 \, p^2 \frac{b}{e} \quad \text{and } n > p \\[2ex] 0 & \text{if } R < 2 \, p^2 \frac{b}{e} \quad \text{and } n = p \end{cases}$$

2.2 The deterministic cash-credit model

We now turn to the analysis of the case where instead of withdrawing fresh cash whenever they run out, the agent instead uses credit and waits for a free withdrawal opportunity. After a cash withdrawal, denoted by $W = m^*$, she spends $\tau_a = m^*/e$ units of time paying for consumption with cash, incurring an opportunity cost $Rm^*/2$, where $m^*/2$ are the average cash balances, conditional on $m > 0$, and where R is the opportunity cost of cash. We assume that first the agent consumes using cash, and when cash balances are depleted she pays with credit. Thus, 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 value of the strategy that uses cash or credit is given by:

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

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

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

Denoting M the average real balances, and using the cash share s we can write $M = s \, m^*/2$. The solution of this problem gives

$$\frac{m^*}{e} = \frac{W}{e} = \min \left\{ \frac{1}{p} , \frac{\gamma}{R} \right\} \quad , \quad s \;=\; \min \left\{ 1 , \frac{\gamma \, p}{R} \right\} \quad , \quad \frac{M}{e} = \min \left\{ \frac{1}{2 \, p} , \frac{p}{2} \left(\frac{\gamma}{R} \right)^2 \right\}$$

$$v_r(R, \gamma, p, e) \;=\; \begin{cases} \left(\frac{R}{2 \, p} \right) e & \text{if } \gamma > \frac{R}{p} \\[2mm] \left(1 - \frac{\gamma \, p}{2 \, R} \right) \gamma \, e & \text{if } \gamma \leq \frac{R}{p} \end{cases} \tag{2}$$

All the expressions above have two branches: one when $\gamma \leq R/p$ so that both cash and credit are used, and one when $\gamma > R/p$ so that credit is not used.

2.3 The full deterministic problem

We now analyze the conditions under which is optimal to follow the strategy described in the previous section. That is, to use credit instead of withdrawing fresh cash when $m = 0$. To

do so we compare the steady-state cost of two policies computed above. The value function for the problem is then $v(R, b, \gamma, p, e) = \min \{v_a(R, b, p, e), v_r(R, \gamma, p, e)\}$. We define the threshold function $\underline{b}$, as the value of b that equates the two minimized costs:

$$v_a(R, \underline{b}, p, e) = v_r(R, \gamma, p, e) .$$

Using this definition we obtain that:

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

provided that $\gamma \leq R/p$, so that credit is used when $b \geq \underline{b}$, and provided that $R > 2\,p^2\,b/e$ so that $n > p$ for the case with no credit. Note that for $b < \underline{b}$ credit is not used and $n > p$.[4]

The next proposition summarizes the behavior of the full deterministic model. It is obtained from considering whether the minimum cost is attained by the cash-credit value described in equation (2) or by the value of the variation of the Baumol-Tobin model described in equation (2.1). Note that for some of the calculations it is useful to consider the expression in equation (3). The proposition considers two cases, depending on whether $\gamma \gtrless 2\,b/e\,p$, and for each case analyzes the behavior for all values of R.

PROPOSITION 1.   The interest rate elasticity of the money demand is decreasing in R, and takes either three or two values depending on parameters. Let $p > 0$ and $\gamma > 0$. Then $W/M = 2/s$ and

---

[4] The expression for $\underline{b}$ comes from equating: $v_r = \gamma e[1 - \gamma p/(2R)]$ with $v_a = e\sqrt{2\,R\,\frac{b}{e}} - pb$. Note that inserting the expression of $\underline{b}$ into v_a gives $e\gamma - p\gamma^2/(2R)e = \gamma e[1 - \gamma p/(2R)]$.

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

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

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

if $R \in \left(0\ ,\ 2\,p^2\,\frac{b}{e}\right]$ $\qquad -\frac{\partial \log M/e}{\partial \log R} = 0$ $\qquad$ only cash used $\qquad\quad n = p \quad s = 1$

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

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

if $R \in (0, \gamma p]$ $\qquad -\frac{\partial \log M/e}{\partial \log R} = 0$ $\quad$ only cash used $\qquad\quad n = p \quad s = 1$

This proposition illustrates four properties of the model. First, the interest rate elasticity of money demand is decreasing in the level of the interest rate. There are two cases: the first corresponds to a large cost of credit $(\gamma > 2\,p\,b/e)$, in which case there are three qualitative 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 becuase 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. This is the case discussed in the previous section.

The second property of the model is that the model really has only two parameters, γp and $p^2\,b/e$, as the alert reader will have already noticed. Note that in the model with free withdrawal opportunities analyzed and estimated in Alvarez and Lippi (2009) the shape of the money demand (both in the deterministic and stochastic version) depends only on $\hat{b} \equiv p^2\,b/e$. Thus the cash-credit aspect of the model in this paper depends only on γp, if $\hat{b}$ is kept constant. We will see that this carries through the stochastic model analyzed in Section 3.

Third, there are parameter values for which one obtains three extreme cases for all values of R. The classic Baumol-Tobin case is obtained by having both large cost of credit $(\gamma \to \infty)$ and no free withdrawals $(p \to 0)$. Alternatively, it can be obtained when the cost of

withdrawals is small $(b \to 0)$.[5] The case where only cash is used with an inelastic money demand is obtained when there are very frequent free withdrawal opportunities $(p \to \infty)$. The last extreme case is the one where only credit is used, which obtained when the credit cost is very small $(\gamma p \to 0)$.

Finally while the deterministic steady state model gives several of the insights of the dynamic stochastic model, here we comment on some properties which are both not desirable from an empirical point of view and also that are not robust in the sense that they disappear in the stochastic version of the model. These properties are that: i) the result that as $R \downarrow 0$ then $\underline{b} \to \infty$, so credit is not used, ii) the comparative static of the use of credit w.r.t. p, iii) the discreteness of the money demand and withdrawing elasticities, iv) the lack of "precautionary" hoarding, i.e. the result that cash is zero at withdrawal, e.g. $\underline{M} = 0$, v) the fact that $\frac{W}{M}$ is smaller than 2 if $n = p$ and $s = 1$, and vi) the fact that $\frac{W}{M} \geq 2$ if $s \leq 1$ (with equality holding only at $s = 1$.

3 An inventory model with means-of-payment choice

In this section we solve a discounted, stochastic dynamic problem which incorporates simultaneously: optimal cash management, and the optimal choice of means for payment. As in the deterministic problem of Section 2 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.[6] The state of the agent's problem is 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 Rm 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 $r \geq 0$. We first work with this standard minimization problem, and provide a derivation of the objec-

---

[5]Note that the two possibilities which give rise to the classic Baumol-Tobin model can be captured by simply considering the case where $p^2\, b/e \to 0$.

[6]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.

tive function and its corresponding opportunity cost from first principles in Appendix C. There are three substantive differences between the model 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, and we let the agent minimize the expected discounted cost, allowing any feasible policy.

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:

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

The outer $\min\{\cdot, \cdot\}$ in the functional equation (4) chooses between two strategies. The term in the first line represents the case where cash is not replenished, but where the agent still chooses what fraction of their consumption is paid for in cash, and what fraction by 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 opportunities and to due to the depletion of cash. The minimization with respect to c describes the agent's choice of the optimal means of payment. The second term 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 an adjustment is made or not, and if no adjustment is made which payment instrument (cash or credit) to use. In particular, it does not impose any pattern of when adjustment takes place or when different payment instruments are used.

We maintain the following assumptions for the model of this section.

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

We briefly comment on the assumptions. 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.

While in this section we treat R and r, p, π as parameters that can be set independently, the value of R and $r + p + \pi$ are indeed related –for instance $r + \pi$ should be the shadow nominal interest rates. We return to this relationship in the next section.

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:

$$(r + p)V(m) = mR + pV(m^*) - V'(m)[\pi m + e] \text{ for all } m \in (0, m^*] \text{ and} \tag{5}$$
$$(r + p)V(0) = \gamma e + pV(m^*) \tag{6}$$

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 pay in credit and wait for a free withdrawal opportunity.

For completeness we define an alternative policy, which we refer to as a modified Baumol-Tobin 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 in both policies cash burns in the agent's hands, since in both cases it is used to pay, rather than credit, as long as it is available ($m > 0$). But in the Baumol-Tobin case credit is never

13

used thus the model is unable to speak to the evidence on cash-credit. Note that when a m^*-Baumol-Tobin policy is followed, in the range of inaction $(0, m^*]$ the o.d.e. in equation (5) must still hold. However, under this policy the boundary condition at $m = 0$ is different, and it is given by:

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

For a cash burning policy to be optimal, i.e. to solve the problem in equation (4), one needs to establish three things:

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

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

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

Instead for a BT policy to be optimal, i.e. to solve the problem in equation (4), one needs to establish three things:

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

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

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

3.2 Characterizing the candidate policies

We now 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 (this mirrors our approach in Section 2), see Appendix A for the proofs. The lemmas will allow us to analyze the choice of the optimal policy. 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:

$$V(m) = \left(pV(m^*) - \frac{eR}{r+p+\pi} \right) \frac{1}{r+p} + m \frac{R}{r+p+\pi} + A \left(1 + \frac{\pi}{e} m \right)^{-\frac{r+p}{\pi}} \tag{8}$$

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

$$A = \frac{e}{r+p} \left[\gamma + \frac{R}{r+p+\pi} \right] . \tag{9}$$

14

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.

$$-\gamma - V'(0) = 0 \quad \text{and} \quad -\gamma - V'(m) < 0 \quad \text{for} \quad m \in (0\,,\, m^*]\,. \tag{10}$$

The next lemma characterizes the optimal cash target m^* for a cash-burning policy. When this value for the target m^* is used the condition in Item ii is satisfied.

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

$$V'(m^*) = 0\,, \tag{11}$$

holds, then m^* is given by:

$$0 \leq \frac{m^*}{e} = \frac{\left(1 + (r + p + \pi)\frac{\gamma}{R}\right)^{\frac{\pi}{\pi+r+p}} - 1}{\pi} \leq \frac{\gamma}{R}\,, \tag{12}$$

Moreover m^*/e is increasing in π.

Inspection of equation (12) 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.

LEMMA 3. For each $m^* > 0$ there is a unique value function V for the Baumol-Tobin policy. This has the same functional form as under a cash-burning policy except that now A in equation (8) is given by:

$$A = \frac{e}{r+p}\left[R\frac{m^*}{e} + (r+p)b + \frac{R}{r+p+\pi}\right]\,. \tag{13}$$

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 (11) holds, then m^* solves

$$\left(1 + \frac{m^*}{e}\pi\right)^{1+(r+p)/\pi} = \frac{m^*}{e}(r + p + \pi) + 1 + (r+p)(r+p+\pi)\frac{b}{eR}\,. \tag{14}$$

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. The next remark develops

15

the expressions for the special case where $\pi = 0$, by using L'Hopital rule in all the relevant formulas. These expressions are simpler, so we will focus on them.

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

$$\frac{m^*}{e} = \frac{1}{r+p} \log\left(1 + \frac{\gamma(r+p)}{R}\right) \geq 0. \tag{15}$$

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

$$\exp\left(\frac{m^*}{e}(r+p)\right) = 1 + \frac{m^*}{e}(r+p) + (r+p)^2 \frac{b}{eR}. \tag{16}$$

The analysis of the expression for m^* in equation (15) 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

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

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 (16) satisfies

1. $\frac{m^*}{e}$ is increasing in $\frac{b}{eR}$, $\frac{m^*}{e} = 0$ as $\frac{b}{eR} = 0$ and $\frac{m^*}{e} \to \infty$ as $\frac{b}{eR} \to \infty$.

2. For small $\frac{b}{eR}$, we can approximate $\frac{m^*}{e}$ by the the solution to the Baumol-Tobin model, or

$$\frac{m^*}{e} = \sqrt{2\frac{b}{eR}} + o\left(\sqrt{\frac{b}{eR}}\right)$$

3. The interest rate elasticity of m^*/e is smaller than $1/2$ and it is decreasing in $(r+p)/R$.

3.3 Characterizing the optimal cash-credit choice

The next proposition characterizes the optimal choice between the cash-burning and BT policies. 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 the

16

cost exceeds the lower bound $\underline{b}$, the condition in Item iii is verified, and when the cost is below it, the condition in Item c is satisfied.

PROPOSITION 2. A cash burning policy with m^* given by equation (12) is optimal provided that $b \geq \underline{b}$ where the lower bound for the fixed cost of adjustment is given by

$$\underline{b} = \frac{e}{r+p} \left[\gamma - R \frac{m^*}{e} \right] \tag{17}$$

and m^*/e is given by equation (12). Instead if $b \leq \underline{b}$ a Baumol-Tobin policy with m^* given by equation (14) is optimal.

Note that the results of this proposition for the case where $b > \underline{b}$ correspond to the deterministic steady state case of Proposition 1 where both cash and credit are used and where $n = p$. When $b < \underline{b}$ the optimal decision rule entails no use of credit. In this case the problem is the one analyzed in Alvarez and Lippi (2009), and corresponds to the deterministic steady state problem of Proposition 1 where credit is not used and $n > p$.

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

PROPOSITION 3. The function $\underline{b} \geq 0$ and it is homogenous of degree one in (γ, R). Moreover

$$\frac{\partial \underline{b}}{\partial \gamma} > 0 \quad \text{with} \quad \lim_{\gamma \to 0} \underline{b} = 0 \quad \text{and} \quad \lim_{\gamma \to \infty} \underline{b} = \infty ,$$

$$\frac{\partial \underline{b}}{\partial R} < 0 \quad \text{with} \quad \lim_{R \to 0} \underline{b} = \infty \quad \text{and} \quad \lim_{R \to \infty} \underline{b} = 0 ,$$

$$\lim_{r+p \to 0} \underline{b} = \frac{eR}{\pi^2} \left[\left(1 + \frac{\gamma\pi}{R}\right) \log\left(1 + \frac{\gamma\pi}{R}\right) - \frac{\gamma\pi}{R} \right] = \frac{\gamma^2}{2R} e + \pi o\left(\frac{\gamma^2}{R}\right) \geq 0 ,$$

$$\lim_{\pi \to 0} \underline{b} = \frac{e\gamma}{r+p} \left[1 - \frac{\log\left(1 + (r+p)\frac{\gamma}{R}\right)}{(r+p)\frac{\gamma}{R}} \right] = \frac{\gamma^2}{2R} e + (r+p) o\left(\frac{\gamma^2}{R}\right) \geq 0 ,$$

$$\frac{\partial \underline{b}}{\partial \pi} < 0 , \lim_{\pi \to \infty} \underline{b} = 0 , \frac{\partial \underline{b}}{\partial (p+r)} < 0 , \text{ and } \lim_{r+p \to \infty} \underline{b} = 0 .$$

We use Proposition 3 to analyze the circumstances under which a policy of cash-burning is optimal. The pattern of using cash to pay as long as it is available reflects that cash is "burning" in the hands of the agent. Yet even though cash burns if $b \geq \underline{b}$, then the agent finds it optimal to use credit until a free withdrawal opportunity arrives, replenishing her cash balances at that point. The proposition shows that this strategy id optimal if the fixed cost b of adjusting directly is high relative to the cost of using credit. An analogous condition holds in the deterministic steady-state model. Specifically, $\underline{b}$ 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. Equations equation (18) and equation (18) reflect this: as $\gamma/R \to 0$ then the threshold $\underline{b} \to 0$ so that only cash will be used. If γ^2/R is small then $\underline{b}$ coincides with the one of the deterministic model. Finally, the threshold $\underline{b}$ is decreasing in inflation: higher inflation increases the range of parameters for which the cash burning policy is optimal.

We note that the threshold $\underline{b}$ and the cash-target level m^* are both functions of the sum $p + r$, instead of p and r separately. From the point of view of the agent, they both enter as discount rates for the cost. Yet they have very different effects on other observables, as we will see below. For instance, consider two situations where $p + r$ is the same and where the parameters are such that cash burning is optimal, i.e. $b > \underline{b}$. Assume that in one case $p = 0$, so that agents will never use cash in the invariant distribution; once they have depleted their cash holdings they will never get another free opportunity to adjust. In contrast, for any other choice of p which keeps $p + r$ constant (and so $\underline{b}$ unchanged), agents will use cash in the long run.

The threshold itself is decreasing in $p + r$. This is because the cost of paying with credit is a completely static choice and so is independent of m, whereas the cost of paying with cash depends on m. With a higher discount rate, the future cost of holding cash is small. For instance, if $p + r \to \infty$, then there is no cost of holding cash, and hence $\underline{b} \downarrow 0$. Thus cash burning is optimal. Below we analyze a case where R and r, π, p are related, such as when take the opportunity cost is $R = r + \pi + p$. In this case, given the results above, we see that $\partial \underline{b}/\partial(p + r) < 0$, which confirms the intuition given by for the effect of the discounting $p + r$. Finally, note that in the deterministic case of the previous section, since there is no discounting, and since the free adjustment opportunities occur at fixed times, then $\underline{b}$ was not a function of p. This is also consistent with the intuition given for the effect of $r + p$.

3.4   Some predictions of the model

Next we consider several statistics of interest generated by a household who follows the optimal policy described above. We denote by s the cash share: the long run average fraction of purchases paid with cash. We denote by M the quantity of money. 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. We denote by 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.

Before analyzing the cash-burning policy we briefly review optimal cash management statistics when $b < \underline{b}$ and hence credit is not used. This case was studied in Alvarez and

Lippi (2009) and the reason for recalling it here is to compare it with the case that occurs when $b \geq \underline{b}$. To simplify the expressions we consider the case where $\pi = r = 0$. The point of including these expressions is to compare them with the case of $b \geq \underline{b}$ given below in Proposition 5.

PROPOSITION 4. Let $R > 0$, $p > 0$, $\pi = r = 0$ and $b < \underline{b}$. Then

$$n = \frac{p}{1 - e^{-m^* \frac{p}{e}}} > p \; , \; \frac{W}{M} = \frac{m^*}{M} - \frac{p}{n} \in (0, 2) \; ,$$

$$\frac{M}{e} = \left[\frac{1}{1 - e^{-\frac{p}{e} m^*}} \frac{m^*}{e} - 1/p \right], \quad \text{and} \quad \underline{M}/M = \frac{p}{n} \; .$$

where the optimal replenishment target m^* was characterized in Lemma 3.

Next we consider several statistics of interest for a household who follows the cash burning policy which is optimal when $b > \underline{b}$, so that credit is used.

PROPOSITION 5. Let $r + p > 0$, $\pi = 0$, $R > 0$. Assume that $b \geq \underline{b}$. 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 (15), one obtains closed form expression for s, M, W and $\underline{M}$. For $r \downarrow 0$ these expressions are simple functions of $\gamma p / R$:

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

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

The fact that the average cash at the time of withdrawal $\underline{M}$ equals the mean cash holdings M 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 4.

When both cash and credit are used, there are a number of interesting similarities and differences of the statistics in this proposition with those in the steady-state model analyzed in Proposition 1. We discuss these now. As in other dynamic inventory models, the first order expansion about the no-discounting case of many expressions coincides with those of

the deterministic steady state cash-credit model. The cash share in Proposition 5 for the case of $r = 0$ can be written as

$$s = 1 - \left(1 + \frac{\gamma\, p}{R}\right)^{-1} \approx \frac{\gamma\, p}{R}, \;\; \text{with } s \to 1 \text{ as } \frac{p\gamma}{R} \to \infty, \text{ and with } s \to 0 \text{ as } \frac{p\gamma}{R} \to 0 \;,$$

where $\approx$ denotes a first order expansion of s on $\theta \equiv \gamma\, p/R$ around 0. Thus the share of expenditure made in cash is increasing in the "credit cost" γp: the product of the cost of using credit per transaction and the rate of arrival of free opportunities to withdraw cash. It is also decreasing in the opportunity cost of carrying cash, R. The elasticity of s with respect to θ is $1/(1+\theta)$, which implies that the absolute value of the interest elasticity of s is given by:

$$-\frac{\partial \log s}{\partial \log R} = \frac{1}{1 + \frac{\gamma\, p}{R}} = 1 - s \equiv 1 - \text{cash share} \;,$$

which is increasing in R, decreasing in γp, and smaller than one. We use the expression for the cash share to illustrate the tension on the model between the use of cash and 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 $b > \underline{b}$. For these values, *only* credit is used, i.s. $s = 0$. Thus we find that the interesting case is the one where $b > \underline{b}$ but yet $p > 0$, so that $0 < s < 1$.

For $r = 0$ the scaled money demand Mp/e can be written as a function $\theta \equiv \gamma p/R$. The elasticity of this function is:

$$0 \le \frac{\partial \log Mp/e}{\partial \log \theta} = \frac{\left(\frac{\theta}{1+\theta}\right)^2}{\log(1+\theta) - \theta/(1+\theta)} \le 2$$

and is decreasing in θ. Thus money demand is decreasing in the opportunity cost R, with an interest rate elasticity *increasing* in R satisfying:

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

For comparison the model with cash purchases only ($c = e$) summarized in Proposition 4, which obtains if $b < \underline{b}$, 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 $1/2$.[7] This difference reflects the elasticity of the cash share s, which is between 0 and 1, and the

---

[7]So the elasticity is smaller or equal to the 1/2 implied by the Baumol-Tobin model.

elasticity of the money demand relative to *cash* consumption M/c, also between 0 and 1:

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

Now we comment briefly on W/M. Letting $r = 0$ and $b \geq \underline{b}$, note that W/M is decreasing in $\gamma p/R$, ranging from $W/M \to 0$ as $(\gamma \ p)/R \to \infty$, to $W/M \to \infty$ as $(\gamma \ p)/R \to 0$. One can see that $W/M \geq R/(\gamma \ p)$. Indeed a first order expansion of on γ around 0 gives $W/M \approx R/(\gamma \ p)$, i.e. it approaches the lower bound. For comparisons in the model with $b < \underline{b}$ we have that $W/M \leq 2$.

4   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 ($b < \underline{b}$) then credit is never used, i.e. $s = 1$ and the model becomes a version of Baumol-Tobin. These predictions 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 significantly 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 delivers a significant improvement of the model's predictions. The variation in b implies that agents follow a cash-burning policy when they face a high withdrawal cost, while they follow a Baumol-Tobin policy in the times when the value of b is above some critical threshold. The analysis can equivalently be conducted assuming the cost of using credit is 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 $\prime$ the derivative of V with respect to m. The value function solves the following functional equation:

$$0 = \min \left\{ \min_{0 \leq c \leq e} R\,m + \gamma[e - c] + p \min_{z \geq 0} [V(z, b) - V(m, b)] + \lambda \left[\int V\left(m, \tilde{b}\right) dF\left(\tilde{b}\right) - V(m, b) \right] \right.$$
$$\left. -V'(m, b)(c + \pi m) - r\,V(m, b) \ , \ \ b + \min_{z \geq 0} V(z, b) - V(m, b) \right\} \quad \forall\,(m, b) \in \mathbb{R}_+^2 \quad (18)$$

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. 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 $V(m, b) \geq V\left(m, \tilde{b}\right)$.

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 $b \geq 0$ agents use only cash. If $m = 0$ agents withdraw cash when $0 \leq b \leq \underline{b}$, and they use 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. We now turn to characterize the value function under a threshold cash-burning policy.

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

$$V^*(b) = \min_{z \geq 0} V(z, b) \quad \text{and} \quad \bar{V}^* = \int_0^\infty V^*(b) \, dF(b) , \tag{19}$$

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

$$V(m, b) = \left(p\bar{V}^* - \frac{eR}{r+p+\pi} \right) \frac{1}{r+p} + m\frac{R}{r+p+\pi} + \bar{A} \left(1 + \frac{\pi}{e}m \right)^{-\frac{r+p}{\pi}}$$
$$+ \frac{p}{r+\lambda+p} \left(V^*(b) - \bar{V}^* \right) + A(b) \left(1 + \frac{\pi}{e}m \right)^{-\frac{r+p+\lambda}{\pi}} \quad \text{for all } (m, b) \in \mathbb{R}_+^2, \tag{20}$$

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 (19), the parameters $(r, \lambda, p, \pi, R, \gamma, e)$, and the function F (see Appendix B 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

22

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 B.2, 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

$$V(0, b) = \begin{cases} b + V^*(b) & \text{if } b \leq \underline{b} \\ \frac{\gamma e + p V^*(b) + \lambda \bar{V}(0)}{r + p + \lambda} & \text{if } b \geq \underline{b} \end{cases} \tag{21}$$

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

4.1   A quantitative assessment

We use the model with the random b in a calibration that illustrates how it can produce cash management behavior featuring both cash and credit usage and where the amount of cash the time of a withdrawal is smaller than the average cash holdings, i.e. $M < \underline{M}$. As mentioned, the latter feature is key in the data but was not produced by the model with a constant cost of withdrawal studied in Section 3.

The relevant observable statistics that are used as benchmarks of our calibration are taken from the household evidence collected in Alvarez and Lippi (2009) (for Italy) and Alvarez and Lippi (2013) (for Italy and Austria). Table 1 in both of these papers shows that for the households with an ATM card, i.e. the ones who have access to a cash vs credit/debit choice, the share of cash expenditures s is in between 50 to 90 percent of total (non-durable) expenditures and that the average currency holdings is about 10 to 17 days of cash expenditures.[8] Moreover, the data shows that the amount of cash at the time of withdrawal relative to the average money balances $\underline{M}/M$ is between 0.3 and 0.4, that the ratio of the average withdrawal to the average money balances is about W/M is in between 1.2 and 1.6 and that the households with ATM withdraw cash about 40 to 70 times per year, i.e. $n \approx 60$. The left panel of Table 1 summarizes the relevant range for these variables.

The right panel of Table 1 reports the comparable moments predicted by the 3 cash management models described above. Column [1] uses the cash-only model described in Proposition 4. This model corresponds to the best cash management strategy when the withdrawal costs are low ($b < \underline{b}$), so that credit is not used. As discussed in details in Alvarez and Lippi (2009) this model has essentially 2 independent parameters, p and $\frac{bp^2}{eR}$, which are

---

[8]The use of debit payments is widespread in Europe. In Italy the vast majority of ATM cards also serve as debit cards.

Table 1: Selected moments on cash holding patterns : model vs data

		Range in Data	Model		
			[1]	[2]	[3]
			cash-only	cash-credit	mixed
Average Currency	M/c	(10, 17)	10	10	10
Withdrawal size	W/M	(1.3, 1.6)	1.4	1.4	1.4
Cash at withdrawals	$\underline{M}/M$	(0.3, 0.4)	0.4	1	0.5
Number of cash withdrawals	n	(40, 70)	26	18	26
Cash share of expenditures	s	(0.5, 0.9)	1	0.7	0.9

The table reports the unconditional sample means. Average currency is measured relative to cash expenditures per day, $c = se$. The number of cash withdrawals is per year. The parameters for model [3] are: $p = 13, r = 0.02, R = 0.02, \pi = 0, \gamma = 0.0035, \lambda = 10$. The low value of $b = 0.005$ or 0.5% percent of daily consumption.

set to reproduce a mean level of cash holdings $M/c = 10$, where $c = se$ per day so that M/c amounts to 10 days of cash consumption of average holdings, and a ratio of $M/W = 1.4$. The predictions for the other sample moments can be seen as an over-identifying test of the model. The share of cash expenditure is equal to 1 in this model, clearly higher than in the data where credit is also used.

The second model in column [2] is the cash-credit model described in Proposition 5. This model corresponds to the best cash management strategy when the withdrawal costs are sufficiently hight ($b > \underline{b}$) so that credit is 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 M/W. Both cash and credit are used in this specification (notice the cash share is 70% of purchases) although the model unrealistically predicts a very high level of cash at the time of withdrawals, namely $\underline{M}/M = 1$, as stressed above.

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 (0.5% of daily expenditures) or a high value (10% 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 when b is high the agent uses credit at $m = 0$ (since $b > \underline{b}$) waiting for a free withdrawal opportunity or a change in b. Intuitively, the behavior produced by this model is close (though not identical, because of option value motives) to a weighted average of the behavior of the cash-only and the cash-credit model characterized in columns [1] and [2]. The model parameters are taken from, or are close to, the structural estimates in Alvarez and Lippi (2009): the number of free withdrawal per year is $p = 13$, the opportunity cost of cash is $R = 2\%$ (this includes the nominal interest rate and the probability of cash theft, as

discussed in Appendix C).

In spite of its simplicity all of the model's predictions fall close within, or very close, to the ranges that are observed empirically in the dataset from Austria and Italy (reported in the first column of the table). The main point of this analysis is to illustrate the tractability of the model and its potential for empirical analysis. The simulation fits reasonably well 4 of the 5 data moments. As we argues above Proposition 4 and Proposition 5 show that the model is over-identified since the two underlying models have essentially 2 independent (combination of) parameters each. The mixed model adds one parameters (the mixing between the 2 underlying models) so that it is also over-identified. 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.

5 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 models of the cash-credit choice is the assumption that the choice is sequential, so that the agent chooses whether to use cash or credit purchase after purchase, a decision 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 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. 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.

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.

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.

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.

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.

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

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.

A   Proofs

Proof. (of Lemma 1.) Inserting equation (8) 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 (8) using equation (9) 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 $m \in [0, m^*]$. $\square$

Proof. (of Lemma 2.) Solving for $V'(m^*) = 0$ using equation (8) and equation (9). 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:

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

and

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

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

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

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 $\gamma/R > 0$. $\square$

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 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 form 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 $m \in [0 ,\ m^*]$. $\square$

Proof. (of Proposition 3) Note that $\underline{b} \geq 0$ is equivalent to $m^*/e \leq \gamma/R$. The homogeneity

28

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. $\square$

Proof. (of Proposition 5). 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

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

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^*:

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

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 as 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 times, 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

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

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)

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

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

□

Proof. (of Proposition 6.) To show the equivalence we characterize the value function

30

for the total and shadow cost. Once this is done, the proof follows immediately by using equation (6).

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

$$
\begin{aligned}
-(1+\gamma) - G'(m) &\leq 0 \text{ if } c = e \ , -(1+\gamma) - G'(m) = 0 \text{ if } 0 < c < e \ \text{ and} \\
-(1+\gamma) - G'(m) &\geq 0 \text{ if } c = 0
\end{aligned} \tag{22}
$$

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

$$
m^* = \arg\min_{z \geq 0} \ z + G(z) \quad . \tag{23}
$$

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

$$
\begin{aligned}
rG(m) &= G'(m)(-e - \pi m) + p_1 [m^* - m + G(m^*) - G(m)] + \\
&\quad + p_2 [b + m^* + G(m^*) - G(m)] \quad .
\end{aligned} \tag{24}
$$

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 (24) 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:

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

31

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

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

(25)

with boundary:

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

(26)

Summarizing, a m^* cash-burning policy is optimal if and only if G and m^* satisfy equation (A) and its first order condition in equation (22) for all $m \in [0, m^*]$, the target m^* satisfies equation (23), and the boundary condition equation (A) is satisfied. A m^* Baumol-Tobin policy is optimal if and only if G and m^* satisfy equation (25) and its first order condition in equation (22) for all $m \in [0, m^*]$, the target m^* satisfies equation (23), and the boundary condition equation (26) is satisfied.

Characterization of Shadow Cost. The derivation of the Bellman equation follows by the same logic used to derive equation (24). As in problem (37) 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:

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

(27)

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

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

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

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

(28)

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

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

Combining these equations we have:

$$V(m) = \begin{cases} \dfrac{\gamma e + pV^*}{r+p} & \text{if } m = 0 \\ \dfrac{Rm + pV^* - V'(m)(e+m\pi)}{r+p} & \text{if } m \in (0, m^*) \end{cases} \tag{29}$$

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

$$V(m) \le V^* + b \quad \text{for } m \in [0, m^*] . \tag{30}$$

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

$$V(m) = \begin{cases} b + V^* & \text{if } m = 0 \\ \dfrac{Rm + pV^* - V'(m)(e+m\pi)}{r+p} & \text{if } m \in (0, m^*) \end{cases} \tag{31}$$

with

$$(r+p)(V^*+b) \le \gamma + pV^* . \tag{32}$$

Summarizing, a m^* cash-burning policy is optimal for the shadow cost problem with if and only if V and m^* satisfy equation (29) and its first order condition in equation (27) for all $m \in [0, m^*]$, the target m^* satisfies equation (28), and the boundary condition equation (30) is satisfied. A m^* Baumol-Tobin policy is optimal for the shadow cost problem with if and only if V and m^* satisfy equation (31) and its first order condition in equation (27) for all $m \in [0, m^*]$, the target m^* satisfies equation (28), and the boundary condition equation (32) is satisfied.

□

B Threshold cash-burning policy: proofs.

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

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

$$\bar{V}(m) = \int_0^\infty V(m, b) \, dF(b) \tag{33}$$

33

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

$$(r+p)\bar{V}(m) = mR + p\bar{V}^* - [e + \pi m]\bar{V}'(m) \text{ for all } m \geq 0 \qquad (34)$$

Note that the o.d.e. (34) 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 (8) 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 (B). 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 (B) 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 (21) that at $b = \underline{b}$ we have:

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

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:

$$V^*(\underline{b})(r+\lambda) + \underline{b}\,(p+r+\lambda) - \frac{p\,\lambda\,\bar{V}^*}{(r+p)} + \frac{e\,R}{(r+p)} = \gamma e + \frac{eR}{(r+p+\pi)} + \lambda\,\bar{A} \quad (35)$$

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

$$
\begin{aligned}
0 = &\int_0^{\underline{b}} \left[V^*(b)\frac{r+\lambda}{p+r+\lambda} + b\right]dF(b) + F(\underline{b})\left[-\frac{p\,\lambda\,\bar{V}^*}{(p+r+\lambda)(r+p)} + \frac{e\,R}{(r+p+\lambda)(r+p)} - \bar{A}\right] \\
&+ (1 - F(\underline{b}))\left[\frac{\gamma e}{p+r+\lambda} + \frac{eR}{(r+p+\pi)(r+p+\lambda)} - \frac{p+r}{r+p+\lambda}\bar{A}\right]
\end{aligned} \quad (36)
$$

Note that given the function V^* we can regard equation (35) and equation (36) as a system of two equations in two unknowns $(\bar{A}, \underline{b})$. We can write equation (36) as:

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

35

and replacing $\lambda \bar{A}$ from equation (35) we write an equation for $\bar{A}$:

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

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

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

$\square$

B.1   Analytic expressions for $\bar{A}$, $\underline{b}$ and $A(\cdot)$.

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

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

$$
= \lambda \int_0^{\underline{b}} [V^*(b)(r+\lambda) + b(p+r+\lambda)]\, dF(b) - \lambda F(\underline{b})\left[V^*(\underline{b})(r+\lambda) + \underline{b}(p+r+\lambda)\right]
$$

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:

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

$$
A(b) = V^*(b)\frac{r+\lambda}{p+r+\lambda} + b - \frac{p\,\lambda\,\bar{V}^*}{(p+r+\lambda)(r+p)} + \frac{e\,R}{(r+p+\lambda)(r+p)} - \bar{A}, \;\; \text{if } b \leq \underline{b} .
$$

B.2 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 (B) together with equation (21) 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 $m = 0$.

C 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 cash is stolen, it must necessarily pay b, at which time it can adjust it cash with no cost.

The problem of the agent is:

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

$$\left. + \int_0^\infty \exp\left(-r\,t\right)(e - c(t))(1+\gamma)dt \right\} \tag{37}$$

subject to $dm = -(c + \pi m)dt - m\,dN_2$, where $m\left(t\right) \geq 0$, where N_2 is the Poisson counter of times where the cash was lost or stolen up to t, and where $m\left(0\right) = m$ 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.

C.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:

$$0 = \min \left\{ \min_{0 \leq c \leq e} (1+\gamma)[e - c] + p_1 \min_{z \geq 0} [z - m + G(z) - G(m)] \right.$$

$$+p_2 \min_{z \geq 0} [b + z + G(z) - G(m)] - G'(m)(c + \pi m) - r\, G(m) \quad,$$

$$\left. b + \min_{z \geq 0} [z - m + G(z) - G(m)] \right\} \quad \text{for all} \ \ m \geq 0 \ .$$

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.

C.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.

$$
V(m) = \min_{\{m(t),c(t),\tau_j\}} \mathbb{E}_0 \left\{ \sum_{j=0}^{\infty} \exp\left(-r\,\tau_j\right) \left[I_{\tau_j} b + \int_0^{\tau_{j+1}-\tau_j} \exp\left(-r\,t\right) R\, m\left(t+\tau_j\right) dt \right] \right.
$$
$$
\left. + \int_0^{\infty} \exp\left(-rt\right)(e - c(t))\gamma \; dt \right\} \tag{38}
$$

subject to $dm = -(c + \pi m)dt$, $m(t) \ge 0$, 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 (37) and (38) are equivalent and to characterize the solution.

PROPOSITION 6.   Either a cash-burning or Baumol-Tobin policy with cash target m^* is optimal for the shadow cost problem (38) defined by parameters (e, r, π, b, p, R) if and only if it is optimal for the total cost problem (37) 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 $m \ge 0$.

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.