Excess Capacity and Liquid Accounts as a Store of Value *

Alessandro Mennuni
University of Southampton

January 10, 2018

Abstract

This paper introduces a macroeconomic framework where liquid accounts emerge because of a search friction in the goods market. Specifically, the friction implies a link between money demand and the mismatch between the demand and supply of goods. The model has monetary equilibrium with empirically plausible amounts of monetary holdings despite large credit availabilities, as observed in modern economies. Furthermore, the theory generates an endogenous drop in Total Factor Productivity and a labour wedge, providing an original interpretation of recessions and an explanation for the empirical relationship between the velocity of money and excess production capacity.

JEL Classification Codes: E32, E41, E44
Keywords: Money; Credit; Equilibrium Excess Supply;

*I thank David Andolfatto, Marco Bassetto, Garth Baughman, Francesca Carapella, Chris Carroll, Andrew Clausen, Richard Dennis, Renato Faccini, Jonas Fisher, Chiara Forlati, Gaetano Gaballo, Ivan Jaccard, Paul Klein, John Knowles, Fabrizio Mattesini, Guido Menzio, Cyril Monnet, Pietro Reichlin, Victor Rios-Rull, Thomas Sargent, Kirill Shakhnov, Serhiy Stepanchuk, Roman Sustek, Randy Wright and seminar participants at the NBER SI 2015 consumption, Wharton, St. Louis Fed, New York Fed, Chicago Fed, the Board, the University of Glasgow, Queen Mary, Ensai, Global Macro at NYU Abu Dhabi, Macro Marrakesh, The Bank of England, The European Central Bank, for helpful comments.

1 Introduction

Figure 1 plots the velocity of various monetary aggregates as well as Total Capacity Utilization (TCU), a measure of spare production capacity.[1] Both velocity and TCU drop during the financial crisis. This is also true of all earlier recessions since the 80s. A decline in velocity is an increase in money over GDP. Why would agents hold more liquidity during recessions? This is especially interesting given that recessions are times in which more production capacity remains available as indicated by the decline in TCU.[2]

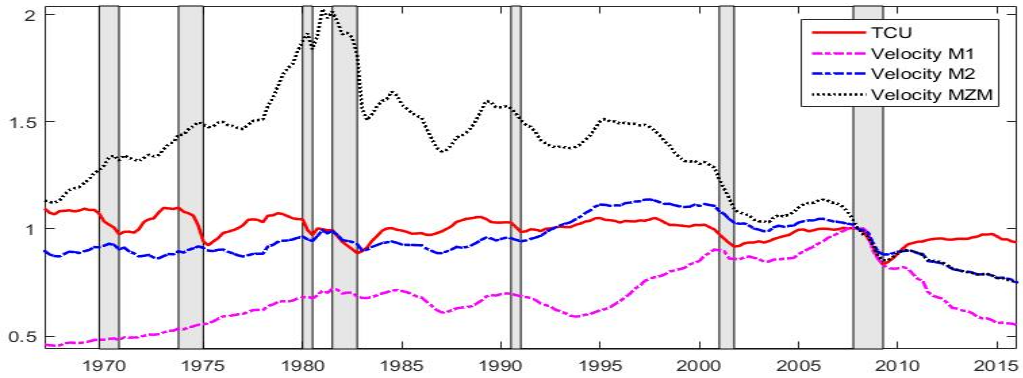

Figure 1: Total Capacity Utilization, and Velocity of M1, M2 and MZM. Source: FRED. Note: Time series are normalized to be 1 in 2008.

This paper develops a standard neoclassical model augmented with a search friction in the goods market where firms and households trade. The search friction induces a portfolio problem: because goods (consumption or investment) are hard to find, not all available funds are used to buy goods but the residual is optimally stored into the asset that is liquid in that it is not subject to the friction. Thus money demand emerges as a need to store unmatched available funds. Furthermore, a deterioration in the search friction causes a recession characterized by an excess in production

[1] FRED defines capacity utilization as how much capacity is being used from the total available capacity to produce demanded finished products.

[2] Mechanically, declines in velocity are partly explained by the decline in GDP, but it is not clear why the denominator (money) did not decline proportionally. In fact, checkable deposits, M1, M2, and MZM, all *increased* in levels at the onset of the financial crisis and before Quantitative Easing. To explain fluctuations in velocity has been a long standing challenge since Hodrick et al. (1991). In fact, at least until the financial crisis, the notorious instability of money demand led academics and central bankers to abstract from monetary targets altogether. Perhaps for this reason, the relationship between velocity and TCU has been neglected in academic circles but —this paper argues— it helps to shape a theory of money and of the business cycle. Recently, Lucas and Nicolini (2015) argued that the interactions between money and financial crises remain poorly understood.

capacity and in liquid savings as in Figure 1.

Another finding is that monetary equilibrium is consistent with large availabilities of credit: money has value even when agents can pay with credit for more than 100% of their income. This is not a common result given that credit crowds out transaction money demand as recently explained by Gu et al. (2016). Given the ease of credit in modern societies, it should be a fundamental issue in monetary economics to reconcile the coexistence of money with such amounts of credit. As the paper clarifies, that monetary equilibrium is so robust to credit has to do with the fact that money is held as a store of value. Coexistence of money and large amounts of credit is also key to match the sizeable levels of broad monetary quantities we observe which include the endogenous creation of inside money via credit: for instance, M2 ranged between 1.8 and 2.8 times quarterly GDP since 1959 in the US.

It is instructive to draw a comparison with the New Monetarist approach — e.g. Kiyotaki and Wright (1989) (KW), Shi (1997), Lagos and Wright (2005) (LW) — where like here, money is microfounded through a search friction in the goods markets. However, in these theories the emphasis is mainly on the transaction role. In contrast in this paper, the transaction motive alone does *not* generate money demand (even without credit). As the paper clarifies, this result relies on the presence of capital which can be turned into consumption. Thus the role of money identified here is polar to that in LW, where the transaction role has its clearest exemplification because the portfolio problem highlighted here is overcome through an ex-post walrasian market. To be clear, both mechanisms are present in KW, where there is no centralized market to solve the portfolio problem induced by the search friction, and there is no reversible asset to neutralize the transaction role of money. However, to focus on this neglected angle seems telling as relative to transaction theories, in a sense the issue is reversed upside-down: money is not demanded in order to make transactions easily, but because one does not make transactions easily.[3]

Two further and related results are that money is important for welfare and that the matching friction is an important driver of the business cycle. These results stem from the following mechanism: the possibility to store value in the liquid asset makes agents less preoccupied about not finding goods, but look for better trading opportunities. Technically, money leads to a market tightness (firms over buyers)

[3]The implications are also different than those of cash-in-advance models like Lucas and Stokey (1987) or Svensson (1985).

where the probability of selling goods is higher for firms relative to the non monetary equilibrium.[4] So money is desirable because it increases firms' ability to sell i.e. their measured productivity. However, the mechanism also opens the door to a source for recessions as trading (or matching) conditions can deteriorate, leading buyers not to spend their money and firms not to sell their goods.

This link between the excess supply of goods and demand for money formalizes an age old economic intuition which relates to Walras' law and can be traced back to Mill (1844): "... there cannot be an excess of all other commodities, and an excess of money at the same time."[5]

This idea, which lies at the core of the neoclassical-keynesian dispute, gained renewed attention in the recent years of increased economic turmoil: according to this view, the financial crisis resulted in a recession because agents stopped spending for consumption and investment but hoarded wealth in unproductive but safe assets.[6]

Indeed there is evidence that the financial crisis was characterized by a surge in the holdings of liquid assets as is reflected in the large decline in the velocity of money and in the record-high amounts of cash held by firms. As mentioned, that this liquidity surge is related to an excess supply capacity is consistent with the drop in TCU during the Great Recession, as well as in all earlier recessions from the 80s onward.[7]

While the paper is mainly theoretical, a simple quantitative section brings this model to the data. First, is the original matching mechanism consistent with the data? I derive matching probabilities and market tightness from data on GDP, production capacity, and available funds, and find that they have the properties of the model: firms and buyers matching probabilities are decreasing and increasing functions of

[4]The ratio of firms over buyers can also be rearranged as the value of supply over demand.

[5]Here the excess supply of goods is an equilibrium outcome given the search friction.

[6]Search can be seen as capturing what hinders the ability to trade quickly, such as information acquisition, for example. But that goods are hard to find does not mean that one could not blindly buy, say, a car or stocks in virtually no time. However, to find the right house, or even some items of clothing, can take a while. To an extent, even shares and most financial products have unique and opaque characteristics that justify a search process as is indicated by the fact that financial intermediaries retain a lot of power even though nowadays many such markets are electronic. In this context, that money is a liquid store of value because not subject to the friction relates to the idea that information insensitive securities should serve as liquidity, Gorton and Pennacchi (1990).

[7]Instead, the relationship between velocity and TCU was negative before the 80s. The model explains this changing pattern through a different combination of shocks. These correlations do not vanish when controlling for output growth, interest rates, or inflation: regressing both TCU and velocity over the aforementioned variables does not explain these facts as the residuals of the 2 regressions exhibit similar patterns (positively correlated from the 80s and negatively before).

market tightness respectively. This was not obvious as these probabilities and market tightness were not constructed using the matching function. These results suggest that this theory provides a good description of the matching between demand and supply; furthermore, a structural estimation finds that the matching process plays a predominant role in understanding recessions.[8] In particular, a wedge to the matching function generates an endogenous drop in total factor productivity (TFP), a surge in liquidity (or decline in velocity), disinflation, a drop in capacity utilization, and a labour wedge which is qualitatively consistent with the business cycle accounting of Chari et al. (2007). Intuitively, a drop in the efficiency of the matching process captures some disruption in the intermediation between buyers and sellers and leads to a drop in sales with a resulting increase in excess capacity and in the holding of liquid assets.[9] This shock is distinguished from a pure demand shock (a shock to search effort similarly to Bai et al. 2017), which is also present and played a role for the recessions prior to the 80s, explaining why velocity did not decline.[10]

The paper proceeds as follows. Section 2 discusses the literature further, Section 3 sets up the model, Section 4 offers a theoretical characterization and Section 5 includes the quantitative analysis. Section 6 concludes and suggests possible applications. The appendixes contain proofs and a description of the data.

2 Literature Review

Liquid assets are present in many other branches of the literature. First, a micro-foundation of fiat money typically interpreted as a store of value is offered by the overlapping generations model pioneered by Samuelson.[11]

[8] The structural estimation also highlights that the model is numerically tractable, despite the microfoundation of money. Since the model is robust to credit, it is possible to relax the assumption of anonymity and also have insurance markets and study a representative agent model grounded into a modified neoclassical setting which can be linearized.

[9] In traditional models this wedge is by and large imputed to shocks to the production technology. While this efficiency shock too can be seen as a "measure of our ignorance", it seems instructive to suggest an identification for part of the Solow residual through the matching between demand and supply, especially given that it also explains movements in velocity, TCU, and in the labour wedge.

[10] According to this model, technology shocks are not an important cause of recessions. One reason for this result is that labour slightly decreases in response to a positive technology shock. So a technology driven recession would be accompanied by an increase in labour, which is counterfactual. This result is reminiscent of Galí (1999); intuitively, demand plays a role.

[11] However, Wallace (1980) challenges the existence of a clear cut distinction between medium of exchange and store of value in overlapping generations models.

In the Baumol-Tobin (BT) framework an ad hoc cost generates a portfolio problem in reallocating wealth between the liquid and the illiquid asset: recent examples include Alvarez and Lippi (2009), Ragot (2014), and Kaplan and Violante (2014). Although BT models often assume that purchases need to be mediated by the liquid asset (cash-in-advance), this is not crucial so there too money is inherently a store of value. So one can think of the search friction in this paper as a different way to induce a portfolio problem that makes money a useful store of value. Furthermore, the implications are rather different and it is made explicit that money need not be the only form of payment. The two frictions also seem distinguished conceptually, in fact, a cost can be added on top of the search friction (for example, search effort is introduced in the model; this is conceptually similar to a cost, but there is monetary equilibrium also without it). Distinguishing between the cost and the search mechanism seems fruitful given the size and the many varieties of liquid assets.[12]

In the Bewley models (Bewley 1980) consumption uncertainty and the lack of insurance, lead to precautionary liquid savings. This liquidity in excess of expected consumption is usually interpreted as a store of value. However, for money to coexist with assets that pay higher financial dividends, it is necessary to give to money the transaction advantage of being the only asset that can be quickly exchanged for goods to buffer idiosyncratic shocks: see Wen (2015).[13] Other related frameworks where liquidity arises as a combination of timing and credit frictions are Holmstrom and Tirole (1998) and Diamond and Dybvig (1983). One contribution relative to these literatures is to offer a motive for liquidity that is robust to the presence of insurance and credit technologies. Besides the different microfoundations, the business cycle implications of these theories are rather different.

It should be noted that these theories are not mutually exclusive. For instance Telyukova and Visschers (2013) have both precautionary, and transaction money demand through a cash in advance constraint to account for the variance of velocity. Wang and Shi (2006) also account for the variance of velocity with search intensity, and with a transaction motive.[14] Furthermore, Telyukova (2013) reconciles the

[12]For instance, given the emphasis on the store of value, at least to a first approximation, the notion of liquidity could include other assets such as government debt.

[13]Alternatively, it is possible to have capital subject to uninsurable idiosyncratic shocks directly so that it is more risky than money as in Brunnermeier and Sannikov (2016).

[14]While these papers match the unconditional variance of velocity, they do not tackle the issue of generating a liquidity surge during recession of the observed magnitudes. Furthermore, these models (including Bewley and BT models) do not relate velocity to TFP, TCU, and the labour wedge.

coexistence of money holdings and rolled over credit card debt in a model where consumption uncertainty cannot be fully insured through credit cards. See also Wen (2015) for some forms of credit insurance in the Bewley framework.

There is a growing literature with search frictions in the goods market: examples include Bai et al. (2017), Huo and Ríos-Rull (2013), Petrosky-Nadeau and Wasmer (2011) and Den Haan (2014). The main contribution to this literature is to use it to construct a theory of liquidity. Furthermore, disciplining the model through monetary quantities elicits the distinction between matching shocks and demand shocks.

The paper is also related to the vast literature that models the financial crisis through credit constraints building on Bernanke and Gertler (1989) and Kiyotaki and Moore (1997). In a sense, these theories work in the opposite way to that in this paper. In these models during recessions firms *wish* to produce more but are constrained.[15] Here instead, firms do *not wish* to produce more because of their difficulties to sell.[16] These two channels are possibly both real. However, since the present paper is not based on borrowing constraints, but on the incapacity to spend even having access to liquidity, it offers an alternative financial explanation for the recession. And in this model too borrowing declines during recessions, however this does not happen because of tighter credit but because of fewer spending opportunities.[17]

Finally, like the New Keynesian framework, this one may prove useful for policy analysis. An advantage is that it explains monetary quantities.[18]

[15] See Kiyotaki and Moore (2012), Shi (2012), Jermann and Quadrini (2012), Brunnermeier and Sannikov (2014), Christiano et al. (2014), Iacoviello (2015) and Cui and Radde (2016) among others.

[16] The difference is also reflected in the aspects of liquidity that are emphasised: there money can be more easily *spent* than how other assets can be pledged for credit (this is a transaction role). Here money can be more easily *acquired* than other goods.

[17] This paper does not study monetary policy other than showing that money is neutral but not superneutral and that the Friedman rule is optimal. But it is worth pointing out that while the two theories are both consistent with a shrink in loans, they may have different policy implications: open market operations aimed at easing credit conditions can be effective in models with credit constraints (see Kiyotaki and Moore 2012) but may not be as effective in this model which, to a degree, subscribes to the adage: you can lead a horse to water but you cannot make it drink. Evidence on the effects of Quantitative Easing is mixed, see Williamson (2015) and references therein.

[18] In relation to this issue, Woodford (1998) showed in an influential paper that even in the presence of credit that lead to a cash-less limit, the framework remains useful for monetary policy. This raises the issue of whether monetary aggregates are important at all. This theory accounts for the fluctuations in monetary aggregates. In fact, they are key in developing the theory and in the identification of the sources of the business cycle.

3 The Model

Time is discrete. The economy is populated by a continuum of measure one of households that live forever. In each period static firms produce goods for consumption and investment purposes with a neoclassical production function of labour and capital. Besides consumption and capital, there is a costlessly storable object, called money, which is divisible and intrinsically useless. The money supply M is constant in most parts of the paper; money growth is considered in Section 4.

Similarly to the standard neoclassical model, in each period firms sell goods to households while labour and capital inputs are supplied by households and demanded by firms. These two latter input markets are competitive. Instead, the market for goods is subject to a search friction.[19]

The market structure for goods is as in Menzio et al. (2013). There is a continuum of submarkets indexed by the terms of trade $(p, q) \in \mathbb{R}_+ \times \mathbb{R}_+$ where p is the price per unit of good paid by the household (the buyer) and q the quantity that goes from the firm (the seller) to the buyer. So pq is the actual payment made by the buyer. A firm chooses how many trading posts to create in each submarket (i.e. how many units of size q to put for sale in each submarket) and a household chooses which submarket to visit. It is convenient to use one of these submarkets as the numeraire.

As is typical in search models, the buyer cannot visit multiple submarkets in the same period and can at most find one trading post. So the matching process is such that a household and a trading post meet in pairs; let the matching function μ be concave and homogeneous of degree one in the number of trading posts f and households h, with continuous derivatives. In a sub-market with tightness $\theta = \frac{f}{h}$, let $\psi(\theta) = \mu(f, h)/h = \mu(\theta, 1)$ denote the probability with which a household or buyer finds a trading post, and $\phi(\theta) = \mu(f, h)/f = \mu(1, 1/\theta)$ the probability with which a trading post is matched with a buyer. The function ψ is strictly increasing with $\psi(0) = 0$ and $\psi(\infty) = 1$. ϕ is strictly decreasing with $\phi(0) = 1$ and $\phi(\infty) = 0$.

Search is competitive as in Moen (1997). So the terms of trade cannot be ex-post renegotiated. However, similarly to the neoclassical model, the payment pq need not take the form of money: firms also accept to deliver the good for credit. To clarify, it is useful to specify the following timing within the period: the input markets clear at the beginning but payment from firms to households is deferred to the end, after

[19] It would be possible to consider search frictions for the inputs markets too. But to isolate the key novelties, the model is kept as close as possible to the neoclassical one.

firms revenues are realized. After the input markets clear and before inputs are paid, households and firms make transactions in the frictional market. To pay, households can issue an intratemporal bond, a promise for later payment at the end of the period. Furthermore, at the end of the period households can also issue to other households an intertemporal bond to roll over any existing debt. At maturity, bonds clear in money, or by clearance of net bond positions, or by issuance of a new intertemporal bond. Bonds are fully liquid: an agent that holds a bond issued at the end of period t can use it during $t+1$ to pay firms before maturity at the end of $t+1$.[20]

Of course, with enough credit, money has no value. But it is shown that the amount of credit necessary for that is the maximum possible.

Market tightness varies with the terms of trade across the sub-markets according to the equilibrium function $\theta(p,q)$, which is taken as given by firms and households. As a result, the probabilities ϕ and ψ are endogenous functions of (p,q).[21]

3.1 Households

Households liquid funds at the beginning of a period are $p_m m + a$, where m is money, p_m is its the price in terms of the numeraire sub market, and a is the value in terms of the numeraire of the intertemporal bond holdings.[22]

A household enters submarket (p,q) such that

$$pq \leq p_m m + a + B, \tag{1}$$

$B \geq 0$ is the maximum the household can borrow from the firm by issuing the intratemporal bond. Absent credit ($a = B = 0$) Equation (1) would be a pure cash in

[20] It would be possible to allow agents to also issue intertemporal bonds before the end of the period to pay for goods, and then the static firms use them to pay for the inputs at the end of the period. This would be redundant given the presence of an intratemporal and an intertemporal bond.

[21] Competitive search is adopted here because it does not add a bargaining inefficiency, thereby not introducing a further element of departure from the neoclassical framework.

[22] Households also own the capital stock, which –like firms production– could in principle be put on sale in the frictional goods market. However, this market does not operate once insurance markets are introduced later. Clearly a household would never put on sale capital at the same price as the one in which she buys: this is because with the proceedings she may not buy back goods for the same amount given the search friction, and would hold the rest in money, which pays lower return. A household would be willing to sell in a submarket with a higher price than the one at which she buys, but (in the representative agent environment that follows after the introduction of insurance) it would not find anyone willing to buy at that same higher price, so the submarket would not be active. Appendix C in the working paper version Mennuni (2017) shows this result formally and it explicitly introduces a secondary market for capital.

advance constraint. This way to introduce credit resembles Gu et al. (2016) (GMW). While others have introduced credit in different ways, drawing on several specific frictions which all have elements of truth (see Cavalcanti and Wallace (1999) and Berentsen et al. (2007) among others), for this study, this approach seems very clean in that, up to the borrowing limit, credit is perfectly substitutable to money.

With probability $\psi(\theta(p,q))$ there is a match so the household pays pq and buys goods for q which can be used as consumption c or as investment i:

$$c + i = q.$$

Capital accumulates according to $k' = i + k(1-\delta)$, where $\delta \leq 1$ is the depreciation rate. Furthermore, end of period capital $k' \geq 0$. I.e. the household can disinvest to the point of consuming up the entire capital stock.[23]

At the end of the period she receives income payments $wn + kr$ where w is the real wage, n hours of work, and r the rental price of k. The firm pays with its revenues, i.e. either with money, or by turning the bonds it received from other households.[24]

Since the household spent pq (either in money or bonds), her end of period balance after honoring the debt is $p_m m + a + wn + kr - pq$.[25]

This balance is stored in money m' or it can be saved in a'. Each unit pays at par at maturity and costs v:

$$p_m m' + v a' = p_m m + a + wn + kr - pq.$$

With probability $1 - \psi(\theta(p,q))$ the household does not make a transaction; in this case she does not buy any goods and at the end of the period she is left with the

[23] Since I will introduce insurance markets and Inada conditions in the utility function, this constraint is only avoiding Ponzi schemes, but it does not induce the sort of precautionary savings it would in an incomplete market model à la Aiyagari (1994).

[24] At the end of a period each agent i has a net position of intratemporal bonds equal to the ones received by firms (and issued by some other household) less those she issued to a firm. Call $\hat{b}_i$ the bonds issued by agent i and $\hat{b}_{-i,i}$ the bonds agent i receives by firms (the indexes $-i, i$ emphasise that the bond is issued by some household other than i, and passed on to i). So agent i's net position is $\hat{b}_{-i,i} - \hat{b}_i$. The sum of all households' net positions is $\sum_i \hat{b}_{-i,i} - \sum_i \hat{b}_i$. Since firms pass to households all bonds they receive $\sum_i \hat{b}_{-i,i} = \sum_i \hat{b}_i$, i.e. the intra-temporal bond market clears. Similarly, at the end of the period the household has intertemporal bonds issued in the previous period by some other household $\hat{a}_{-i,i}$ less those she issued $\hat{a}_i$ and $\sum_i \hat{a}_{-i,i} = \sum_i \hat{a}_i$. If the net credit position of an individual agent is negative (positive), she can pay (be paid) in money or by issuing (accepting) the intertemporal bond. Since the intertemporal bond is issued by an agent and accepted by an other, $\sum_i \hat{a}_i = 0$. See Ljungqvist and Sargent (2012) Section 18.10.3 for a discussion of the exchange and clearance of IOUs in similar settings.

[25] Like in GMW, the balance is independent of whether pq is paid in money or credit as the intratemporal interest rate is zero.

initial money and bond holdings plus income:

$$\begin{cases} c + k' - k(1-\delta) = 0, \quad k' \geq 0 \\ p_m m' + va' = p_m m + a + wn + kr. \end{cases}$$

One could relax this extreme distribution assumption that one either trades in full or not al tall. However, insurance—included next—smoothes its implications.

It is also necessary to impose a lower bound on the inter-temporal bond: $a' \geq \underline{a}$. This only avoids Ponzi schemes but it is loose enough to never bind in equilibrium.

3.1.1 Insurance

That some agents trade and others do not generates heterogeneity in assets holdings. Ways to maintain tractability in search models are either to assume a big family with many agents as in Shi (1997), or to use the timing and preference structure developed in LW. However, since there is no need to assume anonymity to rule out credit, but actions are monitored, it is possible to have insurance for all households in the same sub-market.[26] Assuming the law of large numbers holds so that ψ is the exact share of the population that successfully made a transaction, all households that participated in the same sub-market by being ready to pay $x = pq$ for q, receive goods for ψq and pay ψx. I.e. the share ψ of households that made a transaction, transfer $(1-\psi)q$ of goods each and are thus left with ψq. The transfers sum up to $\psi(1-\psi)q$ which can be divided among the remaining $(1-\psi)$ share of the population so that each receives ψq. In turn, those that receive the goods transfer, make a payment of ψx in liquid assets. It is easy to check that this way each agent in the same sub market receives goods for ψq and pay liquid assets for ψx so that the end-of-period liquid balance is $p_m m + a + wn + kr - \psi pq$ for all. It is shown later that equilibrium in which liquid balance is positive and $p_m > 0$ exists. So while insurance removes heterogeneity, it does not remove the portfolio problem that makes money balances positive.

[26] Monitoring is necessary because without it, a possible strategy is to go to a market that one cannot afford but where one's full balance is equal to the ex-post payment $\psi(\theta)x$. This way the household is not able to pay x in case of a successful match but could then pretend to have not matched and claim a transfer from the other households. Of course, this way overall transactions would not be enough to sustain the insurance scheme. For this reason, anonymity rules out the presence of insurance. Notice that the insurance suggested here does not have the incentive issues present in the big family assumption in Shi (1997): there agents do not respond to their individual incentives but act in the interest of the entire family even though they are not monitored.

As it is well known, with monitoring it is possible to construct non monetary equilibria with the same or better allocation than the monetary equilibria (Kocherlakota (1998)). In fact, Aliprantis et al. (2007) show that even with full anonymity such equilibria may exist. But these non monetary equilibria require strategies as a function of other people observed actions so that punishment of defection is possible by selecting bad sub-game perfect Nash equilibria. Similarly to LW and GMW among many others, this paper does not focus on such equilibria and agents behaviour is function of economic state variables, but not of the actions that led to such outcomes.

Insurance simplifies the analysis and makes it clear that the role of money does not depend on the absence of insurance. However, it is worth discussing the meaning of goods redistribution with search frictions. An interpretation consistent with ex-post redistribution is that goods come in different varieties and each household can only store (and therefore buy) a subset of such varieties but a variety is not known before visiting the trading post. After purchases are made, there can be perfect insurance between households that like the same variety. It should be noticed that this theory of money does not hinge on this insurance assumption: it would be possible to solve the model without the insurance market and allow for heterogeneity to spread.

3.1.2 The representative household problem

With insurance markets it is possible to study the problem of a representative household: she starts each period with capital k, money m and bonds a. For recursive equilibria, the aggregate state Ω is composed of the aggregate capital stock K and money M, and of a vector of shocks with a known Markov process to be defined later.

The household solves the following problem with rational expectations:

$$V(k,m,a,\Omega) = \max_{\{c,n,d,k',m',q,p\}\geq 0, a'\geq \underline{a}} u(c,n,d) + E\beta V(k',m',a',\Omega') \tag{2}$$

$$s.t. \quad pq \leq p_m m + a + B(k,n,\Omega,\exists), \tag{3}$$

$$q \leq A_d d, \tag{4}$$

$$c + k' - k(1-\delta) \leq \psi(\theta(p,q))q, \tag{5}$$

$$p_m m' + va' \leq p_m m + a + wn + kr - \psi(\theta(p,q))pq. \tag{6}$$

Where β is the discount factor. The utility function $u(\cdot)$ is increasing in c, and decreasing in n and in shopping effort d, whose role is explained later. $u(\cdot)$ is concave and has continuous derivatives with $\lim_{c\downarrow 0} u_c = \infty$, $\lim_{n\downarrow 0} u_n = 0$ $\lim_{d\downarrow 0} u_d = 0$.

The household takes input market prices w and r as given. E indicates rational expectations taken over next period aggregate state Ω' given Ω.

Equation (3) restates (1) where the borrowing constraint B is allowed to be a generic function of $(k, n, \Omega, \exists)$ where $\exists$ stands for equilibrium meaning that the borrowing constraint may be equilibrium specific (e.g. whether the equilibrium is monetary or not). This form allows for all cases considered in this study, for instance, in one case studied later agents will be allowed to borrow their entire end of period income so that $B = wn+rk$. This specification also spans the case of exogenous credit limits (as the function B can depend trivially on its explanatory variables) and —even though not studied here— that of endogenous constraints due to limited enforceability with credit either unsecured or collateralized through capital.

Equation (4) allows for a demand constraint as effort d is needed to look for goods, and $A_d > 0$ is an effort productivity parameter. Effort is usually introduced in theories that incorporate search frictions in the goods market such as Bai et al. (2017). However, in usual specifications effort enters directly into the matching function. Instead with my specification the matching theory also works without this constraint and the theoretical and quantitative implications of effort are disentangled from those of the search friction. This seems an advantage because how to specify effort generated some controversy: see Bai et al. (2017), Huo and Ríos-Rull (2013), Kaplan and Menzio (2016). The motivation for including the effort constraint even though it is not crucial is that it helps qualify some of the theoretical results, and distinguish matching shocks from demand shocks in the quantitative section.

Equation (5) shows that only a fraction ψ of demand q is matched with investment and consumption goods. What is left is invested in liquid assets as shown in (6): the right hand side shows the end of period balance after insurance, as elaborated earlier. From this latter equation, it is intuitive why money may have value in this economy: $\psi < 1$ implies that not all available funds $p_m m + a + B$ can be spent in goods. As formalized later in Proposition 1, if B is not at its loosest implementable level, the right-hand-side of (6) is positive, i.e. there is left over wealth which gives rise to money or bond demand. Since bonds are in zero net supply, in equilibrium $a' = 0$ and there is positive money demand.[27]

[27]In this representative agent environment, for intertemporal bonds to be in positive supply they would have to be a liability of the government. But since there is no liquidity difference, the distinction between money and government bonds would be intangible. This can also be appreciated by the first order conditions for m' and a' in Appendix A; they imply an arbitrage Fisher equation

So, other things equal, the smaller ψ the larger $p_m m'$. However, it should be noticed that households effectively choose ψ (and thereby end-of period money holdings) by choosing p and q, which determines market tightness given the equilibrium function $\theta(p, q)$. They can also choose $\psi \to 1$. So why are agents willingly holding money?

The Portfolio Problem

The first order condition for p illustrates the key trade-off in the decision of buying goods versus holding money. Focusing on an interior solution, the equation is

$$\frac{\partial \psi}{\partial \theta}\lambda_3 = \frac{\partial \psi}{\partial \theta}\lambda_4 p + \frac{\partial p}{\partial \theta}\left(\lambda_1 + \lambda_4 \psi\right), \tag{7}$$

where λ_1 — λ_4 are the Lagrange multipliers on Constraints (3)—(6).

The left-hand-side shows the marginal gain: with a higher p, θ increases (it is shown later that θ is increasing in p), this increases ψ so that agents end up with more goods, thereby relaxing (5) as captured by λ_3. However, with the increase in ψ agents are left with less liquid funds, this tightens (6) as captured in the first term in the right hand side by $\lambda_4 p$. Finally, there is a last term which is positive. So the net of the first two effects is positive and $\lambda_3 > \lambda_4 p$: put differently, agents prefer goods to money and if it was for these two effects only, they would put $\psi = 1$ so that all funds are turned into goods. The last term captures the fact that, to increase θ one has to pay more per unit of good (higher p): this tightens constraints (3) and (6). So the last term captures the portfolio problem.

Store of value versus transaction motive

To see how the transaction motive alone is not sufficient to generate money demand and instead appreciate the role of the portfolio problem, notice that here, if the portfolio problem was addressed by giving agents the chance to re-balance their money-goods holdings through an end-of-period centralized market as in LW, they would leave with $m' = 0$.[28] This does not mean that money does not have a transaction role: indeed money relaxes the Transaction Constraint (3). But this motive is

that pins down v so that money and bonds pay the same return. It would be possible to relax the assumption of perfect substitutability and distinguish between money and government bonds by assuming a small search friction for bonds. The intertemporal bond highlights that this theory of money does not rely on the fact that agents are not allowed intertemporal credit. In fact, without insurance, the bond would be traded but it would still be in aggregate zero net supply, thus leaving space for money demand to store the remaining unmatched aggregate savings.

[28] See Lagos and Rocheteau (2008). In particular, they find conditions under which coexistence of money and capital in a LW framework is possible even when capital can be used for transactions in the decentralized market. However, capital cannot be consumed in the decentralized market.

not sufficient to demand money because, as mentioned, agents prefer capital. Intuitively, capital pays higher dividend and can be consumed at any time. So, if there was no portfolio problem, agents would relax the goods constraint and tighten the transaction constraint by turning m' in k'. In models where money is demanded for its transaction role, agents do not turn all money into goods even if they can. Indeed in LW they could leave the centralized market with no cash but they choose to hold it so that they can make transactions in the decentralized market.

This store of value motive due to unmatched funds is also what makes money so robust to credit because so is the portfolio problem generated by the search friction: here for money to loose value, credit must be such that the matched agents can spend not just their end of period income, but also the funds of the unmatched agents. As formalized in Proposition 1, this requires agents not to be credit constrained and to borrow up to the maximum implementable limit. This is not the case in the LW framework where the level of credit from which money has no value is binding and it is not the maximum implementable limit as shown in GMW, Proposition 1.

3.2 Firms

Firms can choose to open trading posts in any market identified by price and quantity. A trading post in market (p, q) has a match with probability $\phi(\theta(p, q))$, in which case it sells q. To open a trading post, a firm needs production capacity $Ak_d^\alpha n_d^{1-\alpha} \geq q$, where k_d and n_d are the capital and labour inputs.[29]

A trading post in market (p, q) gives expected profits

$$\pi(p, q) = \max_{k_d, n_d} \phi(\theta(p, q))pq - wn_d - rk_d \tag{8}$$

s.t.

$$q \leq Ak_d^\alpha n_d^{1-\alpha} \tag{9}$$

[29]Otherwise a firm could open many trading posts and exploit the law of large numbers across them to have production capacity only for sales: $Ak_d^\alpha n_d^{1-\alpha} = \phi(\theta(p, q))q$. Ruling this out implies some excess production capacity and an endogenous Solow residual. As in Bai et al. (2017), it is assumed that excess production capacity is not storable. This assumption seems reasonable for services and nondurables, which form the large majority of GDP. In future it may be interesting to allow for inventories, but to match its rich dynamics (e.g. procyclical inventory investment) the model should be complicated for instance by introducing S-s policies or stockout-avoidance motives; see Wen (2011) for a recent analysis.

The first order conditions for capital and labour are

$$\xi(p,q)\alpha A\left(\frac{n_d}{k_d}\right)^{1-\alpha} = r, \tag{10}$$

$$\xi(p,q)(1-\alpha)A\left(\frac{k_d}{n_d}\right)^{\alpha} = w. \tag{11}$$

Where $\xi(p,q)$ is the lagrange multiplier on the production constraint. These two first order conditions imply that $\frac{k_d}{n_d}$ is the same in any trading post. Then $\xi(p,q)$ is equal for all (p,q). Thus it is going to be called ξ from now onward.

Using the 2 first order conditions, maximized profits can be written as

$$\pi(p,q) = \phi(\theta(p,q))pq - \xi q \tag{12}$$

Since firms can choose between any market (p,q), all potentially active markets must give the same profits. Furthermore, free entry implies that such profits must be zero: if profits were positive there would be infinite posts and $\phi(\theta(p,q)) = 0$, which contradicts that profits are positive. Equation (12) and $\pi(p,q) = 0$ imply

$$\phi(\theta(p,q))p = \xi. \tag{13}$$

Equation (13) has to hold for a market to be active and defines function $\theta(p,q)$.

3.3 Equilibrium

Before defining an equilibrium, it is useful to point out a few properties. The next lemma states that Equation (13) implies that $\theta(p,q)$ trivially depends on q.

Lemma 1 *$\theta(p,q)$ does not depend on q.*

The intuition behind the proof in Appendix B is that since the production function has constant returns to scale and input prices are taken as given, production increases proportionally with costs. Then, for profits to be independent of q, ϕ and thereby θ have to remain constant. From now on the function θ will be denoted $\theta(p)$.

It is also immediate from Equation (13) that $\theta(p)$ inherits the differentiability properties of ϕ and that p is a strictly increasing function of θ.

Finally, from Equation (13) it is clear that given θ, p is proportional to ξ. In other words, Equation (13) pins down a functional relationship between θ and p up to a

value for revenues per unit of production ξ. This value is free and can be normalized.[30] As a normalization, ξ is chosen to be equal to the equilibrium value of ϕ. This implies $p = 1$ in the equilibrium submarket as is immediate from Equation (13).[31]

Appendix C contains a formal equilibrium definition. The market clearing conditions are that households purchases are equal to firms sales:

$$\psi(\theta)q = \phi(\theta)fq; \tag{14}$$

and market clearing in the liquid assets, and inputs markets:

$$m' = M, \qquad a' = 0, \qquad fn_d = n, \qquad fk_d = K.$$

4 Characterization

The next subsection clarifies the role of the two frictions —search in the goods market and imperfect credit— for the existence of monetary equilibrium.

4.1 Money and Credit

At the end of each period money demand is equal to the liquid funds (initial money holdings + end-of-period income) not spent to buy goods because of the search friction, which implies that not all agents get to spend.[32] So money demand is what is leftover. Then, for it to be zero, credit has to be large enough so that the agents that are matched spend their liquid funds plus those of the agents that are unmatched. This leaves no leftover. For this to happen, the agents that are matched need to borrow against their end-of period income plus the liquid funds of the agents that

[30] For this one has to show that all other prices (r, w and p_m) also change proportionally to ξ, so that no relative price is changed. It is immediate from Equations (10) and (11) that given an allocation, r and w are also proportional to ξ. Expressions for r and w and p from Equations (10), (11) and (13), can then be substituted into the budget constraints —Equations (3) and (6)— to show that p_m is also proportional to ξ. Since no constraint is changed, neither will the optimal choices and thus the equilibrium allocation.

[31] Following Moen (1997), I assume that if agents are indifferent between multiple submarkets, only one will open. Multiple active submarkets could be possible in principle because both households and firms arbitrage between submarkets call for increasing functions between p and θ, so they may be tangent more than once for some special parameterization. As discussed in Appendix F, numerically I find that the possibility of multiple active markets does not occur given the restrictions imposed to the matching function.

[32] Liquid funds also include a but $a = a' = 0$ in equilibrium. So the intertemporal bond is often ignored from here onward.

are unmatched. Whenever credit is below that level, some funds remain unspent giving rise to money demand. As formalized in Proposition 1 and in Corollary 1, such pivotal credit level is

$$L = wn + rk + \frac{(1-\psi)}{\psi}(p_m m + wn + rk). \tag{15}$$

Furthermore, L is the maximum credit that can be reached in equilibrium. So the credit friction necessary for money to have value is very mild.[33] The intuition why L is the maximum equilibrium credit is that it would not be possible for agents to borrow more than against their own income $wn + rk$ plus all the funds of the unmatched $(1-\psi)(p_m m + wn + rk)$, divided by the number of matched agents ψ. Finally, the amount of credit can be appreciated in terms of income which in this model is equal to $wn + rk$. When money has no value $p_m = 0$ so from Equation (15) credit for this occurrence must be at least $\frac{wn+rk}{\psi}$ i.e. larger than 100% of income when $\psi < 1$.

Proposition 1 *Let intratemporal borrowing be defined as* $\hat{b} = pq - p_m m - a$. $p_m = 0$ *if and only if* $\hat{b} = L$ *where* L *— defined in Equation* (15) *— is the maximum implementable borrowing limit.*

Corollary 1 $p_m > 0$ *if and only if* $\hat{b} < L$.

It should be noticed that the results do not imply that whenever agents are allowed to borrow L, (i.e. when the borrowing limit $B \geq L$) then money has no value; it remains to be seen if and when that amount of borrowing occurs in equilibrium. It may be possible for instance, that agents choose not to borrow that much even if the are allowed to, i.e. the borrowing limit would be slack.

The next lemma shows the value of money when the credit limit is not binding: $\hat{b} < B$. Without costly effort there is no monetary equilibrium (because with no credit restriction, they borrow to the maximum implementable credit limit so $B \geq L$ for it not to bind), with costly effort this has only be proven in steady state and in deterministic equilibrium pathes converging to a steady state: the reason why the result may not always hold is that agents may not always want to borrow up to the maximum implementable limit (thereby demanding money to store residual wealth) because to buy goods requires effort.

[33] This is not true of other models of money and search where there are credit constraints below the maximum implementable limit for which there cannot be monetary equilibrium, see GMW.

Lemma 2

1. Without effort costs and with no binding borrowing limit, money has no value.

2. With effort costs and with no binding borrowing limit, money has no value in steady state and in any deterministic path that converges to a steady state.

As shown in in the proof, the result relies on the Euler equation for m', Equation (28). From this condition it does not seem possible to rule out coexistence of money and non binding credit limits when risk is present and effort is costly. Intuitively, if the expected value of money is positive (even if money may loose value in some future state that occur with probability smaller than 1) there is an incentive to holding $m' > 0$ even with perfect credit because one saves in effort.[34] This form of money hoarding is also akin Baumol-Tobin models where to hold money saves some costs.

A further corollary follows.

Corollary 2 *If the credit constraint binds, then money has value.*

Intuitively, because default is not allowed, agents never want to borrow above the natural limit L, then credit constraints can only bind if below L. But with credit below L, money has value from Corollary 1. Put differently, money has value as long as there are some operating credit frictions, no matter how mild. This result stands in contrast to other monetary models. For instance, GMW in Proposition 1, case 2, show that there are debt levels that are binding and yet money has no value. So there the credit restrictions need to be sufficiently severe for money to have value.

4.2 Money is essential

It is worth noticing that the propositions above characterize monetary and non monetary equilibria, but do not provide parametric conditions that mark the existence and non-existence of monetary equilibria. Put differently, L is endogenous. Then, for some exogenous credit limit B it is in general possible to find an equilibrium in which $B \geq L$ and thus the equilibrium is non monetary. But for that same B there might be a monetary equilibrium too. This is because money is essential so that in a non monetary equilibrium, the economy contracts so much relative to a monetary equilibrium, that the borrowing limit B is not binding as L has contracted.

[34] This money hoarding behaviour can be especially appealing during recessions when the return from capital goods may be very low, akin to the keynesian liquidity trap.

To show how money can improve the allocation relative to a non-monetary equilibrium Proposition 2 covers a special case, when agents can borrow up to their end-of-period income, $B \equiv wn + rk$. Although this credit limit seems high as it generates enough liquidity that money is useless in cash in advance models, that is not the case here. In fact, it does not generate enough liquidity to sustain any production in a non-monetary equilibrium; this highlights how dramatically bad the non-monetary allocation can be without enough credit relative to a monetary equilibrium where for the same credit limit, production takes place thanks to the extra monetary liquidity.

Proposition 2 *If $B \equiv wn + rk$, then no production takes place in a non monetary equilibrium.*

Intuitively, if buyers cannot store value in the form of the liquid asset, they prefer markets where it is inefficiently too easy to buy goods, but they don't internalize that this hinders firms ability to sell goods.[35]

$B \equiv wn + rk$ is also a pivotal case as with more credit, e.g. if $B \equiv \gamma(wn + rk)$ with $\gamma > 1$, some production takes place in the non-monetary equilibrium.[36] Instead with less credit ($\gamma < 1$) there can *only* be monetary equilibrium.[37] Intuitively, in this case the credit constraint always binds thus Corollary 2 applies.

4.3 Changes in money supply

The next proposition shows that money is neutral, but not superneutral. To allow for money to change over time, households receive a lump sum monetary transfer $dm = M' - M$. Thus $p_m dm$ is added to the right hand side of Equations (3) and (6).

Proposition 3 *Money is neutral but not superneutral.*

[35] The argument formalized in the proof goes as follows. For any finite market tightness not all liquid funds are turned into spending. Since in equilibrium spending equals income, liquidity > spending = income. But in a non-monetary equilibrium liquidity = credit so putting credit = income as in the proposition, one gets income > spending = income. Evidently market tightness must be infinite so that each agent is matched to a firm with probability 1 and all liquid funds are turned into spending: liquidity = spending = income. However, this implies zero probability of each firms to match and hence no hiring and no production.

[36] If $p_m = 0$, from Equation (3) $q > wn + rk$, which combined with Equation (32) implies $q > \phi A k^{\alpha} n^{1-\alpha}$. Then (34) implies $\psi < 1$ and thus $\theta > 0$ and $\phi > 0$.

[37] Suppose $p_m = 0$, then $q < wn + rk$, then (34) implies $\psi > 1$ which is impossible.

4.4 Changes in the credit limits

What are the effects of changing the function B? GMW show that changes in credit conditions can be neutral. The result holds here too in steady state if the debt limit is "lump sum" in the sense that it is independent of the households inputs of the credit limit (individual k and n). Intuitively, a level change in B should affect the total liquid balance $p_m m + B$, but this is neutral because p_m responds endogenously to keep total liquidity constant. And since changes in p_m are neutral, there are no other effects. Of course, just like in GMW, once B is enlarged enough so that the $p_m = 0$, then B matters, but the equilibrium must be nonmonetary.[38]

Proposition 4 *Take the steady state of a monetary equilibrium given a credit function B independent of k and n. Change B to zB with $z > 0$ and such that the new equilibrium price $\hat{p}_m > 0$, then the steady state allocation is unchanged.*

To appreciate the importance of the assumption of lump sum debt suppose for instance that B is moved from $B \equiv wn + rk$ to $z(wn + rk)$. Then the first order conditions for labour and k' are affected (with possible real effects) because $\partial B/\partial n$ and $\partial B/\partial k'$ are affected. This finding is akin the well known result that Ricardian equivalence can hold with lump sum taxes, but not when taxes are distortionary.

Finally, it should also be noted that lump sum debt is neutral, but not superneutral: a change in the growth rate of B would affect the growth rate of p_m with real effects.

4.4.1 The Friedman rule is optimal

Since money is not superneutral, this section discusses monetary policy in order to achieve efficiency. It is first necessary to define efficiency. For that, I construct a planner problem. Since this subsection characterizes deterministic steady state results, for simplicity, the planner problem abstracts from the shocks.

Definition 1 *An allocation $\{c, n, d, q, k', \theta\}$ is efficient if it solves the following:*

$$\widetilde{V}(k) = \max_{\{q,c,k',\theta,d,n\}\geq 0} u(c, n, d) + \beta E\widetilde{V}(k') \tag{16}$$

s.t.

$$\theta q \leq Ak^{\alpha} n^{(1-\alpha)} \tag{17}$$

[38] The next proposition is restricted to the steady state because it is shown in the proof that steady state inflation is not affected by a change in B. However, the inflation rate can be affected outside the steady state, with consequential real effects.

$$q \leq A_d d \tag{18}$$

$$c + k' - k(1-\delta) \leq \phi(\theta)\theta q \tag{19}$$

The planner chooses market tightness θ (or equivalently the number of trading posts f as households have measure 1 so $f = \theta$).

Equation (17) ensures that total production is not smaller than the quantity offered by each trading post (q) times the number of trading posts θ. Constraint (18) states that the planner has to respect the household's effort constraint, this is equivalent to Equation (4) in the household problem and it is repeated for convenience. The resource constraint, Equation (19), is derived from Equation (5), the equilibrium condition (14), and the fact that θ is equal to the number of trading posts f.

The next proposition shows that in steady state, the first order conditions of the planner and the household coincide at the Friedman rule. If the household problem is concave, this implies that the planner outcome is an equilibrium. It should be noted that concavity does not hold for any parameterization: as discussed in Appendix F, it is necessary to have sufficient complementarity in the matching function.

Proposition 5 *In the steady state of a monetary equilibrium, the first order conditions of the planner and the household coincide at the Friedman rule.*

It should be noticed that Proposition 5 holds irrespective of the credit limit: intuitively the credit limit is not binding at the Friedman rule (FR).

Corollary 3 below clarifies how inflation distorts the allocation in the case of no effort costs. With inflation agents choose a market with too high θ relative to the planner solution; this way it is easy for buyers to find goods so they remain stuck with less money. However, they do not internalize that this reduces the productivity of firms. In fact, the corollary shows that the FR calls for large amounts of liquid savings: the possibility to store in money with high return makes agents choose a lower market tightness which is efficient because it makes firms sell more goods. A low market tightness also implies a low ψ and thus more savings in money, but this is not a cost at the FR where money gives the same return as capital.

Before moving to the corollary it is useful to discuss the planner solution with no effort costs. In this case it is optimal to put $\theta \to 0$. This is because Constraint 18 in the planner problem does not bind and Constraints (17) and (19) imply

$$c + k' - k(1-\delta) \leq \phi(\theta) A k^{\alpha} n^{(1-\alpha)}.$$

From this last equation it is evident that θ approaching zero is optimal because then ϕ tends to one and all production is either consumed or invested, so there is no waste. With θ approaching zero, Equations (17) and (18) imply that q and d approach infinity. Intuitively, the number of trading posts tend to zero, but become large.[39] This extreme result with no effort costs also highlights the role of demand in this model: there is a benefit for the planner to make households search in crowded markets (where the ratio of households per trading posts is high), because the higher the demand for each trading post, the higher ϕ. This also implies high d and low ψ, but it is not a cost if effort is free. Is this implementable? The next corollary shows that θ is chosen optimally at the FR and not otherwise.[40]

Corollary 3 *Assume that effort is not costly and consider the steady state of a monetary equilibrium. Then at the Friedman rule, $\theta \to 0$, $\phi \to 1$, $\psi \to 0$. When inflation is above the Friedman rule, $\theta \nrightarrow 0$, $\phi \nrightarrow 1$, $\psi \nrightarrow 0$. Furthermore, assume any bounded credit limit B, then $p_m m \to \infty$ at the Friedman rule and $p_m m$ bounded otherwise.*

Of course, when effort is costly, there is a further cost of choosing a lower market θ because it implies a larger q and hence more effort. Thus θ is bounded away from zero and the value of money is bounded even at the optimal allocation. In fact, through numerical exercises it is found that when effort is not costly, inflation is highly distortionary, while when effort is costly, the cost of inflation is very small as θ is fairly insensitive to inflation. With the benchmark parameters estimated in the quantitative section (with costly effort and low effort supply elasticity), agents would only be willing to give up 0.03% of their steady state consumption to be at 0 inflation rather than having 10% annual inflation. This is much lower than what found by LW (between 1.4 and 3.2%). However, with no effort costs, the number is 7.3%.

Finally, it is worth mentioning that bargaining systems other than competitive search may introduce effort and market tightness inefficiencies even at the FR, possibly implying different optimal inflation rates.

[39] To understand this it is useful to draw a comparison with labour search models such as Mortensen and Pissarides (1994); there market tightness is given by the ratio between vacancies and unemployment. If vacancies were free to post, free entry would imply infinite vacancies. Hence the cost of search here takes the role played by vacancy posting costs in Mortensen and Pissarides (1994).

[40] In this case with no effort cost, that θ is optimal at the FR is proven without assuming concavity: at the FR the household first order conditions are *only* consistent with the efficient level of θ.

5 Mapping the Model to Data

In this section, I map this model to the data. In particular, it is shown that the matching function can be recast to take as inputs aggregate demand and aggregate production capacity. Thus, it is interesting to find empirical counterparts and to see whether this modelling device offers a reasonable and useful description of the data. Subsequently, I study the business cycle implications of the model; the upshot is that the matching friction accounts for a large share of the business cycle fluctuations.

5.1 Matching

I assume the following matching function:

$$\mu = z_m^{1/\rho} \left(\alpha_m f^\rho + (1-\alpha_m) h^\rho\right)^{1/\rho}. \tag{20}$$

μ is the number of matches and z_m is the matching shock. This specification is convenient because as ρ approaches minus infinity, the function converges to $\min(f, h)$, and the model becomes perfectly competitive and with $\theta = \psi = \phi = 1$. See the working paper version Mennuni (2017) for a more detailed discussion.

5.1.1 Recasting the Matching in terms of aggregate demand and supply

The variables in Matching Function (20) μ, f, and h do not have empirical counterpart so how to bring the matching function to the data? Multiplying the right and left hand side by q one gets

$$y = z_m^{1/\rho} \left(\alpha_m y_s^\rho + (1-\alpha_m) y_d^\rho\right)^{1/\rho}. \tag{21}$$

Where $y \equiv \mu q$ are total transactions which —since the model abstracts from inventories— are equivalent to GDP. Written this way, the matching function takes as inputs production capacity or supply $y_s \equiv fq$, and households demand $y_d \equiv hq$. This is convenient because below I find empirical counterparts to y_s and y_d.

In equilibrium $y_d = p_m m + y$ thus y_d is constructed using data on money and GDP.[41] I take m to be M1. Later it is shown how the model compares to broader measures of money such as M2 or MZM. y_s is constructed through GDP and data on total

[41] $y = \phi f q$ and firms' maximization imply $y = wn + kr$. Then from Equation (3), $y_d \equiv pq = p_m m + y$. Furthermore, since $\psi = y/y_d$, it is possible to construct $\psi = (p_m m/y + 1)^{-1}$.

capacity utilization (τ^{CU}): the percentage of total available capacity being used to produce demanded finished products. This matches closely with what ϕ means in the model. In particular, the literature on capacity utilization measures output as

$$y = (\tau^{CU} k)^{\alpha} n^{1-\alpha}. \tag{22}$$

Since $\tau^{CU} \in [0, 1]$, total production capacity y_s is obtained putting $\tau^{CU} = 1$ in Equation (22). Then $\phi = y/y_s = (\tau^{CU})^{\alpha}$, and y_s can be backed out as y/ϕ.[42]

Having constructed ϕ, ψ and $\theta = y_s/y_d$, it is possible to check whether they behave consistently with the novel matching process. In particular, ψ should be increasing and ϕ should be decreasing in θ. Notice that the matching function has not been used to construct these variables so these properties do not hold by construction. However, Figure 2 shows that the data line up with this matching theory rather well.[43]

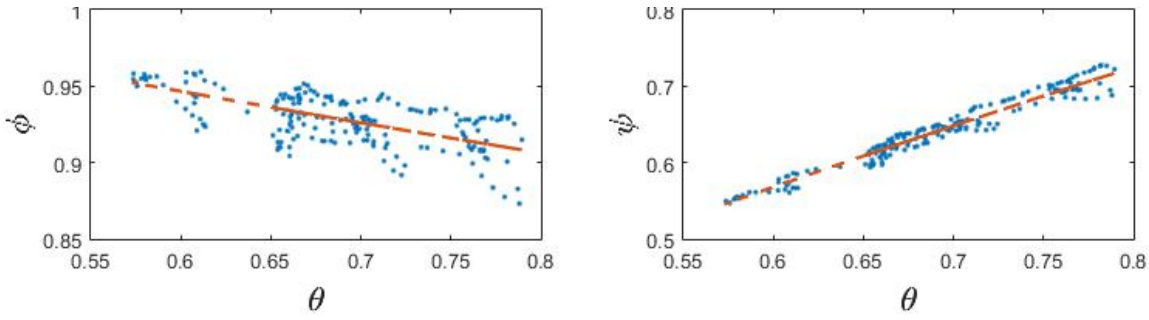

Figure 2: Empirical ϕ and ψ as a function of θ. Note: the trend is the prediction of the matching function with the estimated parameters α_m and ρ, and with steady state z_m.

Next, I parameterize the rest of the model and study its business cycle implications.

5.2 Parametrization

Parameter values are summarized in the table in Appendix D. I focus on quarterly data from 1967.Q1 (when data on total capacity utilization start) to 2016.Q1.

[42] Since the mapping between data on capacity utilization and the notion in this model may not be perfect, I experimented perturbing τ^{CU} with measurement error (40% of its variance) and found negligible differences. What matters is that capacity utilization is procyclical, which would be arguably true of other measures of ϕ.

[43] To appreciate that this result was not obvious: suppose that y, y_d and y_s were positively correlated (as they indeed are) but changes in y were in general smaller than changes in y_d, which in turn were smaller than changes in y_s. Then θ and ψ would have been negatively correlated, inconsistently with the predictions of the matching function.

The utility function is: $u = \log(c) - \chi_n \frac{n^{1+1/\nu_n}}{1+1/\nu_n} - \chi_d \frac{d^{1+1/\nu_d}}{1+1/\nu_d}$.

As a credit limit I assume $B = wn + rk$. This seems a natural benchmark as it is the implicit assumption in the neoclassical model, where a tighter limit would induce a cash in advance constraint. This assumption also makes B procyclical which is consistent with the notion that credit conditions deteriorate in recessions. Subsection 5.4 discusses the implications of this assumption further.

I assume z_m, A_d, A, χ_n and β to be AR1 stationary independent stochastic processes. I estimate the persistence and innovation variances for each stochastic process, the Frisch elasticities of labour and effort supply ν_n and ν_d, and the complementarity of the matching function ρ. A detailed explanation for the priors and posteriors is offered in Appendix F. The remaining parameters are calibrated, the choice of targets is for the most part standard and it is detailed in Appendix E.

The observables in the Bayesian Estimation are the growth rates of consumption, market hours, real GDP per capita, and the mentioned time series of capacity utilization and money-output ratio. Intuitively, consumption and GDP elicit the discount factor shock. Consumption and hours help identify the labour supply shock and the Frisch elasticity. Capacity utilization and the money-output-ratio imply ϕ and θ, which elicit the process for z_m and the matching complementarity ρ. Finally, GDP and θ imply y_d and thus effort via Equation (4); this disciplines the effort supply shock and Frisch elasticity.

5.3 Variance Decomposition

To appreciate how the model accounts for business cycle volatility through each shock, the variance decomposition of some variables of interest is reported in Table 1.

Table 1: Variance decomposition

	GDP	C	N	ϕ	Y_d	Y_s	Velocity
z_m	0.3761	0.1708	0.0537	0.6815	0.0199	0.0214	0.4923
A_d	0.2816	0.0658	0.2486	0.1934	0.8742	0.1038	0.3057
A	0.2218	0.2910	0.0088	0.0839	0.0533	0.5923	0.1326
β	0.0352	0.2923	0.1157	0.0008	0.0409	0.0482	0.0013
n	0.0847	0.1817	0.5695	0.0387	0.0161	0.2384	0.0611

First, the matching and the effort shocks alone, (z_m and A_d in the first two rows), explain 38+28 or 66% of the variance of GDP, 80% of the variance of velocity $\equiv$

$y/(p_m m)$, and 87% of that of ϕ, i.e. the movements in the Solow residual that are endogenous and not due to the technology shock A.

To disentangle the role of each shock it is useful to highlight that A_d is the most important shock for aggregate demand y_d (5th column) whereas the technology shock A is the most important one for aggregate supply y_s (6th column). In this sense A_d may be interpreted as a demand shock (similarly to Bai et al. (2017)). Instead z_m has little effect on both aggregate demand and supply; z_m is more like an intermediation shock that affects the matching of demand and supply. As it is usual with matching functions, the exact underlying frictions are not explicit. But a matching shock may be interpreted as capturing a more cautious behavior due to frictions such as information, screening, monitoring, agency and retail costs.

The working paper version Mennuni (2017) also reports impulse response functions of z_m and A_d. The main take away points are that an expansionary shock to both z_m and A_d induce increases in ϕ i.e. an endogenous surge in the Solow residual $y/\left(k^{\alpha}n^{1-\alpha}\right)$. This induces the usual real business cycle implications that—consistently with the data—there is comovement of hours, consumption and real input prices with output.[44] So far the two shocks are essentially isomorphic, however, they are distinguished in that a z_m shock causes a surge in velocity (therefore, after a negative shock velocity declines as in many recessions) and a drop in p_m, *i.e.* inflation is procyclical. Instead, an increase in A_d causes a decline in velocity. Intuitively, A_d does not affect the matching efficiency as z_m does, but changes the way people search, i.e. market tightness, thereby causing an increase in ϕ, but a decrease in ψ, with a resulting increase in money demand, or equivalently, a drop in velocity.[45]

As a result, z_m is key in recent recessions from the 80s onward, all characterized by a liquidity surge, or decline in velocity. This is illustrated for the case of the

[44] In future one could distinguish the matching function between consumption and investment. But absent further bells and whistles, if the matching shocks were not correlated, the impulse response to each shock would not make consumption and investment co-move. A case for an aggregate matching function is that consumption and investment products are intertwined. For instance, both consumption and investment often come with credit and insurance contracts, so a shock to, say, the ability to sell financial products would affect both consumption and investment.

[45] So velocity is a key observable to tell these two shocks apart. Bai et al. (2017) estimate a model with demand shocks. They do not have velocity and they do not distinguish between demand shocks and z_m. However, like here they find their wedge to be important. A demand shock (an A_d shock or equivalently an effort supply shock), has similar effects when modelling effort as in Bai et al. (2017). In particular, irrespective of the specification, as long as a positive demand shock increase matches and is expansionary, it will cause a drop in velocity.

great recession in Appendix G, Figure 4, which shows a peak to trough analysis from 2007.IV onward when including only one shock at the time versus the baseline path with all shocks. In particular, z_m shocks account for virtually the entire drop in output and ϕ (thereby generating an endogenous drop in productivity), as well as the increase in liquidity.[46] The role of z_m can also be appreciated by directly looking at the time series for z_m in Figure 5, closely related to velocity and τ^{CU} in all recessions in which velocity declines, i.e. all those from the 80s onward.[47] Instead the demand shock A_d played a negligible role in these recessions characterized by a liquidity surge. However, in earlier recessions (those started in 1969.IV and 1973.IV) there was no surge in liquidity as is illustrated in Figure 6 for the 1973.IV recession. For recessions where liquidity does not increase, a combination of z_m and A_d shocks is necessary: both push ϕ down, but they neutralize each other for liquidity as is evident from Figure 6, last panel.[48] The A_d shock is also behind the increase in θ (south-west panel), which did not occur during the Financial crisis.[49,50]

According to this model, technology shocks played a negligible role for recessions. One reason is that A makes ϕ (or τ^{CU}) countercyclical, which is counterfactual.

Finally, the model explains 43% of the variance of hours without labour supply shocks. While hours movements still require a strong ad hoc labour supply shock, neoclassical models with a similar calibration (low Frisch labour supply elasticity) explain about 10% of the variance of hours, see Ríos-Rull et al. (2012). To appreciate the propagation mechanism, it is instructive to rearrange the first order conditions (24), (25), (26) and (30) in Appendix to

$$-u_n = \left(u_c\psi - \frac{u_d}{A_d} + \lambda_4(1-\psi)\right) w, \tag{23}$$

[46]The model abstracts from monetary policy intervention, which may have contributed to the increase in liquid assets. However, the last panel in the figure includes a vertical line at the time of the first round of Quantitative Easing to highlight that the liquidity surge already took place. That a liquidity surge also characterized earlier recessions corroborates the view that such surge is not just an artifact of Quantitative Easing. Williamson (2015) argues that it is not obvious if and how a swap of assets between the central bank and private financial institutions affects money supply.

[47]To show where the identification of z_m comes from, the figure plots the time series identified with the Kalman Filter with no measurement error, and the one identified through Matching Function (21) given data on y, y_s, y_d and the estimated α_m and ρ: the two are virtually identical.

[48]The model abstracts from oil shocks which might cause a drop in τ^{CU} and an increase in prices, see Finn (1996). The increase in prices might prevent velocity from declining. Given the emphasis on the post 80s period, the oil channel is neglected.

[49]A_d and z_m combined also help account for the post great recession period characterized by a declining velocity due to a declining z_m and an recovering τ^{CU} (or ϕ) due to an increasing A_d.

[50]See the working paper Mennuni (2017) for peak to trough figures at all NBER recession dates.

which for simplicity, abstracts from the corner multiplier on q. There is a wedge relative to the neoclassical labour supply equation $-u_n = u_c w$. Intuitively, the benefit from working is not just $u_c w$, but there is the added issue that goods have to be found which reduces incentives to work. This wedge is procyclical which makes hours more volatile: the correlation between detrended GDP and the wedge (equal to $-u_n - u_c w$) constructed simulating the model with the identified shocks is 0.82. However, the model dampens the hours response to technology shocks, which is not significantly different from zero, and with negative mode. A factor that contributes to this is that labour demand increases much less than in the neoclassical model because ϕ drops.[51] Furthermore, households do not turn the wage into goods 1 for 1 but $(1-\psi)$ goes in money, and the gap between the return of capital and that of money widens after a technology shock; this reduces labour supply relative to the neoclassical model. On the other hand ψ and d increase after a technology shock. Other things equal this increases labour supply, but not enough to overcome the other effects.

5.4 Broader monetary aggregates

In this model credit $B = wn + rk$ is equal to inside money (a liability of the private sector used as a medium of exchange, see Lagos 2008) so inside and outside money together amount to $p_m m + B$. How does this compare to broader monetary aggregates not used in the estimation? Figure 3 below plots the time series of M2 and MZM over GDP, and the model simulation of $p_m m + B$ over output with the identified shocks.

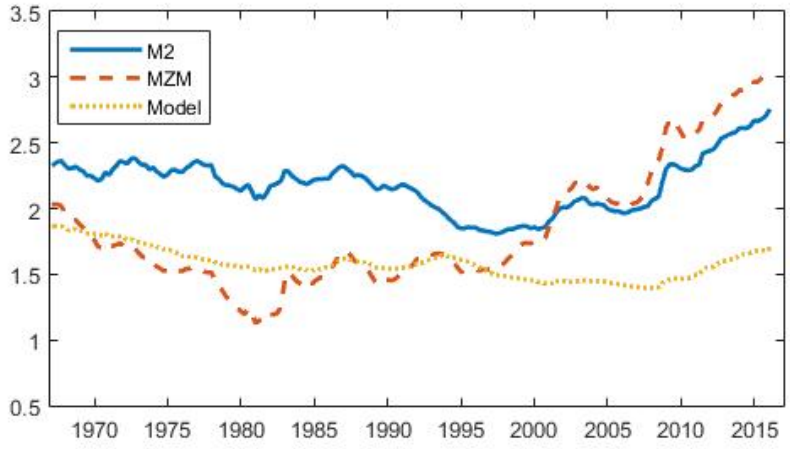

Figure 3: M2, MZM, and model simulation of Broad Money.

[51] Intuitively, demand bites more after a positive technology shock, akin Galí (1999).

On average the model implies less liquid assets than these other measures but it does not seem far off.[52] M2 and MZM are also more volatile than the model counterpart; this suggests that credit is more sensitive to the business cycle than $B = wn + rk$ and in future it might be fruitful to introduce financial frictions in this framework.

6 Conclusions

Motivated by the observed amounts of monetary holdings in times when several means of payment do not require to hold liquid funds, this paper developed a theory of liquid assets as a store of value due to a search friction between buyers and sellers. In particular, monetary equilibrium is consistent with large availabilities of credit, which is key to match the levels of broad monetary quantities we observe.

Money also enhances productivity and welfare: when people are less worried about not finding goods because they have easy alternative means to store value, they search for better deals (lower prices but longer queue length) making firms more productive. Put differently, the presence of money increases aggregate demand relative to aggregate supply. These implications are very different than those of a cash-in-advance, or other monetary set-ups.

By linking demand, supply, and the value of money, the search friction is also a natural source of the business cycle. A shock to the matching function emerges as an important source of recessions, while generating a surge in liquidity and spare production capacity: the paper documents these two patterns (surge in liquidity and drop in capacity utilization) for the financial crisis and several earlier recessions. Furthermore, in line with business cycle accounting, changes in the matching efficiency also induce a TFP wedge and, to some degree, a labour supply wedge.

While the story has elements of popular narratives, the framework is novel and could have many uses. For instance, here liquidity can be extended to a larger set of assets differing in their liquidity (captured by the severity of the search friction which stands for differences in information acquisition costs, risk, maturity etc.) reflect-

[52] It should be stressed that to get close to these broader aggregates while disciplining m through M1 is only possible because the model reconciles monetary equilibria with large credit. In fact, it would be possible to make B even larger and have more inside money. For instance, given Corollary 1, to the extent that $B < L$, it may be possible to put $B = z(wn + rk)$ with $z > 1$ and stochastic so that $m = M1$ and $p_m m + B$ matches the real value of M2 or MZM. Without a substantive theory of endogenous credit limits, this seems beyond the scope of this paper. In fact, to be below average seems reasonable given that the theory abstracts from other motives of money demand.

ing empirical counterparts ranging from government bonds, equity shares and other financial products, to possibly far less liquid assets such as houses.[53]

Furthermore, it is possible to embed more finance in this model to address further positive and normative questions. In particular this framework may provide a rationale for why policies aimed at easing credit conditions may not always work: in this model the drop in lending is not due to credit constraints but to the lack of "appetite" from the private sector. This implication stands in contrast to models with credit constraints, which are relaxed by quantitative easing policies as shown in Kiyotaki and Moore (2012). While a debate between the two channels may prove healthy, the two can be studied jointly in this framework. In fact the model shows theoretically when credit limits are neutral.

The framework may also prove useful to study fiscal policy, for instance: being TFP endogenous and affected by the demand-supply ratio, government spending could increase TFP. Furthermore, the money hoarding behaviour during recessions is akin to the keynesian liquidity trap.

Finally, it is well known that a lot of liquidity is held by firms and corporations. It is easy to envisage extensions of this model where households buy consumption goods while heterogenous firms trade capital subject to severe search frictions reflecting low arrival rates of big investment opportunities such as takeovers, thereby generating large holdings of liquid assets. In this context, households and firms may also prefer different types of assets reflecting their different needs for liquidity.

References

Aiyagari, S. R. (1994, August). Uninsurable idiosyncratic risk and aggregate savings. *Quarterly Journal of Economics 109*(3), 659–684.

Aliprantis, D. C., G. Camera, and D. Puzzello (2007, October). Anonymous markets and monetary trading. *Journal of Monetary Economics 54*(7), 1905–1928.

Alvarez, F. and F. Lippi (2009). Financial Innovation and the Transactions Demand for Cash. *Econometrica*.

[53] An attractive feature is that liquidity premia are endogenous: agents choose assets trading off their liquidity and their return so that in equilibrium the more liquid the asset the lower its return.

Aruoba, B. and F. Schorfheide (2011, January). Sticky prices versus monetary frictions: An estimation of policy trade-offs. *American Economic Journal: Macroeconomics* (3), 6090.

Bai, Y., J.-V. R. Rull, and K. Storesletten (2017). Demand Shocks that Look Like Productivity Shocks. *Mimeo*.

Berentsen, A., G. Camera, and C. Waller (2007, July). Money, credit and banking. *Journal of Economic Theory 135*(1), 171–195.

Bernanke, B. and M. Gertler (1989, March). Agency Costs, Net Worth, and Business Fluctuations. *American Economic Review 79*(1), 14–31.

Bewley, T. (1980). The optimum quantity of money. In J. H. Kareken and N. Wallace (Eds.), *Handbook of Macroeconomics*, Models of Monetary Economies, pp. 49–82. Minneapolis: Federal Reserve Bank of Minneapolis.

Brunnermeier, M. K. and Y. Sannikov (2014, February). A macroeconomic model with a financial sector. *American Economic Review 104*(2), 379–421.

Brunnermeier, M. K. and Y. Sannikov (2016, August). The I Theory of Money. CEPR Discussion Papers 11444, C.E.P.R. Discussion Papers.

Cavalcanti, R. d. O. and N. Wallace (1999, August). Inside and Outside Money as Alternative Media of Exchange. *Journal of Money, Credit and Banking 31*(3), 443–457.

Chari, V., P. Kehoe, and E.-R. McGrattan (2007). Business cycle accounting. *Econometrica 75*(3), 781–836.

Christiano, L. J., R. Motto, and M. Rostagno (2014, January). Risk shocks. *American Economic Review 104*(1), 27–65.

Cui, W. and S. Radde (2016, May). Money and asset liquidity in frictional capital markets. *American Economic Review 106*(5), 496–502.

Den Haan, W. J. (2014, January). Inventories and the Role of Goods-Market Frictions for Business Cycles. Discussion Papers 1402, Centre for Macroeconomics (CFM).

Diamond, D. W. and P. H. Dybvig (1983, June). Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy 91*(3), 401–19.

Ferroni, F., S. Grassi, and M. A. Leon-Ledesma (2015, May). Fundamental shock selection in DSGE models. Studies in Economics 1508, School of Economics, University of Kent.

Finn, M. G. (1996). A theory of the capacity utilization/inflation relationship. *Economic Quarterly* (Sum), 67–86.

Galí, J. (1999, March). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review 89*(1), 249–271.

Gorton, G. and G. Pennacchi (1990). Financial intermediaries and liquidity creation. *Journal of Finance 45*(1), 49–71.

Gu, C., F. Mattesini, and R. Wright (2016). Money and credit redux. *Econometrica 84*(1), 1–32.

Hodrick, R. J., N. R. Kocherlakota, and D. Lucas (1991, April). The Variability of Velocity in Cash-in-Advance Models. *Journal of Political Economy 99*(2), 358–84.

Holmstrom, B. and J. Tirole (1998, February). Private and Public Supply of Liquidity. *Journal of Political Economy 106*(1), 1–40.

Huo, Z. and J.-V. Ríos-Rull (2013, September). Paradox of Thrift Recessions. Nber working papers, National Bureau of Economic Research, Inc.

Iacoviello, M. (2015, January). Financial Business Cycles. *Review of Economic Dynamics 18*(1), 140–164.

Jermann, U. and V. Quadrini (2012, February). Macroeconomic effects of financial shocks. *American Economic Review 102*(1), 238–71.

Kaplan, G. and G. Menzio (2016). Shopping Externalities and Self-Fulfilling Unemployment Fluctuations. *Journal of Political Economy 124*(3), 771–825.

Kaplan, G. and G. L. Violante (2014). A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica*.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

Kiyotaki, N. and J. Moore (1997, April). Credit Cycles. *Journal of Political Economy 105*(2), 211–48.

Kiyotaki, N. and J. Moore (2012, March). Liquidity, Business Cycles, and Monetary Policy. NBER Working Papers 17934, National Bureau of Economic Research, Inc.

Kiyotaki, N. and R. Wright (1989, April). On Money as a Medium of Exchange . *Journal of Political Economy 97*(4), 927–954.

Kocherlakota, N. (1998). Money is memory. *Journal of Economic Theory 81*(2), 232–251.

Lagos, R. (2008). *Inside and Outside Money*. The New Palgrave Dictionary of Economics. Palgrave Macmillan.

Lagos, R. and G. Rocheteau (2008, September). Money and capital as competing media of exchange. *Journal of Economic Theory 142*(1), 247–258.

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

Ljungqvist, L. and T. J. Sargent (2012). *Recursive Macroeconomic Theory*. MIT Press.

Lucas, Robert E, J. and N. L. Stokey (1987, May). Money and Interest in a Cash-in-Advance Economy. *Econometrica 55*(3), 491–513.

Lucas, R. J. and J. Nicolini (2015). On the stability of money demand. *Journal of Monetary Economics 73*, 48 – 65.

Mennuni, A. (2017). Liquid Accounts as a Store of Value and Excess Capacity. Discussion Papers in Economics and Econometrics 1704, University of Southampton.

Menzio, G., S. Shi, and H. Sun (2013). A monetary theory with non-degenerate distributions. *Journal of Economic Theory 148*(6), 2266 – 2312.

Mill, J. S. (1844). Essay ii. of the influence of consumption upon production. In *Essays on Some Unsettled Questions of Political Economy*. London: Parker.

Moen, E. (1997). Competitive Search Equilibrium. *Journal of Political Economy 105*(2), 385–411.

Mortensen, D. T. and C. A. Pissarides (1994, July). Job Creation and Job Destruction in the Theory of Unemployment. *Review of Economic Studies 61*(3), 397–415.

Petrosky-Nadeau, N. and E. Wasmer (2011, June). Macroeconomic Dynamics in a Model of Goods, Labor and Credit Market Frictions. IZA Discussion Papers 5763.

Ragot, X. (2014). The case for a financial approach to money demand. *Journal of Monetary Economics 62*(C), 94–107.

Ríos-Rull, J.-V., F. Schorfheide, C. Fuentes-Albero, M. Kryshko, and R. Santaeulàlia-Llopis (2012). Methods versus substance: Measuring the effects of technology shocks. *Journal of Monetary Economics 59*(8), 826–846.

Shi, S. (1997, January). A Divisible Search Model of Fiat Money. *Econometrica 65*(1), 75–102.

Shi, S. (2012, July). Liquidity, Assets and Business Cycles. Working Papers tecipa-459, University of Toronto, Department of Economics.

Svensson, L. E. O. (1985, October). Money and Asset Prices in a Cash-in-Advance Economy. *Journal of Political Economy 93*(5), 919–944.

Telyukova, I. A. (2013). Household Need for Liquidity and the Credit Card Debt Puzzle. *Review of Economic Studies 80*(3), 1148–1177.

Telyukova, I. A. and L. Visschers (2013). Precautionary money demand in a business-cycle model. *Journal of Monetary Economics 60*(8), 900–916.

Wallace, N. (1980). The overlapping generations model of fiat money. In J. H. Kareken and N. Wallace (Eds.), *Handbook of Macroeconomics*, Models of Monetary Economies, pp. 49–82. Minneapolis: Federal Reserve Bank of Minneapolis.

Wang, W. and S. Shi (2006, April). The variability of velocity of money in a search model. *Journal of Monetary Economics 53*(3), 537–571.

Wen, Y. (2011, October). Input and Output Inventory Dynamics. *American Economic Journal: Macroeconomics 3*(4), 181–212.

Wen, Y. (2015). Money, liquidity and welfare. *European Economic Review 76*(C), 1–24.

Williamson, S. D. (2015, July). Current Federal Reserve Policy Under the Lens of Economic History: A Review Essay. Working Papers 2015-15, Federal Reserve Bank of St. Louis.

Woodford, M. (1998, January). Doing Without Money: Controlling Inflation in a Post-Monetary World. *Review of Economic Dynamics 1*(1), 173–219.

Appendixes

A First order conditions of the household

Let $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$ be the lagrange multipliers on the constraints (3)—(6) and let $\lambda_{k'}, \lambda_{m'}, \lambda_{a'}, \lambda_q, \lambda_p \geq 0$ be the multipliers respectively on $k' \geq 0$, $p_m m' \geq 0$, $a' \geq \underline{a}$, $q \geq 0$, $p \geq 0$, with complementary slackness between each multiplier and the respective constraint. The households first order condition for c, n, d, k',m',a',q,p are

$$u_c = \lambda_3, \tag{24}$$

$$u_n = -\lambda_4 w - \lambda_1 B_n, \tag{25}$$

$$u_d = -\lambda_2 A_d, \tag{26}$$

$$\lambda_3 - \lambda_{k'} = \beta E\left(\lambda_4' r' + \lambda_3'(1-\delta) + \lambda_1' B_k'\right) \tag{27}$$

$$\lambda_4 p_m - \lambda_{m'} = \beta E\left((\lambda_1' + \lambda_4') p_m'\right) \tag{28}$$

$$\lambda_4 v - \lambda_{a'} = \beta E\left(\lambda_1' + \lambda_4'\right) \tag{29}$$

$$(\lambda_3 - \lambda_4 p)\psi = \lambda_1 p + \lambda_2 - \lambda_q \tag{30}$$

$$\lambda_1 + \lambda_4 \psi = \frac{\partial \psi}{\partial \theta}\frac{\partial \theta}{\partial p}(\lambda_3 - \lambda_4 p) + \lambda_p / q, \tag{31}$$

B Proofs

Lemma 1

Pick p, q such that Equation (13) holds. Now suppose $\theta(p,q)$ depended on q. Then it would be possible to change q holding p constant so that $\phi(\theta(p,q))$ increases. But from Equation (12) this makes profits positive; then the assumed $\theta(p,q)$ was not profit maximizing.

Proposition 1

It is first shown that if $\hat{b} = L$ then $p_m = 0$. From (3) with equilibrium $a = 0$, if agents borrow L then $pq = (p_m m + wn + rk)/\psi$. The latter implies that the right hand side of (6) is zero. Then the left hand side must be zero too. Since $a' = 0$, this requires $p_m = 0$.

It is now shown that if $p_m = 0$ then $\hat{b} = L$. From Constraint (6), $pq \leq (wn + rk)/\psi$ (because $p_m = 0$ and $a = a' = 0$ in equilibrium). $pq < (wn + rk)/\psi$ violates Equilibrium Condition (14) and zero profits, so $pq = (wn + rk)/\psi$. The latter and the definition of $\hat{b}$ imply $\hat{b} = (wn + rk)/\psi$ (which with $p_m = 0$ is equal to L).

To see that L is the maximum implementable credit notice that borrowing more than L would violate Market Clearing Condition (14). This is because $q > 1/\psi(p_m m + wn + rk)$, $p_m m \geq 0$ and the zero profits condition $\phi f q = wn + rk$, imply $\psi q > \phi f q$. [54]

Corollary 1
An equilibrium has to have $p_m \geq 0$ and $\hat{b} \leq L$. Since from Proposition 1 $p_m = 0$ iff $\hat{b} = L$, the intersection $p_m = 0$ and $\hat{b} < L$, and the intersection $p_m > 0$ and $\hat{b} = L$, are empty.

Lemma 2
The first order conditions to the households problem are reported in Appendix A. With the borrowing constraint not binding, Equation (3) is not binding, so the associated multiplier $\lambda_1 = 0$. With no effort costs also Effort Constraint (4) is not binding so that $\lambda_2 = 0$: this is immediate from First Order Condition 26. Then Condition (30) implies $\lambda_3 = \lambda_4$ (this requires interior p and q which is guaranteed by Equation 13 and Inada conditions on the utility function). Then for any positive p_m, Euler Equations for k' and m' —(27) and (28)— imply $\lambda_{m'} > 0$ which means $m' = 0$. Hence p_m cannot be positive as it violates $m' = M$.

In the case with effort cost $\lambda_2 > 0$, so the argument above does not work. However, in steady state λ_4 has to be constant from (25) and (27). Then (28) in steady state implies $\lambda_{m'} > 0$ for any inflation up to the Friedman Rule limit $p_m/p'_m \to \beta$. Then rolling back (28) without uncertainty gives $p_m = 0$ on the entire transition path (a positive p_m would need $\lambda_{m'} > 0$ which would make $m' = 0$, violating market clearing). Notice that this argument requires there to be no risk as otherwise the fact that $p_m = 0$ in some future state does not imply that the right hand side of (28) is equal to zero.

Corollary 2
The proof consists in showing the following claim:

Claim 1 *in equilibrium the borrowing constraint can only bind if the credit limit is lower than the natural limit: $B < L$.*

Then the result follows because from Proposition 1, if $\hat{b} < L$, money has value. To show Claim 1, suppose instead that $B > L$ was binding in equilibrium. Then agents would borrow $\hat{b} = B$. But then Equation (6) can only be satisfied with $va' + p_m m' < 0$. Since $m' \geq 0$ this requires $a' < 0$, but then $v = 0$ as in equilibrium $a' = 0$. Then the only option would be to default on $\hat{b}$. Since agents are not allowed to default, they must choose $\hat{b} < B$.[55]

[54] An alternative way to show that borrowing cannot be greater than L is that (3) and (6) imply that it would be impossible to repay such debt without aggregate intertemporal $a' < 0$.

[55] This implies that even when $B = L$ and they *choose* $\hat{b} = B$ the constraint is not binding: if B was relaxed, $\hat{b}$ would not increase.

Proposition 2

With $p_m = 0$ firms' first order conditions, the fact that the production technology has constant returns to scale, and market clearing for the capital and labour imply that

$$\phi f A k_d^{\alpha} n_d^{1-\alpha} = \phi A k^{\alpha} n^{1-\alpha} = wn + rk. \tag{32}$$

If $p_m = 0$, from Equation (3) $q = wn + rk$, which combined with Equation (32) implies

$$q = \phi A k^{\alpha} n^{1-\alpha}. \tag{33}$$

Equation (14), the capacity constraint (9), and the fact that $f A k_d^{\alpha} n_d^{1-\alpha} = A k^{\alpha} n^{1-\alpha}$, imply

$$\psi q = \phi A k^{\alpha} n^{1-\alpha}. \tag{34}$$

Equations (33) and (34) then imply $\psi = 1$ and, through the matching function, $\theta = \infty$ and $\phi = 0$. Since production is bounded, $\phi = 0$ implies that no goods are sold.

Proposition 3

To show neutrality, take an equilibrium allocation with constant money supply $m > 0$. Let p_m be the equilibrium function. It is possible to change the money supply to zm with $z > 0$ and pick a new price function $p_m^z = p_m/z$ so that all equilibrium conditions are satisfied with the same allocation.[56]

Superneutrality does not hold because the Euler equation (28) depends on p'_m/p_m which is affected by a change in money growth. Thus the inflation rate affects the dynamics of the Lagrange multipliers λ_1 and λ_4 defined in Appendix A. It follows trivially from the other first order conditions that the allocation is also affected.

Proposition 4

Change p_m to $\hat{p}_m$ so that,

$$\hat{p}_m m + zB = p_m m + B \tag{35}$$

then all equilibrium conditions are satisfied with the steady state allocation associated to B. To see this notice that the equations where the two variables that change (p_m and or B) appear are Equations (3), (6), and (28). Equation (3) is clearly satisfied with all other variables unchanged. (6) is not affected because in equilibrium $p_m(m + d_m) = p_m m'$ for all p_m, so they cancel out from the right and left hand side. Equation (28) is not affected

[56] The equilibrium conditions in which money or p_m appear are Equations (3) and (6), and the Euler equation for m', (28) in Appendix A, which must hold in an equilibrium. Since $p_m^z zm = p_m m$, (3) and (6) are satisfied with the original allocation. In a monetary equilibrium (where $p_m > 0$ and $\lambda_{m'} = 0$) Equation (28) can be rearranged so that prices enter as a ratio p'_m/p_m, but this ratio is equal to ${p_m^z}'/p_m^z$. It follows that the original allocation satisfies these conditions.

because a level change in steady state B does not affect p'_m/p_m. To see this take a steady state with $m'/m = \pi$; it is easy to verify that $p'_m/p_m = 1/\pi$ and $B' = B$. Then using (35) it is easy to check that changing steady state B to zB does not affect the inflation rate:

$$\hat{p}'_m/\hat{p}_m = (p'_m + B'/m'(1-z))/(p_m + B/m(1-z)) = \\ (1/\pi p_m + B/(\pi m)(1-z))/(p_m + B/m(1-z)) = 1/\pi.$$

This is in general not true outside the steady state because B is not constant.

Proposition 5

The proof consists of showing that when $\frac{p'_m}{p_m} \to \frac{1}{\beta}$, the first order conditions necessary for a solution to the household problem (in Appendix A), are identical to those of the planner in steady state. It is then trivial to show that all other conditions are also identical.

In a monetary equilibrium $p_m m' \geq 0$ is not binding, thus the Kuhn-Tucker multiplier $\lambda_{m'}$ defined in Appendix A is equal to zero. Then the first order condition for m', Equation (28), implies $\lambda_1 = 0$ in steady state at the Friedman rule.

Since firms price condition, Equation (13), is true for all p, and using Lemma 1 (that θ is only function of p), one can differentiate Equation (13) with respect to p and get

$$\frac{\partial p}{\partial \theta} = -\frac{\partial \phi}{\partial \theta}\frac{p}{\phi}; \tag{36}$$

The properties of the matching function and Inada conditions on the utility function ensure that λ_p and λ_q are both zero. Then substituting $\frac{\partial p}{\partial \theta}$ from Equation (36) into Equations (30) and (31), normalizing the equilibrium $p = 1$, and noticing that $\psi = \theta\phi$, one gets

$$-\frac{\partial \psi}{\partial \theta}\frac{\partial \theta}{\partial \phi}\frac{\psi}{\theta}(\lambda_4 - \lambda_3) = \lambda_2 - \lambda_3\psi. \tag{37}$$

From Equations (30) and (37) with $\lambda_1 = 0$ and substituting ψ and $\frac{\partial \psi}{\partial \theta}$ from $\psi = \theta\phi$ (which implies $\frac{\partial \psi}{\partial \theta} = \phi + \theta\frac{\partial \phi}{\partial \theta}$) one gets

$$-\lambda_2 = \lambda_3\frac{\partial \phi}{\partial \theta}\theta^2. \tag{38}$$

This condition characterizes the decentralized choice of θ. Next, I obtain the same condition for the planner. The Planner first order conditions for θ and q are

$$\tilde{\lambda}_3\phi = \tilde{\lambda}_1 - \tilde{\lambda}_3\frac{\partial \phi}{\partial \theta}\theta, \quad \text{and} \quad \tilde{\lambda}_3\phi\theta = \tilde{\lambda}_1\theta + \tilde{\lambda}_2,$$

where $\tilde{\lambda}_1$, $\tilde{\lambda}_2$, and $\tilde{\lambda}_3$ are the Lagrange multipliers on constraints (17), (18), and (19).[57] The latter two conditions imply

$$-\tilde{\lambda}_2 = \tilde{\lambda}_3\frac{\partial \phi}{\partial \theta}\theta^2. \tag{39}$$

[57] $d \geq 0$ and $n \geq 0$ and Inada conditions ensure that the non negativity constraints on θ and q do not bind, so Kuhn-Tucker multipliers are not included for θ and q.

This planner condition coincides with the equilibrium condition (38) *iff* $\tilde{\lambda}_i = \lambda_i$ for $i = 2, 3$. It is trivial to verify that this is the case from the first order conditions of the household and of the planner for d and c. It is equally trivial to verify that the other equilibrium conditions and planner conditions are identical, which completes the proof.[58]

Corollary 3

I first show that $\theta > 0$, (and hence ϕ and $\psi \in (0, 1)$) when inflation is above the Friedman rule ($\frac{p'_m}{p_m} < \frac{1}{\beta}$). When inflation is above the Friedman rule, Equation (28) and steady state imply $\lambda_1 > 0$. Furthermore, with the marginal utility of effort $u_d = 0$, Equation (26) implies $\lambda_2 = 0$. Then, Equation (30) implies $\psi > 0$ and hence $\theta > 0$.[59]

I now show that when $\frac{p'_m}{p_m} \to \frac{1}{\beta}$ then $\theta \to 0$. From Equation (28) and steady state, when $\frac{p'_m}{p_m} \to \frac{1}{\beta}$, $\lambda_1 \to 0$. Then from Equation (30) $\psi \to 0$ and/or $\lambda_3 - \lambda_4 \to 0$. But from Equation (31), if $\lambda_3 - \lambda_4 \to 0$ then $\psi \to 0$ because from Equation (25) $\lambda_4 > 0$ as $\lambda_1 = 0$. Therefore $\psi \to 0$ and from the matching function $\theta \to 0$.

I now show the results about $p_m m$. It has been shown that $\theta > 0$ tends to zero at the Friedman rule. With total output positive and bounded, Equation (17) implies that $\theta \to 0$ if and only if production per trading post $q \to \infty$. Then with a bounded credit limit B, Equation (3) requires that for θ that tends to 0, $p_m m \to \infty$, and for $\theta \nrightarrow 0$, $p_m m$ bounded.

C Equilibrium

Definition 2 *Let $B(k, m, \Omega, \exists)$ denote a credit limit function where $\Omega = \{K, M, z_m, A_d, A, \chi_n, \beta\}$ with K and M denoting the aggregate capital and money stock. An equilibrium $\exists$ is composed of a value function V and policy rules c, k', n, d, m', a', q for the household as function of a, k, m, Ω, a function $\theta(p; \Omega)$ and prices w, r, p_m, v, measure of firms f, input demands per firm k_d and n_d, revenues per unit of production ξ, all functions of Ω, such that:*

1. *Household: The household's decisions and value function solve the problem in 3.1*
2. *Firms: k_d, n_d and $\theta(p; \Omega)$ satisfy Equations (10), (11), and (13) with $\xi = \phi(\theta(1; \Omega))$.*[60]
3. *The goods, the liquid assets, and the inputs markets clear as detailed in Section 3.3.*
4. *Aggregate and individual state variables are consistent: $k = K$, $m = M$.*
5. *$z_m, A_d, A, \chi_n, \beta$ follow Markovian processes of order one.*

[58] As usual, $p'_m/p_m = \beta$ violates the transversality condition for money because the value of money grows too fast but the limit of p'_m/p_m to β is implementable.

[59] In a monetary equilibrium $p, q, p_m m' > 0$ so $\lambda_p = \lambda_q = \lambda_{m'} = 0$.

[60] *The latter condition implies that the equilibrium p is normalized to 1.*

Notice that Equation (14) implies $\psi(\theta(1;\Omega))/\phi(\theta(1;\Omega)) = f$. Furthermore, the functions ψ and ϕ are such that $\psi(\theta)/\phi(\theta) = \theta$, so $\theta = f$. This is consistent with the definition of the market tightness $\theta = f/h$ because the measure of households is 1.

D Data and Table of Parameters

Nominal and Real GDP and consumption are taken from the NIPA Tables 1.1.5 and 1.1.6 of BEA. Consumption is defined as personal consumption expenditures on non-durables and services +government spending and net export. Hours per capita are constructed by dividing Hours by population taken from the Bureau of Labor Statistics. Hours, ID PRS85006033. Civilian Noninstitutional Population, ID LNU00000000. Velocity of M1, M2 and MZM, retrieved from FRED, Federal Reserve Bank of St. Louis. Capacity Utilization: Total Industry [TCU], retrieved from FRED, constructed by the Board of Governors.

		Prior			Posterior				
Name	Description	Density	Para(1)	Para(2)	Mode	Median	Std	[5 ,	95]
α	Capital income share	Calibrated	0.34						
δ	Capital depreciation	Calibrated	0.014						
β	Discount factor	Calibrated	0.99						
γ_a	TFP growth	Calibrated	1.0022						
α_m	matching function share	Calibrated	0.44						
χ_n	scaling labour supply	Calibrated	4.02						
χ_d	scaling effort supply	Calibrated	6.43						
ν	Frisch labour supply	Gamma	0.85	0.10	1.004	0.995	0.099	0.839	1.164
ν_d	Frisch effort supply	Gamma	0.85	0.10	0.711	0.731	0.091	0.593	0.891
$-\rho$	Matching Compl.	Gamma	2.00	1.0	1.573	1.384	0.188	1.086	1.707
			Persistence of shocks						
ρ_a	TFP	Beta	0.88	0.10	0.984	0.985	0.007	0.972	0.995
ρ_β	β	Beta	0.88	0.10	0.999	0.999	0.001	0.998	0.999
ρ_n	Labour supply	Beta	0.88	0.10	0.998	0.994	0.003	0.989	0.999
ρ_d	Effort productivity	Beta	0.88	0.10	0.999	0.997	0.003	0.990	0.999
ρ_{z_m}	Matching	Beta	0.88	0.10	0.969	0.970	0.008	0.956	0.983
			Std of shocks						
σ_a	TFP	Exp	0.20		0.651	0.652	0.034	0.599	0.713
σ_β	β	Exp	0.20		0.025	0.025	0.002	0.023	0.028
σ_n	Labour supply	Exp	0.20		1.316	1.334	0.103	1.183	1.519
σ_d	Effort product.	Exp	0.20		0.977	0.964	0.064	0.865	1.074
σ_{z_m}	Matching	Exp	0.20		0.712	0.628	0.091	0.486	0.786

Notes: Para (1) and Para (2) list the means and the standard deviations of the Gamma and Beta distributions. Para(1) indicates the rate parameter for the exponential distribution. For the parameters that are calibrated, Para (1) indicates the value. For the structural shocks, values in the last 5 columns are multiplied by 100.

Online Appendix

E Further Calibration Details

To make the model consistent with a balanced growth path with the observed mean growth rate of GDP, and stationary market hours and effort, A and A_d have to grow over time with $\gamma_d = \gamma_a^{1/(1-\alpha)}$, where γ_d is the deterministic growth factor of A_d and γ_a that of A. The other processes must have zero growth.

The steady state level for A can be normalized to one and that for A_d is set to match a steady state ratio of y_d and effort: similarly to market hours, there is no natural units for their measurement and I put both hours and effort equal to $1/3$ in steady state.[61] The steady state level for the stochastic discount factor β is 0.99. I also fix the depreciation rate $\delta = 0.014$ as in Aruoba and Schorfheide (2011) among others, and the capital income share $\alpha = 0.34$, conventional values in the DSGE literature.

The steady state level for the labour supply shock χ_n and for the effort parameter χ_d are piked to match the targeted steady state market hours and effort. To find α_m and the steady state level for z_m, I target steady state values for ϕ, the money output ratio $p_m m/y$, and the consumption output ratio c/y: it is possible to show that given an estimate for the matching function complementarity ρ, there is a unique value of α_m and z_m, that imply a steady state consistent with the above mentioned targets.

The target for c/y is 0.87, which is the sample average using personal consumption expenditures plus government spending and net exports over GDP. Steady state for ϕ is 0.93, the average of the time series constructed earlier, and for $p_m m/y$ is 0.57, this is the average of M1 over GDP.

F Priors

I assume a Gamma distribution for $-\rho$. This is because I restrict ρ to be smaller than zero. With $\rho \leq 0$ one cannot have sales with only aggregate demand, or only aggregate supply. Furthermore, as detailed in the working paper version Mennuni (2017), $\rho \leq 0$ ensures that the slope of marginal cost of choosing submarket with higher p is larger than that of the marginal gain, which is consistent with the objective function being increasing to the left of the crossing point and decreasing thereafter. I find a posterior mode for $\rho = -1.57$.

[61] For instance, it is possible to re-scale effort and change A_d with no effect on any other variable as it is clear from Equation (4).

I assume the Frisch elasticity of labour supply to be Gamma distributed with mean 0.85 and variance 0.1. This way the posterior mode is 1, in line with the micro studies surveyed by Keane (2011), who tend to find smaller elasticities relative to what macro models need to match hours volatility. I use the same prior for the Frisch elasticity of effort given the absence of external evidence. The posterior elasticity is 0.71. Doubling the variance for this prior leaves all substantive results unchanged.

Following the literature, the persistence parameters of the stochastic processes have a Beta prior and the standard deviation of the measurement error innovations follow an inverse Gamma prior. I set the prior of the measurement error so that the posterior measurement error variance does not exceed 1% of the variance of the observed series. The structural shocks standard deviations have an exponential prior as suggested by Ferroni et al. (2015). The covariance matrix of the innovations is diagonal.

G NBER Recessions

Figure 4 shows a peak to trough analysis by depicting a counterfactual path from 2007.IV onward when including only one shock at the time versus the baseline path with all shocks (which generates the exact data because shocks for these simulations are identified assuming no measurement error). The figure shows all the observables (ϕ is a monotone transformation of τ_{CU} so is essentially an observable) and θ, which is a function of the other observables and —as explained earlier— helps appreciate the presence of A_d shocks.[62]

[62]GDP, consumption and market hours are in logs and linearly detrended, so the figure shows the percentage deviation from a linear trend. Since the other variables are ratios, they are linearly detrended in levels. So if $p_m m/y$ goes from 0.1 to 0.2, the figure shows a 0.1 increase.

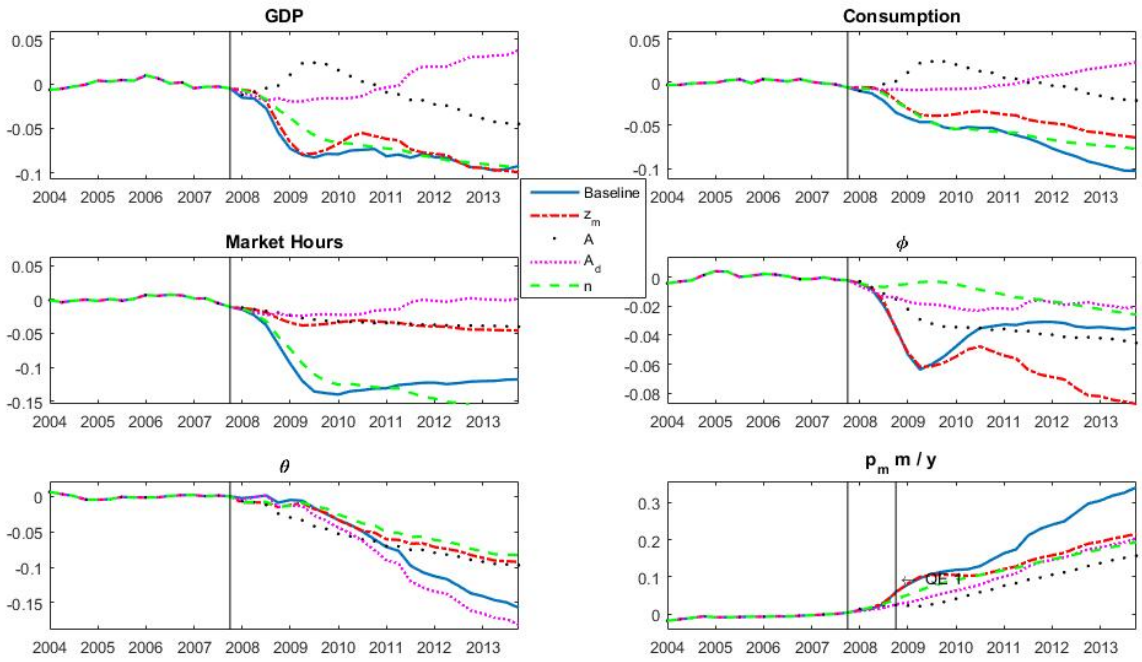

Figure 4: 2007.IV recession due to each shock. Note: The last panel includes a vertical line at the time of the first round of QE to highlight that the liquidity surge already took place.

The path with shocks on β is not shown because not important.

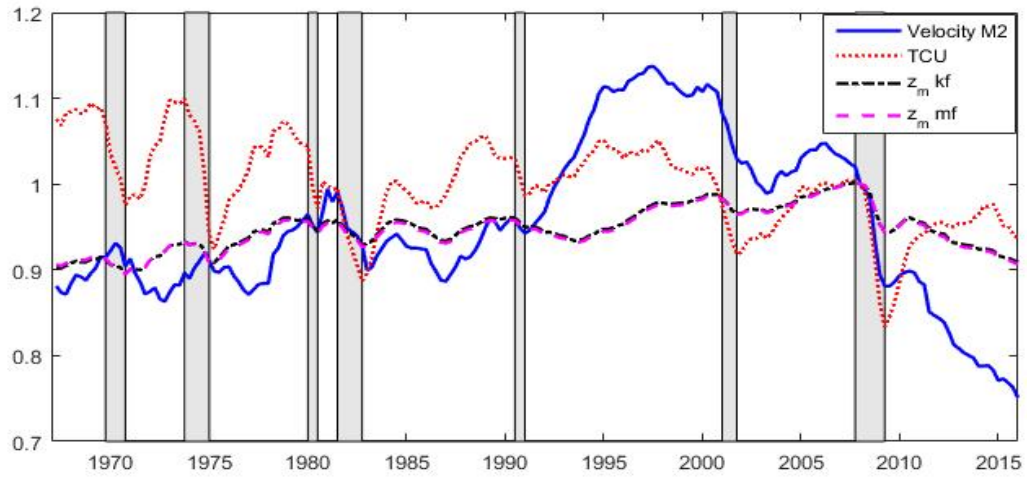

Figure 5: Time series of $z_m^{1/\rho}$ vs velocity and TCU. Notes: z_m kf is identified with the Kalman Filter with no measurement error, z_m mf is identified through the matching function given data on y, y_s, y_d and the estimated α_m and ρ. Since $\rho < 0$ the figure plots $z_m^{1/\rho}$: an increase in this variable is an increase in matching efficiency.

Figure 6 depicts a counterfactual path from 1973.IV onward

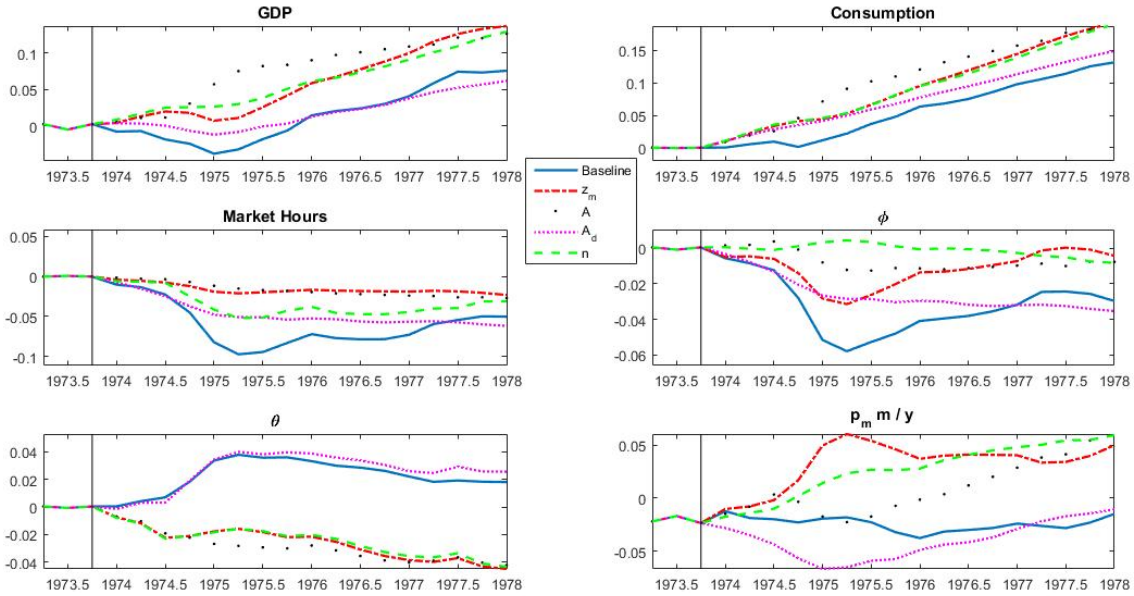

Figure 6: 1973.IV recession due to each shock