UNIVERSITY OF Southampton
School of Social Sciences

Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Liquid Accounts as a Store of Value and Excess Capacity

Alessandro Mennuni (University of Southampton)

No. 1704

ISSN 0966-4246

Liquid Accounts as a Store of Value and Excess Capacity*

Alessandro Mennuni
University of Southampton

July 25, 2017

Abstract

This paper introduces a macroeconomic framework where money emerges endogenously as a store of value because of a search friction in the goods market. With the novel microfoundation, monetary equilibria are robust to large amounts of credit. The model also implies a link between the value of money and the matching between the demand and supply of goods. As a result, the implications are different than those of existing models. In particular, the model offers an original interpretation of recessions which accounts for how monetary holdings and excess production capacity vary over the business cycle.

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, 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 and Queen Mary, Ensai, Global Macro at NYU Abu Dhabi, Macro Marrakesh, The Bank of England, The European Central Bank, for helpful comments.

`The use of esteemed articles as a store or medium for conveying value may in some cases precede their employment as currency [...] Such a generally esteemed substance as gold seems to have served, firstly, for ornamental purposes; secondly, as stored wealth; thirdly, as a medium of exchange.` Jevons (1875)

1 Introduction

This paper introduces a macroeconomic framework that emphasises the role of money as a store of value. This focus is motivated by the fact that a non negligible share of savings is held in liquid form. For instance, M2 over yearly GDP ranged between 0.45 and 0.70 since 1959 in the US. While most theories (rightfully) emphasise money demand for transactions, this motive alone may not explain all the liquidity holdings we observe in modern societies, especially given the vast availability of credit which crowds out transaction money demand. Other theories have emphasized the role of money as a store of value, but the mechanism through which money demand emerges here seems novel. Furthermore, the mechanism accommodates a new explanation for recessions which accounts for why they are often accompanied by spare production capacity and a liquidity surge: as shown later, this has occurred on several occasions, including the financial crisis.

The model is 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.[1]

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

[1] 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 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).

contrast in this paper, the transaction motive alone does *not* generate money demand (even without credit). This result relies on the assumption of reversible capital: that money can be more easily transacted with goods gives money no advantage when capital can be directly consumed at any time. Thus, without the portfolio problem, money would not be held. This is in contrast to LW where the presence of an ex-post centralized market (that removes the portfolio problem by enabling agents to readjust the goods-money portfolio), and the absence of a competing reversible asset, result in positive money holdings for transaction purposes. To be clear, both roles of money for transactions and as a store of value due to a mismatch are present in other models, such as 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. So in essence this paper departs from KW in a polar way to the one of LW. To focus on this other 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. Put differently, the theory exploits the fact that in search models there is an equilibrium excess demand and supply. In the goods markets demand consists of the funds available to the buyers. Since those are in excess, there is a role for a liquid store of value. An important result that stems from this microfoundation is that money has a role as long as there are some credit restrictions, albeit very loose: there is monetary equilibrium with credit above 100% of GDP. With the resulting creation of inside money, the model is qualitatively consistent with broad monetary aggregates such as MZM. The paper relates these findings to Gu et al. (2016) who study the coexistence of money and credit in a LW framework. The implications are also different than those of typical cash-in-advance models like Lucas and Stokey (1987).

The main implications of the model are that TFP, money velocity, and excess supply i.e. unsold production capacity, are all endogenous; furthermore, the matching friction emerges as a new source 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) where the probability of finding goods is lower for buyers, but higher for firms relative to the non monetary equilibrium.[2] So money is desirable because it increases firms' ability to

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

sell or measured productivity. However, the mechanism also uncovers 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."[3]

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 their wealth in unproductive but safe assets.

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.[4] That this liquidity surge is related to an excess supply capacity is consistent with the decline in capacity utilization during the Great Recession: Total Capacity Utilization (TCU) constructed by Federal Reserve Board declined from 81% in December 2007 to 66.7% in June 2009. Furthermore this pattern is not only true of the financial crisis: Figure 1 shows that in all recessions from the 80s onward, capacity utilization and various measures of velocity, dropped jointly.[5]

While the paper is mainly theoretical, a simple quantitative section shows how the model can be brought to the data and a structural estimation finds that the matching

presence of money increases the equilibrium value of aggregate demand.

[3]Here the excess supply of goods —defined as production capacity minus sales— is an equilibrium outcome given the search friction.

[4]Velocity is an inverse measure of monetary holdings, therefore a drop corresponds to an increase in monetary holdings relative to GDP. 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.

[5]Instead, the relationship was negative before the 80s. The model explains this changing pattern through a different combination of shocks. These correlations are not explained by changes in output, interest rates, or inflation: regressing both capacity utilization and velocity over the interest rate, inflation, and GDP growth does not explain these facts as the residuals of the 2 regressions exhibit similar patterns (positively correlated from the 80s and negatively correlated before). For example, the correlation in the residuals from the two regressions —one for capacity utilization and one for the velocity of M2— is 0.43 from the 80s and -0.54 before. Both strongly significant. Theoretically, to explain fluctuations in velocity has been a long standing challenge since Hodrick et al. (1991). While some progress has been made (see for instance Telyukova and Visschers 2013 and Wang and Shi 2006) Lucas and Nicolini (2015) argue that the interactions between money and financial crises remain poorly understood.

process plays an important role in understanding recessions.[6]

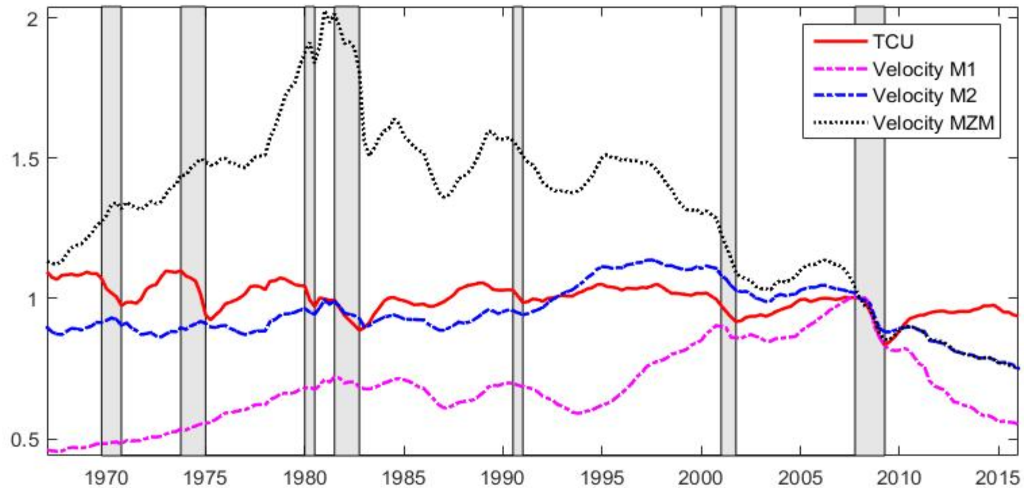

Figure 1: Total Capacity Utilization, and Velocity of M1, M2 and MZM. Source: FRED. Note: Time series are normalized to be 1 in 2008.

In particular, a wedge to the matching function generates an endogenous drop in TFP, a surge in liquidity, disinflation, a drop in capacity utilization, and a labour wedge which is qualitatively consistent with the business cycle accounting of Chari et al. (2007).[7] 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.[8] This shock is distinguished from a pure demand shock (a shock to search effort similarly to Bai et al. 2011), which is also present and played a role for the recessions prior to the 80s, explaining why velocity did not decline.[9]

[6] The structural estimation also highlights that the model is sufficiently tractable numerically, despite the microfoundation of money. By relaxing the assumption of anonymity, it is possible to have insurance markets and study a representative agent model grounded into a modified neoclassical setting which can be linearized. Thus the framework is amenable to further quantitative work.

[7] This labour wedge comes from the fact that income from working cannot be easily spent and makes market hours more volatile than comparable neoclassical business cycle models.

[8] While this shock can be seen as a "measure of our ignorance", the idea that the matching between demand and supply is the place where much of the business cycle is generated seems worth investigating. Indeed, the quantitative analysis also shows that the original matching theory of aggregate demand and supply lines up with empirical counterparts rather well. In traditional models the uncovered wedge is by and large imputed to shocks to the production technology while here the Solow residual is decomposed in conceptually different residuals.

[9] 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.

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, special cases, figures and tables.

2 Literature Review

Liquid assets are present in many other brunches of the literature. First, a microfoundation of fiat money typically interpreted as a store of value is offered by the overlapping generations model pioneered by Samuelson.[10]

In the Baumol-Tobin 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 Baumol-Tobin 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 that M2 alone is more than 2 times quarterly GDP: a long way to go with either the ad hoc cost, or the search friction taken on its own.[11]

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 dividend, 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).[12] Other related frameworks where

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

[11] Furthermore, 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.

[12] Alternatively, it is possible to have capital subject to uninsurable idiosyncratic shocks directly

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.[13] Furthermore, Telyukova (2013) reconciles the 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. (2011), Huo and Ríos-Rull (2013a), 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.[14] Here instead, firms do *not wish* to produce more because of their difficulties to sell.[15] 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 credit declines during recessions, however this does

so that it is more risky than money as in Brunnermeier and Sannikov (2016).

[13]While these mechanisms match the unconditional variance of velocity, it is unclear whether they generate a liquidity surge during recession of the observed magnitudes. Furthermore, these papers do not relate velocity to TFP, capacity utilization, and the labour wedge.

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

[15]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*.

not happen because of tighter credit but because of fewer spending opportunities.[16]

Finally, like the New Keynesian framework, this one may prove useful for policy analysis. An advantage is that it explains monetary quantities.[17]

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.

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.[18]

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.[19] 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

[16]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.

[17]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. Furthermore, they have important consequences for the identification and propagation of the sources of the business cycle.

[18]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.

[19]Terms of trade can be equivalently expressed as the balance pq and the quantity q exchanged. But indexing submarkets by (p, q) it will be more immediate to talk about inflation later.

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 in exchange of an IOU, a promise for later payment in money or by clearance of net IOU positions) at the end of the period. Without default being an option, the seller is indifferent between money or credit as long as credit pays the same return as money (1 within the period). 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 firms revenues are realized. After the input markets clear and before inputs are paid, households and firms make transactions in the frictional market. Households can pay with their initial liquid holdings, or with credit.

Of course, with enough credit, money looses 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).[20]

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 an intertemporal bond issued by households at the end of the previous period. With this bond households can roll over the IOUs issued to firms.[21]

[20] 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.

[21] 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

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 IOUs. Absent credit ($a = B = 0$) Equation (1) would be a pure cash in advance constraint. This way to introduce credit resembles Gu et al. (2016) (GMW). While others have introduced credit in different ways, drawing upon several specific friction which all have some element of truth (see Cavalcanti and Wallace (1999) and Berentsen et al. (2007) among others), for the sake of this study, this approach seems very clean in that, up to the borrowing constraint, 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.[22]

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 IOU's it received from other households.[23] Since the household spent pq (either in money or issuing IOUs), her end of period

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 D shows this result formally and it explicitly introduces a secondary market for capital.

[22]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).

[23]Each agent i has a net position of IOUs equal to the IOUs received by firms (and issued by some other household) less IOUs she issued. Call $\hat{a}_i$ the IOUs issued by agent i and $\hat{a}_{-i,i}$ the IOUs 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 agent i). So agent i receives $\hat{a}_{-i,i} - \hat{a}_i$. The sum of all households' net positions is $\sum_i \hat{a}_{-i,i} - \sum_i \hat{a}_i$. Since firms pass to households all bonds they receive $\sum_i \hat{a}_{-i,i} = \sum_i \hat{a}_i$, i.e. the intra-temporal bond market clears. Credit could also be intermediated through banks that accept deposits and issue loans: that credit is within the period (*i.e.* with zero preference discount) and can be issued at no costs implies zero interest rate. See Ljungqvist and Sargent (2012) Section 18.10.3 for a discussion of the exchange and clearance of IOUs in similar settings.

balance after honoring the debt is $p_m m + a + wn + kr - pq$. This balance is stored in money m' or it can be saved in a'. Each unit pais the monetary value of a unit of goods at the beginning of the next period and costs v:

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

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 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 obviously relax this extreme distribution assumption that one either trades in full or not al tall. However, insurance—included later—smoothes the implications of this assumption.

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.[24] 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

[24] 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.

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.

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 consumption c, and decreasing in market labour n and in shopping effort d. $u(\cdot)$ 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 constraints are usually imposed in theories that incorporate search frictions in the goods market such as Bai et al. (2011). But this constraint is different than the one in Bai et al. (2011) where effort enters directly into the matching function. With my specification, the theory also works without this effort constraint and its theoretical and quantitative implications are disentangled from the rest of the model. This is an advantage because these constraints generated some controversy. Huo and Ríos-Rull (2013b) notice that effort is a substitute to production inputs in Bai et al. (2011). They argue in favour of complementarity, consistently with Equation (4). All these specifications imply that households reduce their search effort during recessions but Kaplan and Menzio (2016) show that the unemployed search more than the employed, which suggests that effort should increase during recessions. However, Huo and Ríos-Rull (2013b) construct a more elaborated search effort process where effort is procyclical, yet higher for the unemployed. 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.[25]

So, other things equal, the smaller ψ the larger money demand is. However, it should be noticed that the household effectively chooses 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 even though it is dominated by capital in return? 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 can be written as

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

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) thereby increasing ψ so that agents end up with more goods, thereby relaxing (5) as captured by λ_3. However, the increase in ψ implies that agents spend more money which tightens (6) as captured in the right hand side by $\lambda_4 p$. The increase in p also means that agents pay more per unit of good: this tightens constraints (3) and (6) as shown in the last term in the right-hand-side.

[25]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 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.

3.1.3 Store of value versus transaction motive

To see how the transaction motive 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 $m' = 0$. This is because here there is capital which pays higher return than money and can be turned into consumption at any time, so turning m' in goods would offer more present and future consumption.[26] 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.[27]

[26] This does not mean that money does not have a transaction role: indeed money relaxes the Transaction Constraint (3). But this motive is not sufficient to demand money because capital pays higher dividend and can be consumed. This can be appreciated from Equation (32) in Appendix A from which $\lambda_3 > \lambda_1$ in an interior solution, i.e. the lagrange multiplier on Constraint (5) is larger than the one on Constraint (3). So, if it was possible, agents would relax the goods constraint and tighten the transaction constraint by relocating money toward capital.

[27] 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. (2011), 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

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.

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

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. Thus Equation (12) implies

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

Since Equation (13) has to hold for a market to be active, it is what defines the function $\theta(p,q)$ that households take into account.

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.*

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 intuition behind the proof in Appendix C is that since the production function has constant returns to scale and input prices are taken as given, production increases proportionally with costs. Then, if profits per unit of production are zero, for a change in q not to affect profits, ϕ and thereby θ have to remain constant. From now on the dependance on q will be omitted and 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.[28] 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).[29]

To close the model it remains to specify the exogenous stochastic variables. They are z_m, A_d, A, χ_n, and β. z_m and A_d —shocks to the matching function and to effort productivity— are natural wedges implied by this paper, so it is interesting to study their implications. Shocks to technology A, labour supply χ_n, and discount factor β, have been shown to be important drivers of the business cycle and will avoid stochastic singularity in the likelihood estimation, see Fernández-Villaverde et al. (2016).

Definition 1 *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 decision rules and the value functions V solve the*

[28] 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.

[29] 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 Section 5.4.2, numerically I find that the possibility of multiple active markets does not occur in the parameter space considered.

household problem in 3.1

2. *Firms:* k_d*,* n_d *and* $\theta(p;\Omega)$ *satisfy Equations (10), (11), and (13) with* $\xi = \phi(\theta(1;\Omega))$*.*[30]

3. *Market Clearing: Households purchases are equal to firms sales:*

$$\psi(\theta(1;\Omega))q = \phi(\theta(1;\Omega))fq; \tag{14}$$

The liquid assets, and inputs markets clear:

$$m' = M, \qquad a' = 0,$$

$$fn_d = n, \tag{15}$$

$$fk_d = K. \tag{16}$$

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.*

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.

4 Characterization

4.1 Money and Credit

In a LW framework with credit, GMW identify cut-off levels of debt that mark the existence and non existence of monetary equilibria. Intuitively, if debt is large enough so that all transactions need not use cash, then cash has no value. That is true here too but the identification of these cut-offs is more complicated because, as is shown later, debt levels are not always neutral so transactions change with debt. However, it is possible to find a condition in terms of equilibrium outcomes that has a meaningful economic counterpart (debt over GDP) and that helps distinguish the implications of my model from GMW in terms of whether and when the credit limit is binding.

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

As it has already been mentioned, if agents borrow up to the maximum credit limit that is implementable in equilibrium, then money has no value: this is shown formally in Proposition 1 below. Importantly, the proposition also shows that the viceversa is true. As a result Corollary 1 follows: money has value whenever credit is below the maximum implementable credit limit. 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. Proposition 1 also shows that the maximum implementable limit is

$$L = wn + rk + \frac{(1-\psi)}{\psi}(p_m m + wn + rk). \tag{17}$$

Intuitively, it would not be possible for agents to borrow more than against their own income $wn + rk$ plus all the funds of the unmatched $(1-\psi)(p_m m + a + wn + rk)$, divided by the number of matched agents ψ.

Proposition 1 *Let borrowing be defined as $\hat{b} = pq - p_m m - a$. $p_m = 0$ if and only if agents borrow $\hat{b} = L$ where L — defined in Equation (17) — is the maximum implementable borrowing limit.*

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

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: without costly effort there is no monetary equilibrium (because with no credit restriction, they borrow to the maximum implementable credit limit), 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.

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 (30). From this condition it does not seem possible to theoretically 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.[31] This form of money hoarding is also akin BT models where to hold money saves some costs.

A further corollary follows: if the credit constraint binds, then money has value. Intuitively, because default is not allowed, agents never want to borrow above the natural limit, then credit constraints can only bind if below the natural limit. But with credit below the natural limit, money has value from Corollary 1. 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.

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

Finally, the tightness of this debt limit can be appreciated in terms of GDP which in this model is equal to $wn + rk$. When money has no value $p_m = 0$ so (17) and Proposition 1 say that the debt limit for this occurrence must be at least $\frac{wn+rk}{\psi}$ i.e. larger than 100% of GDP when $\psi < 1$.

4.2 Money is essential

As it has been mentioned, these propositions characterize monetary and non monetary equilibria, but do not provide parametric conditions that mark the existence and non-existence of monetary equilibria. For example, for some credit limit B it is in general possible to find an equilibrium in which said exogenous credit limit is equal to L and thus the equilibrium is non monetary. But for that same B there might be a monetary equilibrium too. This is because in a non monetary equilibrium, the economy contracts so much relative to a monetary equilibrium, that the borrowing limit is not binding. To illustrate this, let $B \equiv wn + rk$, i.e. agents can borrow their entire income, and construct a non monetary equilibrium with $p_m = 0$ over the entire state space. The next proposition shows that in this case no production takes place.

[31]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.

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

A non monetary equilibrium would clearly be very bad for welfare as without production, consumption would be equal to negative investment, depleting capital. With more credit, e.g. if $B \equiv \gamma(wn + rk)$ with $\gamma > 1$, then some production takes place even in the non monetary equilibrium.[32] Nevertheless, the proposition above helps appreciating the role of money in this model perhaps more than anything else: 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. It should be noted that $B \equiv wn + rk$ is not tight by usual standards: to let people spend their entire income without money is what is necessary to get entirely rid of a cash in advance constraint in the neoclassical model.

Finally, it is worth noting that if $B \equiv \gamma(wn + rk)$ with $\gamma < 1$ then there can *only* be monetary equilibrium.[33] Intuitively, in this case the constraint always binds thus Corollary 2 applies.

4.3 The role of search and effort for money demand

To clarify the role of the key frictions of the model (search and effort), it is instructive to study the case when the matching function is $\min(f, h)$ and effort is not costly (*i.e.* the utility function is flat in d). As discussed in Appendix B, the equilibrium is characterized by $\theta = \psi = \phi = 1$. Furthermore, with $B \geq wn + rk$ money looses all value and the equilibrium boils down to the one of the neoclassical model.

While perhaps not surprising, the result above clarifies what exactly gives rise to valued money: with no effort costs and with $\min(f, h)$ there is no search friction, then money has no value when people can spend their entire income with credit. With $\psi < 1$, not all funds can be spent in goods which gives rise to money demand unless the credit limit is sufficiently above income (at the maximum implementable limit).

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

[33] Suppose $p_m = 0$, then $q < wn + rk$, then (36) implies $\psi > 1$ which is impossible.

4.4 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.*

4.5 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.[34]

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.

4.5.1 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

[34]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.

a planner problem. Since this subsection characterizes deterministic steady state results, for simplicity, the planner problem abstracts from the shocks.

Definition 2 *An allocation $\{c, n, d, q, k', \theta\}$ is efficient if it solves the following:*

$$\widetilde{V}(k) = \max_{\{q,c,k',\theta,d,n\}\geq 0} u(c, n, d) + \beta E \widetilde{V}(k') \tag{18}$$

s.t.

$$\theta q \leq A k^{\alpha} n^{(1-\alpha)} \tag{19}$$

$$q \leq A_d d \tag{20}$$

$$c + k' - k(1-\delta) \leq \phi(\theta)\theta q \tag{21}$$

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

Equation (19) ensures that total production is not smaller than the quantity offered by each trading post (q) times the number of trading posts θ. Constraint (20) 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 (21), 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.[35] It should be noted that concavity does not hold for any parameterization: as discussed in Section 5.4.2, it is necessary to have sufficient complementarity in the matching function. While concavity of the household problem is needed to formally conclude that the Friedman rule is efficient, all other theoretical results, including those using the household first order conditions, do not require concavity: to the extent that the first order conditions are necessary, *i.e.* they must hold in an equilibrium, they characterize equilibrium.[36]

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.*

[35] Without concavity of the household problem in principle the solution to the planner problem may not be an equilibrium, even though it satisfies the household first order conditions.

[36] Also, if the constraint set is not convex the Bellman equation may not be differentiable. However, the first order conditions can be derived through a variation method without relying on the Bellman equation. See Stokey et al. (1989) Section 4.5.

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).

That the FR is optimal may appear counterintuitive at first glance. Intuitively, inflation could have been beneficial in this framework, because it could have forced households to search more. But this channel does not operate because the effort decision is not distorted: with or without inflation there is no wedge between the household first order condition for d—Equation 28—and the one of the planner.

Corollary 3 below clarifies how inflation distorts the allocation in the case of no effort costs: with inflation agents do not want to remain stuck with money and so they choose a market with too high θ relative to the planner solution; this way it is easy for the buyer to find goods, but difficult for the seller. Furthermore, the corollary shows that the FR calls for large amounts of liquid savings. This may seem counterintuitive because in this model money is a measure of savings not matched with goods. However, 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 $\theta = 0$. This is because Constraint 20 in the planner problem does not bind and Constraints (19) and (21) 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 (19) and (20) imply that q and d approach infinity. Intuitively, the number of trading posts tend to zero, but become large.[37] 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

[37] 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).

that θ is chosen optimally at the FR and not otherwise.[38]

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.

These results also highlight that money has a social role in that it affects people search decisions: when money is a good store of value (low inflation), agents are not worried about holding it and search optimally, making the matching efficient.[39]

5 Quantitative exercise

In this section, I map this model to the data and study its business cycle properties.

5.1 Matching

I assume the following matching function:

$$\mu = z_m^{1/\rho} \left(\alpha_m f^\rho + (1-\alpha_m) h^\rho\right)^{1/\rho}. \tag{22}$$

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 as discussed in Appendix B.

5.1.1 Recasting the Matching in terms of aggregate demand and supply

The variables in Matching Function (22) μ, f, and h do not have clear empirical counterpart but multiplying the right and left hand side by q one gets

$$y = z_m^{1/\rho} \left(\alpha_m y_s^\rho + (1-\alpha_m) y_d^\rho\right)^{1/\rho}. \tag{23}$$

[38] 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.

[39] That the matching becomes efficient also hinges on competitive search. Alternative bargaining systems may introduce effort inefficiencies, with possibly different implications for inflation.

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.

y_d is constructed using data on money and GDP.[40] 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 capacity utilization ($\textsc{tcu}$) published by the Federal Reserve Board: $\textsc{tcu}$ is the percentage of total available capacity being used to produce demanded finished products. This matches closely with ϕ in the model. In particular, the literature on capacity utilization measures output as

$$y = (\textsc{tcu}\, k)^{\alpha} n^{1-\alpha}. \tag{24}$$

Since $\textsc{tcu} \in [0, 1]$, total production capacity y_s is obtained putting $\textsc{tcu} = 1$ in Equation (24). Then $\phi = y/y_s = (\textsc{tcu})^{\alpha}$, and y_s can be backed out as y/ϕ.

Having constructed ϕ, ψ and $\theta = y_s/y_d$, it is possible to check whether they behave consistently with the novel matching process, which implies that ψ is an increasing while ϕ is decreasing in θ. This is an interesting test because the matching function has not been used to construct these variables. Figure 2 shows that the data line up with the theory rather well.[41]

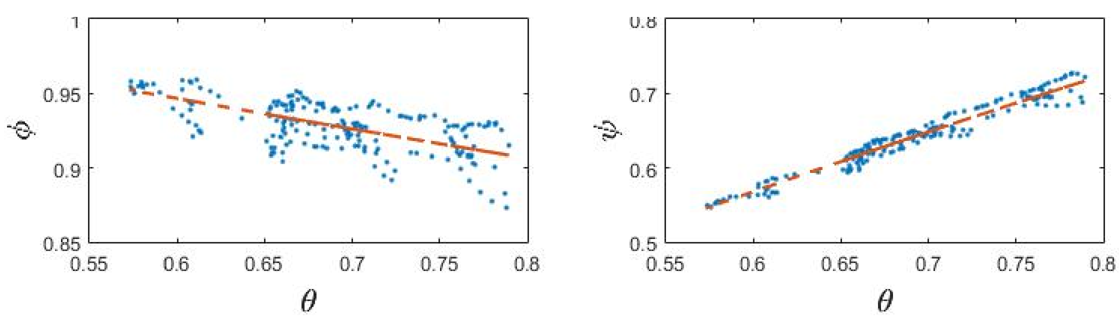

Figure 2: Empirical ϕ and ψ as a function of θ. Note: the trend is the prediction of the matching function with the estimated parameters α_m and ρ, and with steady state z_m.

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

[41] 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.

5.2 Preferences

The utility function is: $u = \log(c) - \chi_n \frac{n^{1+1/\nu_n}}{1+1/\nu_n} - \chi_d \frac{d^{1+1/\nu_d}}{1+1/\nu_d}$. χ_n and χ_d determine average hours and effort. ν_n and ν_d are the Frisch elasticities of labour and effort supply.

5.3 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. In any case, non reported impulse response functions show that the business cycle properties of credit limits have small effects in this model with flexible prices.

5.4 Parametrization

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

I assume z_m, A_d, A, χ_n and β to be AR1 stationary independent stochastic processes. I estimate the persistence parameters and 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. The remaining parameters are calibrated as follows.

5.4.1 Calibrated parameters and targets

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 process 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.[42] 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,

[42]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).

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 depend on the targeted steady state market hours and effort, as well as the Frisch elasticity parameters, to be estimated. 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.

5.4.2 Bayesian Estimation

Observables

The observables 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 data should 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 parameter ρ. Finally, variations in GDP and θ imply variations in y_d and thus in effort via Equation (4); this disciplines the effort supply shock and the effort Frisch elasticity.

An indication that this methodology works is that an estimation on model generated data (with zero weight on the priors) finds the parameters used to generate the data.

Priors and Posteriors

The table in Appendix F summarizes priors and posteriors.

I assume the Frisch elasticity of labour supply to be Gamma distributed with mean 0.85 and variance 0.1. This ensures that the Frisch elasticity is in line with micro studies, who tend to find smaller elasticities relative to what macro models need to match hours volatility. The posterior mode is 1, in line with the studies surveyed by Keane (2011). 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.

I assume a Gamma distribution for $-\rho$. This is because I restrict ρ to be smaller than zero: with too little complementarity (ρ too large), the interior solution of the household's first order condition for p does not maximize the objective function of the household. Increasing complementarity toward the perfect competition case of $\rho = -\infty$ ($\min(f, h)$ implies perfect competition as discussed in Appendix B) is necessary to make the slope of marginal cost larger than that of the marginal gain: Figure E.1 in Appendix E plots for different levels of ρ, the marginal gain and marginal cost in Equation 7 of moving θ (which is equivalent to move p given the function $\theta(p)$). In the first 2 panels, with sufficiently high complementarity, the marginal cost cuts the marginal gain from below, this is consistent with the objective function being increasing at the left of the crossing point and decreasing thereafter. The opposite is true in the last panel with less complementarity (larger ρ), so that the crossing point is not a max. The last 2 panels suggest that there may be a nife-edge ρ between 0 and 1 for which the marginal cost and marginal gain imply a flat objective function so that many markets give same utility. Restricting $\rho \leq 0$ rules this out. This restriction is also consistent with the notion that the two inputs are both essential for a match: with $\rho \leq 0$ one cannot have sales with only aggregate demand, or only aggregate supply. I find a posterior mode for $\rho = -1.57$.

Following the literature, the persistence parameters of the stochastic processes have a Beta prior and the standard distribution 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.

5.5 Variance Decomposition and Impulse Response Functions

To appreciate the importance of each shock the variance decomposition of some variables of interest is reported in Table 1.

First, the matching and the effort shocks alone, (z_m and A_d in the first two rows), explain 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.

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

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). This suggests that A_d should really be interpreted as a demand shock (similarly to Bai et al. (2011)). 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. It may be interpreted as capturing a more cautious behavior due to frictions such as information, screening, monitoring, agency and retail costs.[43]

Figures E.2 and E.3 in Appendix E report the impulse response functions to z_m and A_d. The main take away points are that after a positive z_m shock ϕ increases i.e. there is 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] Velocity surges (therefore, after a negative shock velocity declines as it has happened in many recessions). p_m drops *i.e.* inflation is procyclical.

Similarly to z_m, A_d shocks are also expansionary, induce an endogenous Solow residual, and comovement of consumption, market hours, and input prices with output.

[43]Besides mechanically affecting the matching between demand and supply, there is a further endogenous reason why the matching process affects spending: during recessions the return on capital is lower and that of money is higher because of lower inflation. Then, other things equal, it is optimal for households to choose a market with lower probability of finding goods. Furthermore, given the preference for consumption smoothing, the drop in spending is especially absorbed through a substitution of capital investment with money holdings, akin to a Paradox of Thrift. However, with competitive search, the market choice is not distorted and this endogenous reaction is not inefficient as it lessens the drop in firms finding probability ϕ.

[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.

But search effort, ψ, p_m and velocity move in the opposite direction relative to a z_m shock. The shock also causes a large movement in supply over demand θ, which barely moves after a shock to z_m. These are the crucial differences in the propagation of z_m and A_d shocks which identify the two apart: Bai et al. (2011) estimate a model with demand shocks similar to A_d and find it to be more important than technology shocks. They do not use data on money and TCU in the estimation: this would not allow to distinguish between A_d and z_m.

The model also explains 43% of the variance of hours without labour supply shocks and with low labour supply elasticity. 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 (26), (27), (28) and (32) in Appendix to

$$-u_n = \left(u_c\psi - \frac{u_d}{A_d} + \lambda_4(1-\psi) \right) w, \tag{25}$$

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, as shown in Figure E.4, 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.[45] Furthermore, on the supply side, 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 which other things equal increase labour supply, but not enough to overcome the other effects.

Also counterfactual is that technology shocks make ϕ countercyclical while TCU is procyclical.[46] Not surprisingly given these results, technology shocks are estimated

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

[46] Since the mapping between data on capacity utilization and the notion in this model may not

to play a less important role than in the neoclassical model.

5.6 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 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.

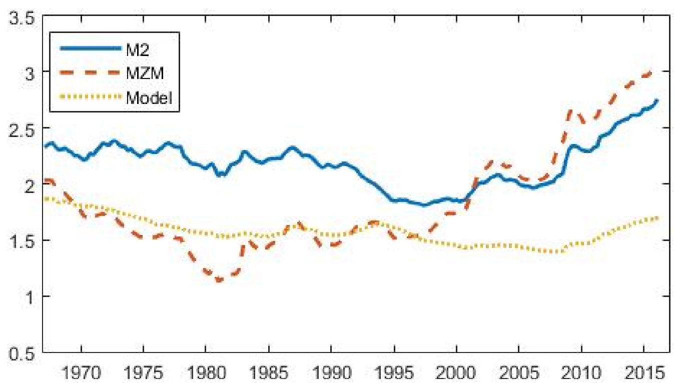

Figure 3: M2, MZM, and model simulation of Broad Money.

On average the model implies less liquid assets than these other measures but it does not seem far off.[47] 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 it might be fruitful to study financial frictions within this framework.

be perfect, I experimented estimating the model with larger measurement error for τ_{CU} (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.

[47]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.

5.7 NBER Recessions

How does the model account for real 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 TCU 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.[48]

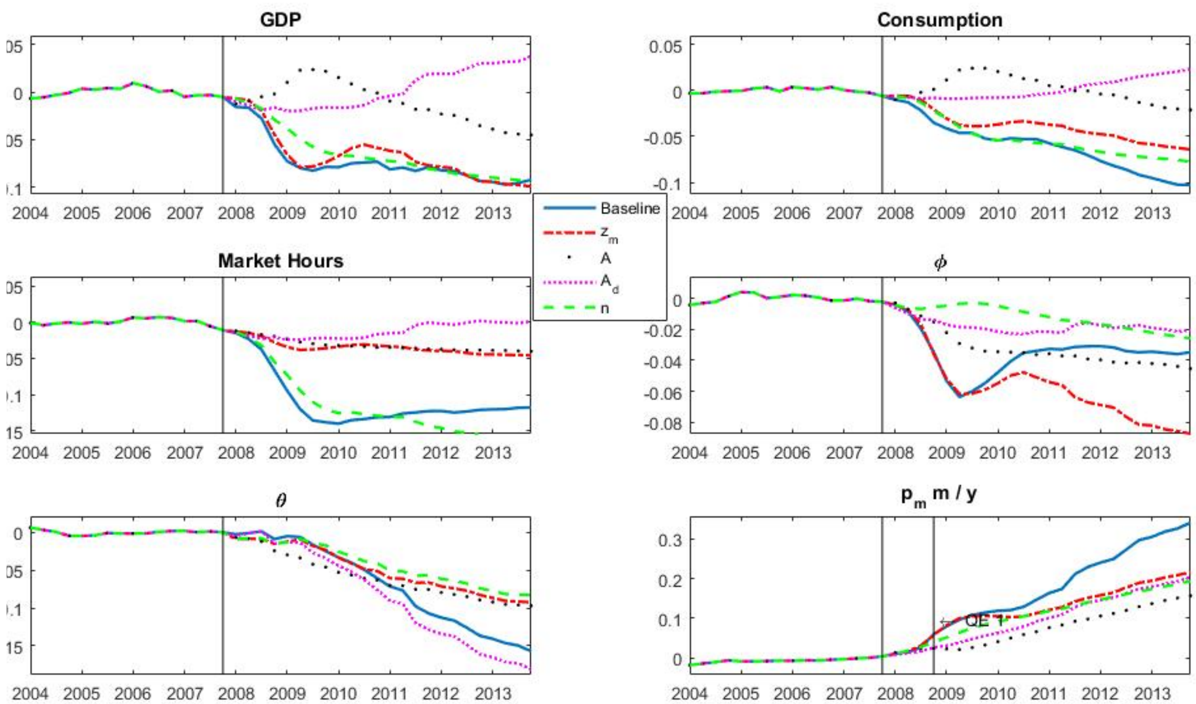

Figure 4: 2007.IV recession due to each shock

First, technology shocks A are not responsible for the crisis. Instead z_m shocks account for virtually the entire drop in output and ϕ (thereby generating an endogenous drop

[48]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.

in productivity), as well as the increase in liquidity.[49] z_m shocks are also responsible for a sizable fraction of the drop in consumption but for a small fraction of the drop in market hours, which is mainly due to the labour supply shock. The demand shock A_d plaid a negligible role in the Financial crisis. The path with shocks on β is not shown because not important.

A few patterns emerge when also looking at past recessions. Figures E.5—E.9 in Appendix E show a peak to trough analysis carried at previous NBER recession dates.[50] All recessions show a drop in ϕ. Some recessions (all those from the 80s onward) are characterized by a liquidity surge.[51] Earlier recessions do not show a liquidity surge (those started in 1969.IV and 1973.IV).

Which shocks account for these facts? i. They have to be shocks that are recessionary and make ϕ drop; those are negative z_m and A_d shocks; instead negative A shocks make ϕ increase which explains why technology shocks are in general not that important during recessions. ii. As seen in the previous subsection, negative z_m shocks make liquidity surge, negative A_d shocks make it drop. So recessions where liquidity increased are characterized by z_m shocks: this alone generates both the drop in ϕ and the surge in liquidity (as well as most of the drop in GDP). 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. So early recessions were also characterized by an A_d shock, this is also evident in the figures from the increase in θ. Finally, labour shocks always play a role for market hours (which is to be expected given the low labour supply elasticity).

To conclude, it is worth stressing that these exercises take this simple model very seriously and thus should not be taken to be more than an illustration of the implications of the model. However, these exercises reinforce the suggestion that the two wedges introduced in this theory —z_m and A_d— are promising to explain the business cycle. Furthermore, they make the case that monetary aggregates and capacity utilization are linked and important in order to understand the business cycle.

[49]Since the model abstracts from monetary policy intervention, which may have contributed to the increase in liquid assets, the last panel in the figure includes a vertical line at the time of the first round of Quantitative Easing to show that the liquidity surge already took place. Indeed, Williamson (2015) argues that it is not obvious if and how a swap of liquid assets between the central bank and private financial institutions affects money supply.

[50]The 1981.I and 1981.III recessions are plugged into the same figure to save space.

[51]That a liquidity surge characterized many recessions corroborates the view that such surge is not just an artifact of Quantitative Easing.

6 Conclusions

Motivated by the observed amounts of savings held in liquid form 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, this theory offers an explanation for money demand when anonymity and the lack of enforceability that limits insurance and credit is relaxed.

Money also enhances productivity and welfare: when people are not 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.) reflecting empirical counterparts ranging from government bonds, equity shares and other financial products, to possibly far less liquid assets such as houses.[52]

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 be ineffective: 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

[52] 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.

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 M&A, 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*.

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 (2011). Demand Shocks that Look Like Productivity Shocks. Technical report.

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 (2014). Search-Based Endogenous Illiquidity and the Macroeconomy. Technical report.

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.

Fernández-Villaverde, J., J. F. R. Ramírez, and F. Schorfheide (2016, January). Solution and estimation methods for dsge models. Working Paper 21862, NBER.

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.

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 (2013a, September). Paradox of thrift recessions. Working Paper 19443, National Bureau of Economic Research.

Huo, Z. and J.-V. Ríos-Rull (2013b, 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.

Jevons, W. S. (1875). *Money and the Mechanism of Exchange*. London: Appleton and Co.

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 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.

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.

Stokey, N. L., R. E. L. (Jr.), and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.

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

$$u_c = \lambda_3, \tag{26}$$

$$u_n = -\lambda_4 w - \lambda_1 B_n, \tag{27}$$

$$u_d = -\lambda_2 A_d, \tag{28}$$

$$\lambda_3 - \lambda_{k'} = \beta E\left(\lambda_4' r' + \lambda_3'(1-\delta) + \lambda_1' B_k'\right) \tag{29}$$

$$\lambda_4 p_m - \lambda_{m'} = \beta E\left((\lambda_1' + \lambda_4')p_m'\right) \tag{30}$$

$$\lambda_4 v - \lambda_{a'} = \beta E\left(\lambda_1' + \lambda_4'\right) \tag{31}$$

$$(\lambda_3 - \lambda_4 p)\psi = \lambda_1 p + \lambda_2 - \lambda_q \tag{32}$$

$$\lambda_1 + \lambda_4\psi = \frac{\partial\psi}{\partial\theta}\frac{\partial\theta}{\partial p}(\lambda_3 - \lambda_4 p) + \lambda_p/q, \tag{33}$$

B Leontief matching function and no effort costs

With matching function $\min(f, h)$, $\phi = 1$ if $\theta \leq 1$ and $\phi = 1/\theta$ if $\theta > 1$. Thus $\phi(\theta)$ is not strictly monotone as assumed in the main text.

Since $\phi = 1$ for $\theta \leq 1$, firm arbitrage ($\phi(\theta)p = \xi$) implies $p = \xi$ if $\theta \leq 1$, $p = \xi\theta$ if $\theta > 1$. Thus θ is not uniquely identified by p but the relationship is a correspondence. As a result, the household's submarket choice does not only involve choosing p, q, but also θ. It is easy to argue that the optimal choice of the household is $p = \xi$ and $\theta = 1$.[53]

But with $\psi = 1$ the natural limit defined in Equation 1 is $L = wn + rk$. So the credit limit $B \geq wn + rk$ is not binding. Then with no effort costs Lemma 2 implies that money has no value.[54] In turn, Equations (5)—(6) and $\psi = 1$ boil down to the usual neoclassical budget constraint $c + k' - k(1-\delta) \leq wn + kr$.

C 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, which implies that 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)

[53] This is because from the matching function, $\psi = \theta$ if $\theta \leq 1$ and $\psi = 1$ if $\theta > 1$. First, ($\theta = 1$, $p = \xi$) is preferred to ($\theta < 1$, $p = \xi$) because the latter implies $\psi < 1$, which induces more savings in money, that pay lower return than capital. On the other hand, $\theta > 1$ would imply $p > \xi$ but would not increase ψ, clearly an inferior choice.

[54] Lemma 2 uses the first order condition for q, which is not affected by the fact that here $\phi(\theta)$ is not strictly monotone.

is zero. Then the left hand side must be equal to zero too. Since $a' = 0$ in equilibrium, 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 $\psi q > \phi f q$. 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 $a' < 0$, which violates equilibrium.

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 28. Then Condition (32) 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' —(29) and (30)— 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 (27) and (29). Then (30) 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 (30) 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 (30) 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]

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 markets imply that

$$\phi f A k_d^\alpha n_d^{1-\alpha} = \phi A k^\alpha n^{1-\alpha} = wn + rk. \tag{34}$$

If $p_m = 0$, from Equation (3) $q = wn + rk$, which combined with Equation (34) implies

$$q = \phi A k^\alpha n^{1-\alpha}. \tag{35}$$

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

$$\psi q = \phi A k^\alpha n^{1-\alpha}. \tag{36}$$

Equations (35) and (36) 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 (30) 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.

[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.

[56]The equilibrium conditions in which money or p_m appear are Equations (3) and (6), and the Euler equation for m', (30) 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 (30) 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.

Proposition 4
Change p_m to $\hat{p}_m$ so that,

$$\hat{p}_m m + zB = p_m m + B \tag{37}$$

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 (30). 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 (30) is not affected 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 (37) 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, are identical to those of the planner in steady state. It is then trivial to show that all other conditions are also identical.

The first order conditions for the household and associated Lagrange multipliers are reported in Appendix A.

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 (30), 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

$$\frac{\partial p}{\partial \theta} = -\frac{\partial \phi}{\partial \theta}\frac{p}{\phi}; \tag{38}$$

The properties of the matching function and Inada conditions on the utility function ensure that λ_p and λ_q are both equal to zero. Then substituting $\frac{\partial p}{\partial \theta}$ from Equation (38) into Equations (32) and (33), normalizing the equilibrium $p = 1$, and noticing that the matching function implies $\psi = \theta\phi$, one gets

$$-\frac{\partial \psi}{\partial \theta}\frac{\partial \theta}{\partial \phi}\frac{\psi}{\theta}(\lambda_4 - \lambda_3) = \lambda_2 - \lambda_3\psi. \tag{39}$$

Substituting $\lambda_1 = 0$ into Equation (32) one gets $\lambda_2 = \psi(\lambda_3 - \lambda_4)$. Substituting this latter condition into Equation (39), one gets

$$\lambda_2(\frac{\partial\phi}{\partial\theta} - \frac{\partial\psi}{\partial\theta}\frac{1}{\theta}) = \lambda_3\psi\frac{\partial\phi}{\partial\theta}.$$

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 finally gets

$$-\lambda_2 = \lambda_3\frac{\partial\phi}{\partial\theta}\theta^2. \tag{40}$$

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 can be arranged as

$$\tilde{\lambda}_3\phi = \tilde{\lambda}_1 - \tilde{\lambda}_3\frac{\partial\phi}{\partial\theta}\theta \tag{41}$$

and

$$\tilde{\lambda}_3\phi\theta = \tilde{\lambda}_1\theta + \tilde{\lambda}_2, \tag{42}$$

where $\tilde{\lambda}_1$, $\tilde{\lambda}_3$ and $\tilde{\lambda}_3$ re the Lagrange multipliers on constraints (19), (20), and (21).[57]

The latter two conditions imply

$$-\tilde{\lambda}_2 = \tilde{\lambda}_3\frac{\partial\phi}{\partial\theta}\theta^2. \tag{43}$$

This planner condition coincides with the equilibrium condition (40) *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 (30) and steady state imply $\lambda_1 > 0$. Furthermore, with the marginal utility of effort $u_d = 0$, Equation (28) implies $\lambda_2 = 0$. Then, Equation (32) 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 (30) and steady state, when $\frac{p'_m}{p_m} \to \frac{1}{\beta}$, $\lambda_1 \to 0$. Then from Equation (32) $\psi \to 0$ and/or $\lambda_3 - \lambda_4 \to 0$. But from Equation

[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.

[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 $\lambda_p = \lambda_q = \lambda_{m'} = 0$.

(33), if $\lambda_3 - \lambda_4 \to 0$ then $\psi \to 0$ because from Equation (27) $\lambda_4 > 0$ 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 (19) implies that θ tends to zero if and only if production per trading post q tends to ∞. Then with a bounded credit limit B, Equation (3) requires that for θ that tends to 0, $p_m m$ tends to ∞, and for $\theta > 0$, $p_m m$ bounded.

D Households do not sell capital

Following up on footnote 21, to enable the household to put on sale some amount of capital d_k in a submarket with a different price than the one at which she buys, one first needs to define a new menu $\theta^b(p, q)$ at which buyers are indifferent relative to the submarket in which the firms operate.[60] Then the seller can pick a price p_k on that schedule. In order to define queue lengths in this representative agent environment in which the household buys and sells "from itself", let h_b be the probability that a household searches in a submarket other than where the firm sells, and h_s the probability that a household puts some of its capital on sale. The budget constraints after insurance, (3)—(6), become

$$pq + \hat{p}\hat{q}h_b \leq p_m m + a + B(k, n, \Omega, \exists) + p_k \phi(\theta^b(p_k, d_k))d_k h_s, \tag{44}$$

$$q + \hat{q}h_b \leq A_d d, \tag{45}$$

$$c + k' - (k(1-\delta) - \phi(\theta^b(p_k, d_k))d_k h_s) \leq \psi(\theta(p, q))q + \psi(\theta^b(\hat{p}, \hat{q}))\hat{q}h_b, \tag{46}$$

$$p_m m' + va' \leq p_m m + a + wn + kr + p_k \phi(\theta^b(p_k, d_k))d_k h_s - \psi(\theta(p, q))pq - \psi(\theta^b(\hat{p}, \hat{q}))\hat{p}\hat{q}h_b. \tag{47}$$

It is also necessary to impose $d_k \in [0, k]$. The left-hand-side of (44) makes clear that households buy goods q from firms in a submarket with some chosen price p from the firms' menu $\theta(p, q)$ and buy goods $\hat{q}$ from themselves (or other households) at some other price $\hat{p}$ from the buyers menu $\theta^b(\cdot)$. Given $\theta^b(\cdot)$, the household as a seller can choose p_k, d_k on the right-hand-side. The way the other constraints are affected should be intuitive. Equilibrium requires that $\hat{p} = p_k$, $\hat{q} = d_k$, $\phi\left(\theta^b(p_k, d_k)\right) d_k h_s = \psi\left(\theta^b(p_k, d_k)\right)\mathrm{h}_b$, and $\theta^b(p_k, d_k) = h_s/h_b$.

From the first order conditions on h_s and h_b, the household is happy to sell if $p_k(\lambda_1 + \lambda_4) \geq \lambda_3$ and it is happy to buy from other households as long as $\hat{p}(\lambda_1 + \lambda_4\psi_b) \leq \lambda_3\psi_b - \lambda_2$, where

[60] This is because on the firm schedule, buyers are not indifferent between submarkets but choose some given p.

$\psi_b \equiv \phi\left(\theta^b(p_k, d_k)\right)$. From the two inequalities above, the equilibrium condition $p_k = \hat{p}$ requires $\psi_b > 1$, impossible with the search friction.[61]

Note that with heterogeneous agents there could be gains from trade so markets for k may open.

E Figures

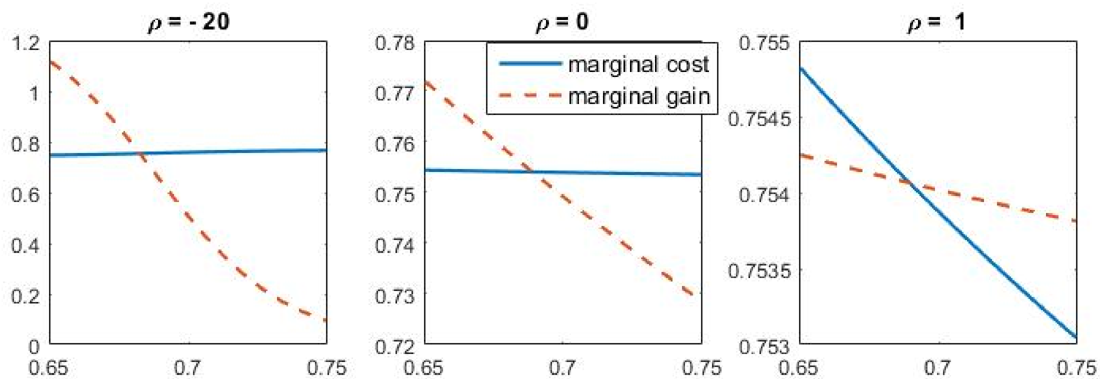

Figure E.1: Marginal gain and marginal cost of submarket choice as a function of θ (condition (7)) for different levels of complementarity in Matching Function 22

[61]It is not necessary to define the menu $\theta^b(p, q)$ for which buyers are indifferent. But for the sake of completeness, here is how it can be done. The household has to be indifferent between the favourite price on the firms menu and any other p, q, θ combination. Suppose the household wishes to spend sum funds x, then the following must hold

$$-\lambda_1 x - \lambda_2 x/p + \lambda_3 \psi(\theta(p,q))x/p - \lambda_4 \psi(\theta(p,q))x = -\lambda_1 x - \lambda_2 x/\hat{p} + \lambda_3 \psi(\theta^b(\hat{p},\hat{q}))x/\hat{p} - \lambda_4 \psi(\theta^b(\hat{p},\hat{q}))x. \quad (48)$$

This can be solved for $\theta^b(\hat{p}, q_k)$.

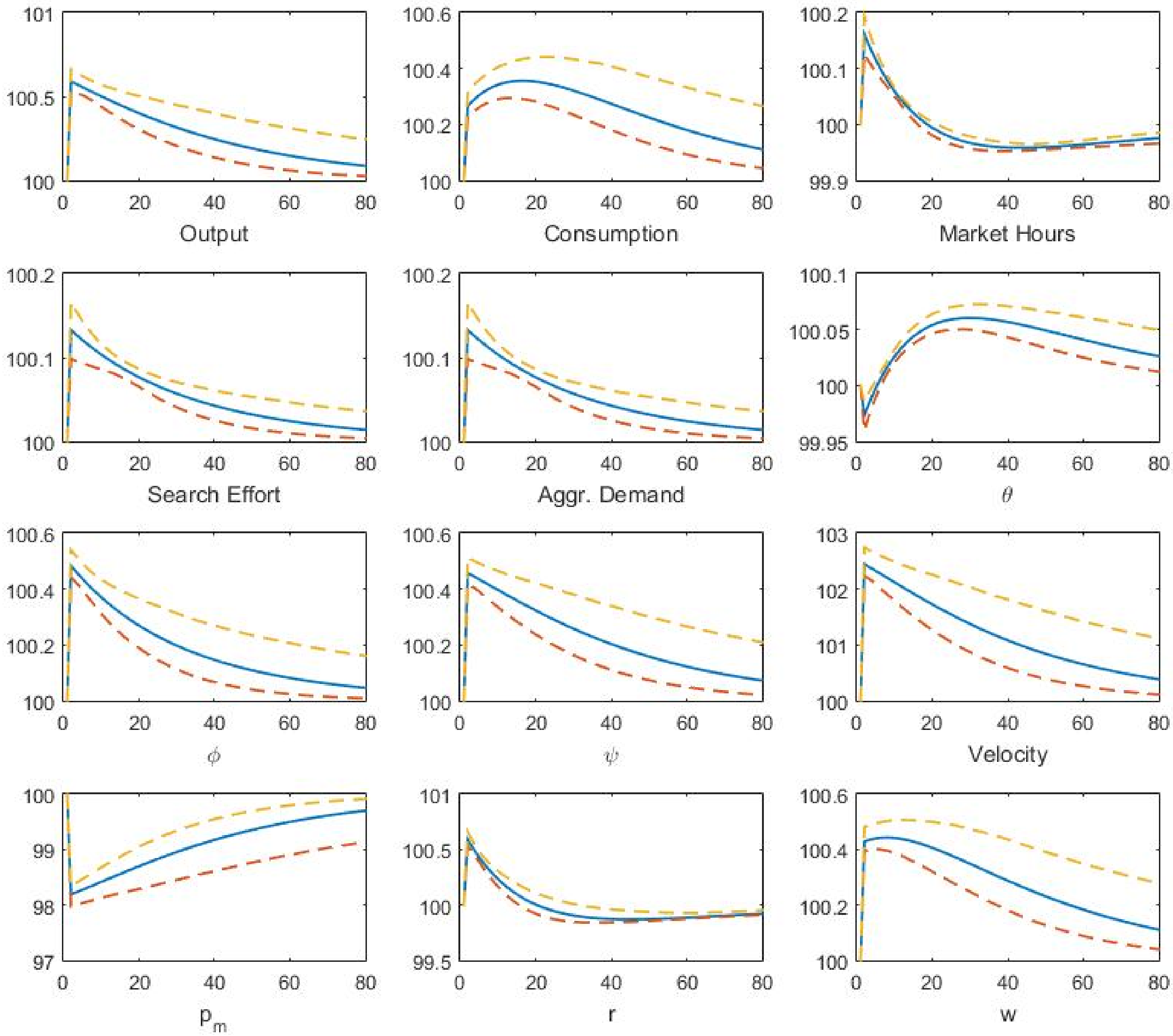

Figure E.2: IRF to Matching Shock z_m

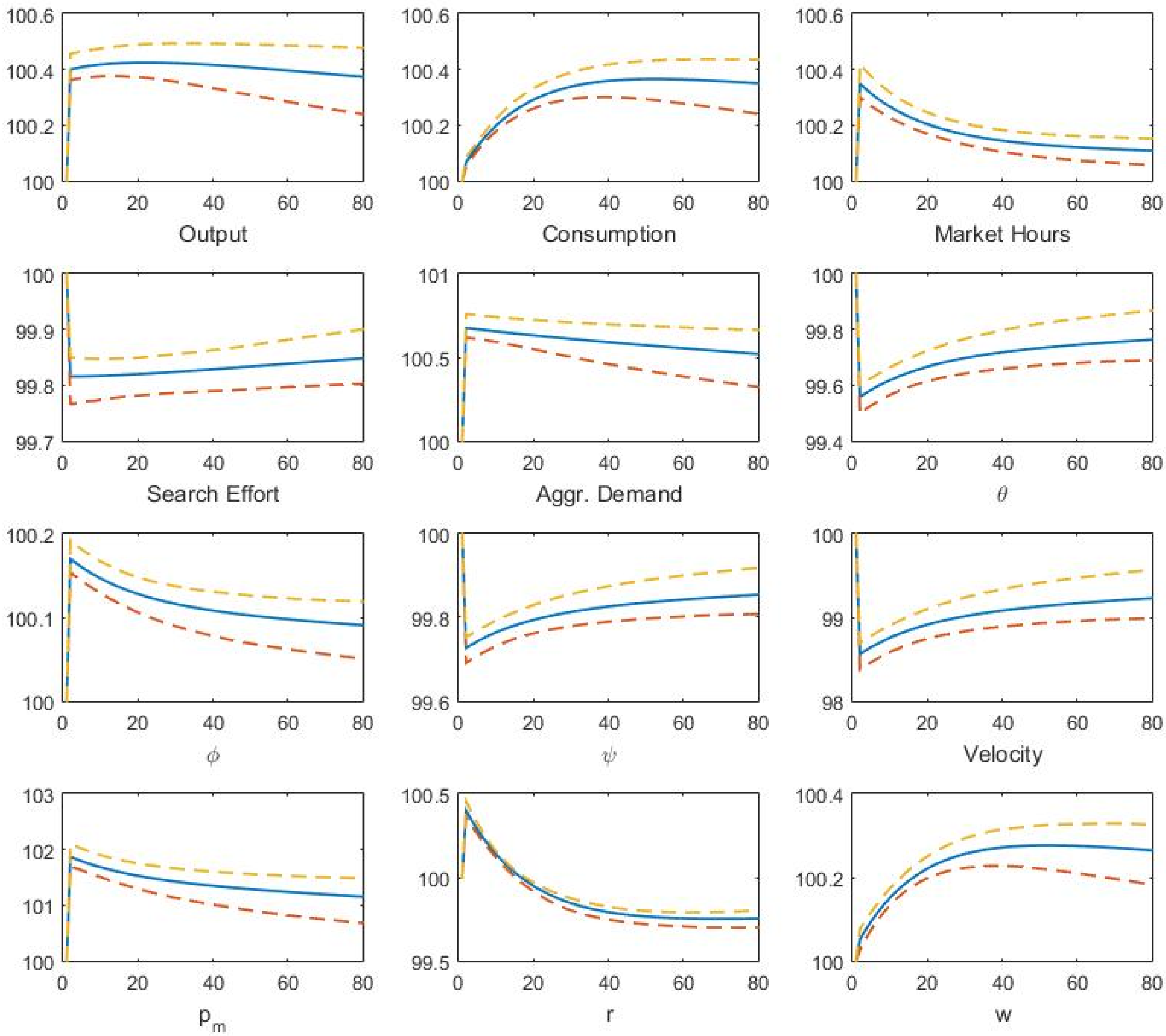

Figure E.3: IRF to Effort Shock A_d

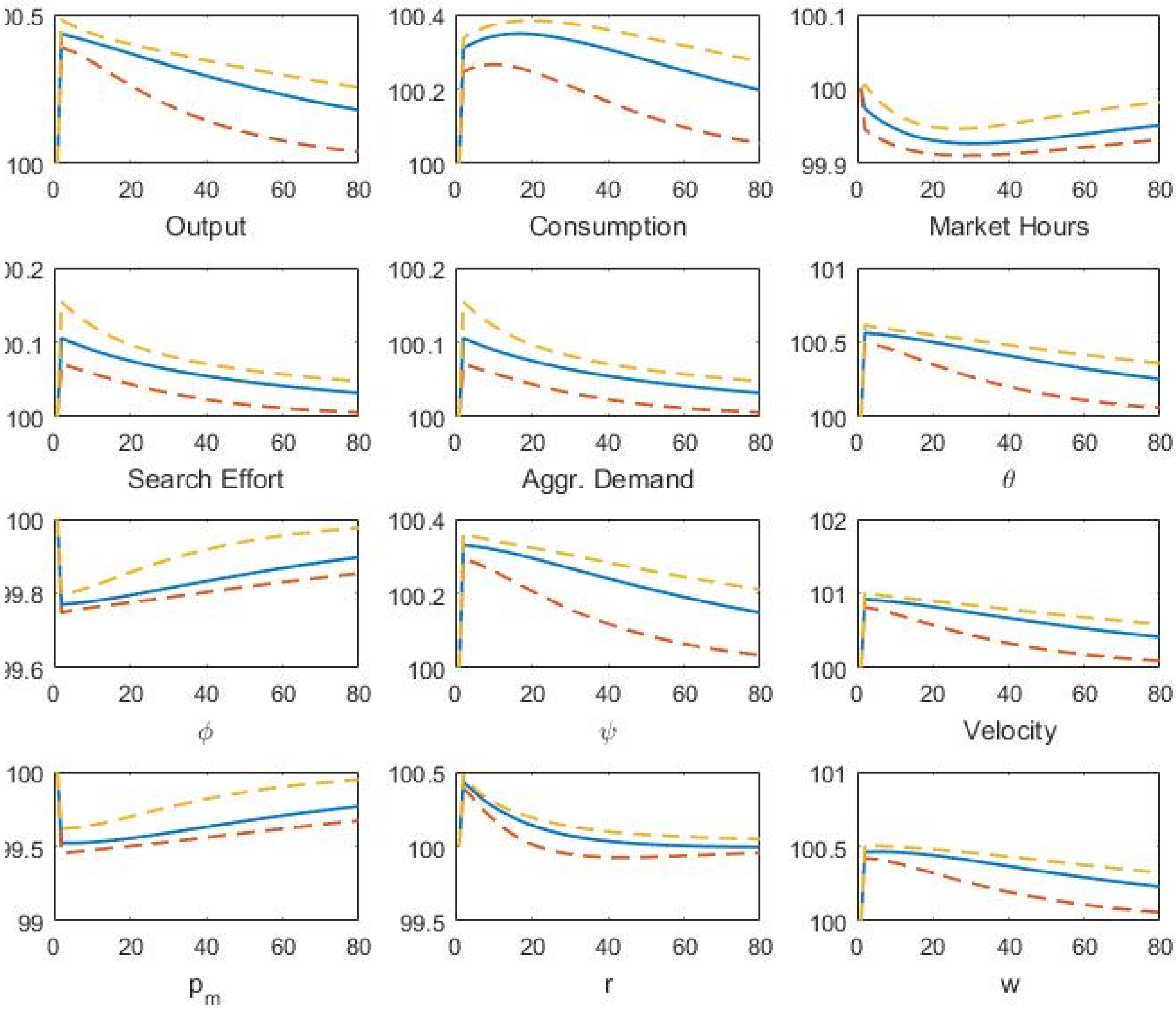

Figure E.4: IRF to Effort Shock A

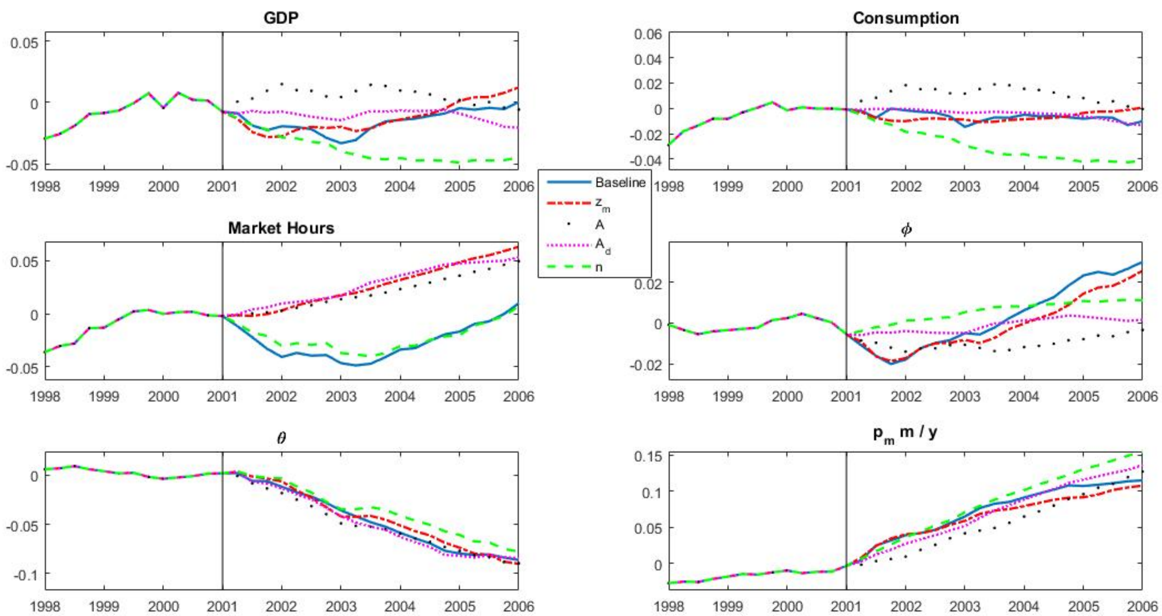

Figure E.5: 2001.I recession due to each shock

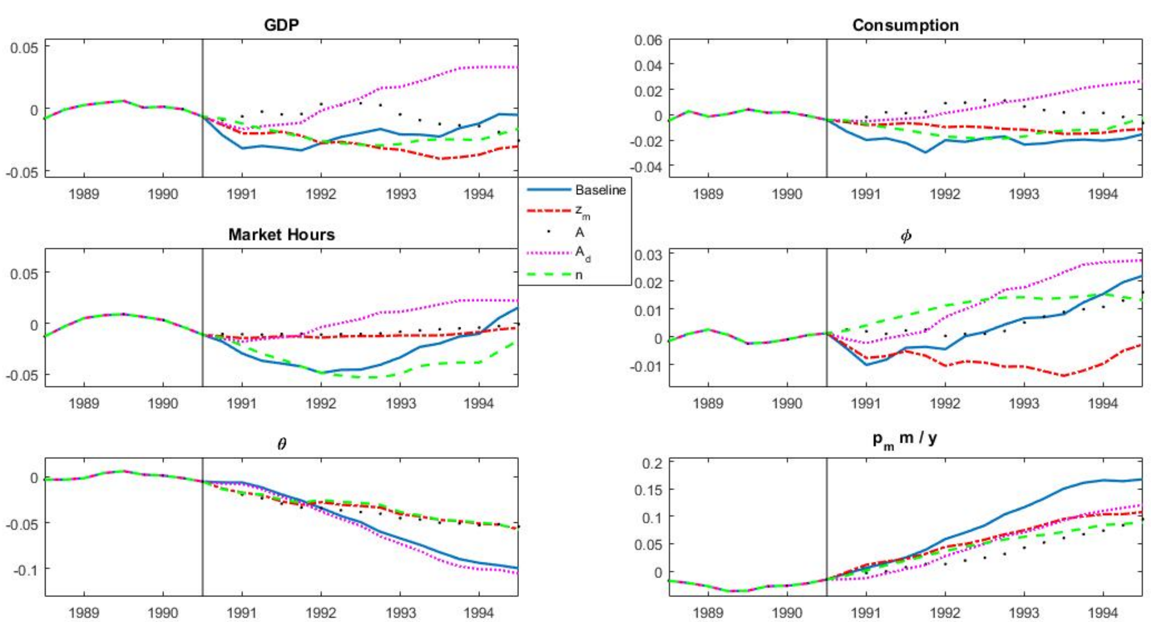

Figure E.6: 1990.III recession due to each shock

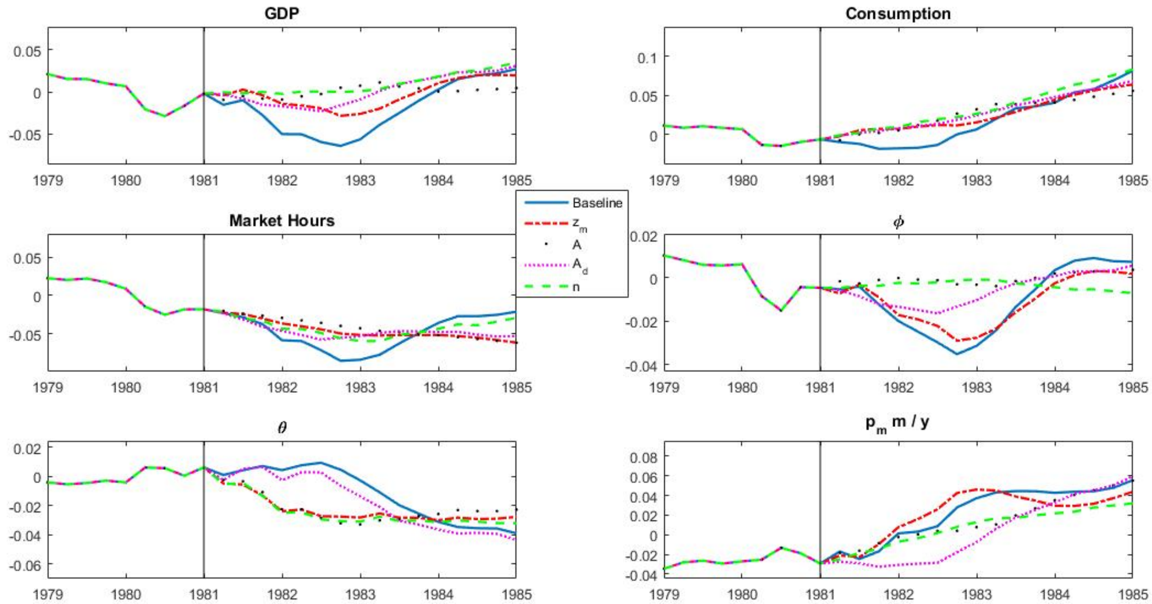

Figure E.7: 1981.I recession due to each shock

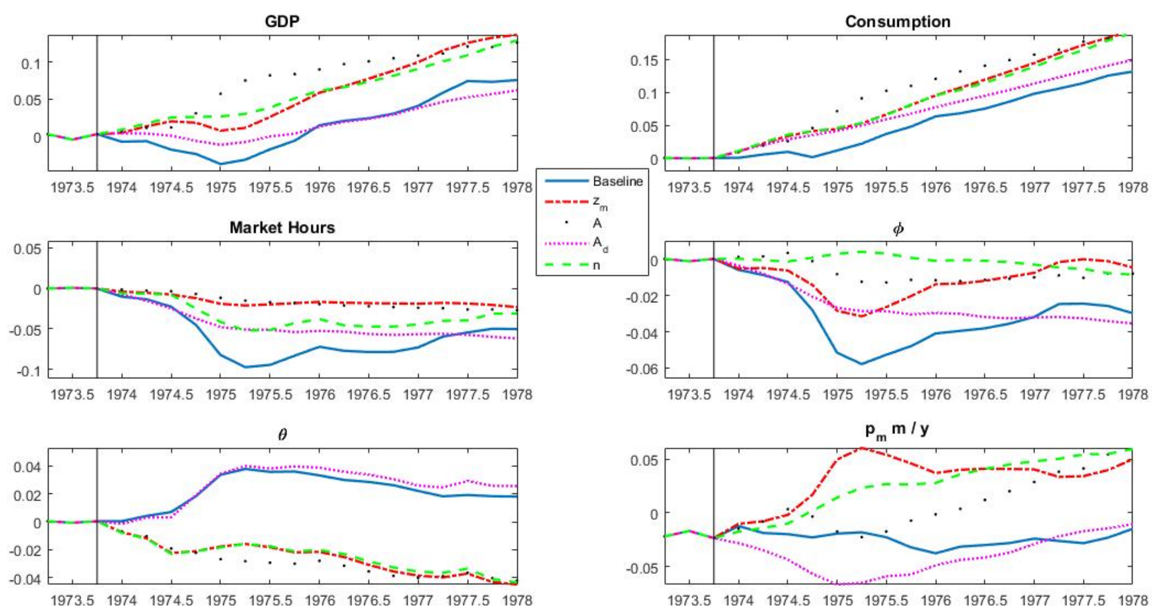

Figure E.8: 1973.IV recession due to each shock

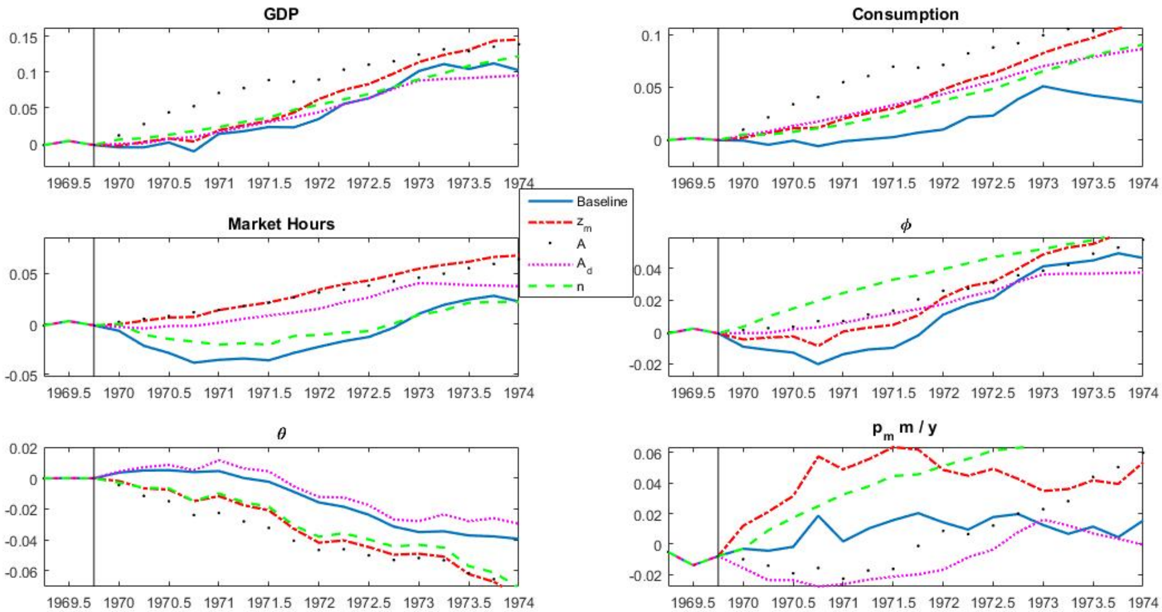

Figure E.9: 1969.IV recession due to each shock

F Data and Summary of Parametrization

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 (BLS). 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.