Geromichalos, Athanasios; Herrenbrueck, Lucas; Salyer, Kevin

Working Paper

A search-theoretic model of the term premium

Working Paper, No. 13-8

Provided in Cooperation with:
University of California Davis, Department of Economics

Suggested Citation: Geromichalos, Athanasios; Herrenbrueck, Lucas; Salyer, Kevin (2013) : A search-theoretic model of the term premium, Working Paper, No. 13-8, University of California, Department of Economics, Davis, CA

This Version is available at:
https://hdl.handle.net/10419/79687

UCDAVIS
DEPARTMENT OF ECONOMICS
Working Paper Series

A Search-Theoretic Model of the Term Premium

Athanasios Geromichalos
UC Davis
Lucas Herrenbrueck
UC Davis
Kevin Salyer
U.C. Davis

June 14, 2013

Paper # 13-8

A consistent empirical feature of bond yields is that term premia are, on average, positive. That is, investors in long term bonds receive higher returns than investors in similar (i.e.\ same default risk) shorter maturity bonds over the same holding period. The majority of theoretical explanations for this observation have viewed the term premia through the lens of the consumption based capital asset pricing model. In contrast, we harken to an older empirical literature which attributes the term premium to the idea that short maturity bonds are inherently more liquid. The goal of this paper is to provide a theoretical justification of this concept. To that end, we employ a model in the tradition of modern monetary theory extended to include assets of different maturities. Short term assets always mature in time to take advantage of random consumption opportunities. Long term assets do not, but agents may liquidate them in a secondary asset market, characterized by search-and-bargaining frictions a la Duffie, Garleanu, and Pedersen (2005). In equilibrium, long term assets have higher rates of return to compensate agents for their relative lack of liquidity. Consistent with empirical findings, our model predicts a steeper yield curve for assets that trade in less liquid secondary markets.

Department of Economics
One Shields Avenue
Davis, CA 95616
(530)752-0741

http://www.econ.ucdavis.edu/working_search.cfm

A Search-Theoretic Model of the Term Premium

Athanasios Geromichalos
University of California - Davis

Lucas Herrenbrueck
University of California - Davis

Kevin Salyer
University of California - Davis

This Version: June 2013

ABSTRACT

A consistent empirical feature of bond yields is that term premia are, on average, positive. That is, investors in long term bonds receive higher returns than investors in similar (i.e. same default risk) shorter maturity bonds over the same holding period. The majority of theoretical explanations for this observation have viewed the term premia through the lens of the consumption based capital asset pricing model. In contrast, we harken to an older empirical literature which attributes the term premium to the idea that short maturity bonds are inherently more liquid. The goal of this paper is to provide a theoretical justification of this concept. To that end, we employ a model in the tradition of modern monetary theory extended to include assets of different maturities. Short term assets always mature in time to take advantage of random consumption opportunities. Long term assets do not, but agents may liquidate them in a secondary asset market, characterized by search-and-bargaining frictions *à la* Duffie, Gârleanu, and Pedersen (2005). In equilibrium, long term assets have higher rates of return to compensate agents for their relative lack of liquidity. Consistent with empirical findings, our model predicts a steeper yield curve for assets that trade in less liquid secondary markets.

JEL Classification: E31, E43, E52, G12

Keywords: monetary-search models, liquidity, over-the-counter markets, yield curve

Email: ageromich@ucdavis.edu, herrenbrueck@ucdavis.edu, kdsalyer@ucdavis.edu.
We are grateful to Darrell Duffie, Richard Green, Guillaume Rocheteau, Ina Simonovska, Adrien Verdelhan, and Randall Wright for useful comments and suggestions.

1 Introduction

An established feature of bond yields is that, on average, investors in long term bonds receive a higher return than investors in similar (e.g. same default risk and tax treatments) short term bonds over a given holding period. As is well known, this observation, which implies the presence of positive term premia, is inconsistent with the expectations hypothesis of the term structure of interest rates. To quote Backus, Gregory, and Zin (1989): "One of the more firmly established facts of financial economics is that the expectations hypothesis of the term structure of interest rates cannot account for observed fluctuations in multiperiod bond returns".

Given this observation, the challenge to economists is to define and quantify the nature of risks and market frictions associated with the purchase of long term bonds. The majority of responses to this challenge have identified the term premium as a risk premium and employed the consumption based capital asset pricing framework in their analyses. The qualitative nature of asset risk is therefore characterized by the covariance between investors' stochastic discount factors and asset returns. Early papers (Backus, Gregory, and Zin (1989) and Salyer (1990)) demonstrated, however, that, due in large part to the autocorrelation properties of inflation, a standard intertemporal asset pricing model would produce counterfactual negative term premia for nominally denominated bonds. More recently, Piazzesi and Schneider (2007) combined Epstein-Zin preferences with a richer stochastic model of inflation and consumption growth (one that includes long-run risks as described by Bansal and Yaron (2004)) and demonstrated that these features can indeed produce positive, time-varying nominal term premia consistent with observation.

While certainly insightful, we do not view the above explanation as wholly satisfactory. For one, positive term premia over holding horizons as short as one quarter are routinely observed (Backus, Gregory, and Zin (1989)). It is difficult to see this as a response to changes in long run risk as explained by Piazzesi and Schneider (2007). In addition, variation in bond supply plays no role in the Piazzesi and Schneider (2007) analysis (which uses the often-employed representative agent assumption). As the recent work by Krishnamurthy and Vissing-Jorgensen (2012) demonstrates, bond supply does appear to affect bond yields.[1] Hence, as an alternative explanation, one that we view as complementary rather than competing, we present a model of the term premium which is based upon the inherent liquidity differentials between bonds of different maturities.

Our explanation is quite simple and intuitive: in a world with uncertain consumption expenditures (e.g. health costs, home repairs, etc.), there may be times when some of the illiquid assets held by households must be converted to liquid assets in order to finance unexpectedly high consumption expenditures. However, in our model, the asset market in which this conver-

[1] Moreover, the Federal Reserve's use of quantitative easing was entirely predicated on the idea that bond supply affects yields and the slope of the yield curve.

sion takes place is not a perfectly competitive Walrasian market but, instead, has the features of an over-the-counter (OTC) market characterized by search and bargaining, such as that described in Duffie, Gârleanu, and Pedersen (2005). Moreover, to provide a precise definition of asset liquidity, we employ the search-theoretic framework of Lagos and Wright (2005), in which a subset of goods markets are decentralized and anonymous, so that a medium of exchange emerges naturally. In this environment, an asset's value is determined both by fundamentals (as in a standard asset pricing model) and by the ease with which the asset can be used for consumption purchases; i.e. its liquidity properties.

In our model, short term assets mature in time to take advantage of random consumption opportunities in markets with imperfect credit. Thus, short term assets are closer substitutes to money, and, in equilibrium, they typically carry a liquidity premium, which reflects the assets' ability to facilitate trade in the frictional goods market. Long term assets (i.e. assets that do not mature in the current period) cannot serve directly as means of payment. However, agents who carry these assets (and have an opportunity to consume) can visit an OTC financial market, where they can try to exchange them for liquid assets, i.e. for a portfolio of money and assets that are about to mature. We show that, although long term assets cannot be used directly to purchase consumption goods, in equilibrium, they can carry an *indirect* liquidity premium, which reflects their ability to help agents avoid the cost of carrying liquid assets (a cost which is strictly positive in all monetary equilibria).

Our model delivers some theoretical predictions which are consistent with empirical observation. The main result of the paper is that, if asset supply is not too large, in a sense to be made precise, long maturity assets will sell at a discount (i.e. "haircut") relative to short maturity assets; that is, investors must be compensated for holding the relatively illiquid long maturity assets so that a positive term premium emerges in equilibrium. We illustrate that this term premium is closely linked to the search and bargaining frictions characterizing the OTC asset market. In particular, we show that the only way to obtain a zero term premium is if the agents who have an opportunity to consume are guaranteed to trade in the OTC market, *and* if they can extract the whole surplus generated from OTC trade.

One of the key insights of our model is that the issue price of long maturity assets is crucially (and positively) affected by the liquidity of the secondary asset market, i.e. how easy it is for agents to liquidate these assets in the OTC. To highlight the importance of this channel, we extend the baseline model to include a second set of assets that only differ from the original ones in that they cannot be traded in secondary markets (i.e. agents have to hold them to maturity). We show that the issue price of long maturities will be higher for the assets that can be traded in secondary financial markets, thus reflecting a liquidity premium. Krishnamurthy and Vissing-Jorgensen (2012) compare the yields on 6-month FDIC-insured certificates of deposit (CDs) and 6-month treasury bills over the 1984-2008 period and provide direct evidence in support of our

finding.[2] In particular, they report that the spread was 2.3 percentage points on average, and they attribute this spread to the higher liquidity of T-bills. Moreover, the authors report that the spread between the yields of the two assets is negatively related to the supply of T-bills, a result that we also obtain in our theoretical analysis. The prediction that bond yields are influenced by the liquidity of secondary markets is also consistent with Gürkaynak, Sack, and Wright's (2010) analysis of the yield curve for inflation-indexed Treasury debt (i.e. TIPS). In particular, they demonstrate that TIPS yields have, in general, fallen as market liquidity (measured by trading volume) in the TIPS market has increased.

Our framework also allows us to compare the price of freshly issued (on-the-run) short term assets with the price of older assets (off-the-run) which mature on the same date. Conventional wisdom suggests that the yields on assets with identical streams of dividends should be equal. However, Warga (1992) documents that the return of an off-the-run portfolio exceeds, on average, the return of an on-the-run portfolio with similar duration. Our model is consistent with this observation. Intuitively, in our analysis, the sellers of off-the-run assets are agents who received an opportunity to consume, and who are desperate for liquidity and, thus, more willing to sell assets at a lower price. Vayanos and Weill (2008) also provide a theoretical explanation of the "on-the-run phenomenon", by building a model where on the on-the-run bonds are more liquid (i.e. easier to sell) because they constitute better collateral for borrowing in the repo market. Although our model attempts to suggest an alternative explanation for the on-the-run phenomenon, it shares a common feature with Vayanos and Weill (2008): the assumption that asset trade takes place in OTC markets. This assumption is essential for both models' ability to capture the on-the-run phenomenon.

It is important to point out that our examination of how liquidity and maturity interact to influence asset yields can be viewed from a broader and, perhaps, more fundamental perspective. Namely, the classic textbook explanations for the existence of an upward sloping yield curve (for example, see Mishkin (2007)) are typically based on two concepts: the preferred habitats theory and a liquidity premium. Recently, Vayanos and Vila (2009) provided a theoretical rationale for segmented markets (i.e. preferred habitats), and we see our work as providing a similar rationale for the notion of a liquidity premium. In the Vayanos and Vila (2009) analysis, the role of arbitrageurs and their risk aversion properties play a critical role in linking the yields on bonds with different maturities. In our analysis, asset maturity is associated with liquidity which, when combined with the aforementioned shocks to consumption and decentralized asset markets, produces an always positive term premium. Hence our model helps to explain why the yield curve has, on average, a positive slope. In general, it is clear that a combination of factors such as risk, market segmentation, and liquidity all influence the term structure, and

[2] Notice that the assets under consideration have the same maturity and the same default risk (they are default-free). However, unlike T-bills, CDs have to be held to maturity.

we see our contribution as providing a theoretical basis of the last factor, i.e. liquidity.[3]

The literature on the term structure is vast but the profession is fortunate to have several excellent survey articles. In particular, Gürkaynak and Wright (2012) provide a nice overview of the testable implications of the expectations hypothesis and the lack of empirical support for them. The literature which characterizes the term premium as a risk premium is also discussed thoroughly. As the authors point out, their review is very much from a macroeconomics perspective (our analysis is as well). For a discussion of analyses of the term structure from a finance tradition, the reader is referred to Singleton (2009) and Piazzesi (2010).

This paper is also related to a growing literature that focuses on the liquidity properties of assets other than fiat money. Lagos and Rocheteau (2008), Geromichalos, Licari, and Suarez-Lledo (2007), Lagos (2011), Lester, Postlewaite, and Wright (2012), and Jacquet and Tan (2012) all develop different versions of monetary-search models where assets compete with money as media of exchange, and study the important question of how monetary policy affects asset prices. Some recent papers exploit the idea that assets can carry liquidity premia in an attempt to offer a new perspective for looking at long-standing asset-related puzzles. Examples of such papers include Lagos (2010) (equity premium and risk-free rate puzzles) and Geromichalos and Simonovska (2011) (asset home bias puzzle). Finally, a number of papers, such as Boel and Camera (2006), Berentsen, Camera, and Waller (2007), and Berentsen, Huber, and Marchesiani (2011) explore the idea that assets may carry liquidity premia because they allow agents to rebalance their money holdings after a consumption opportunity arises. The present paper is uniquely identified from the aforementioned papers, in that it models explicitly the secondary financial market as an OTC (rather than Walrasian) market, and it considers assets of different maturities, thus focusing on the term premium.

The rest of the paper is organized as follows. In Section 2, we first present an intuitive description of the model followed by the formal presentation. In Section 3, we define equilibrium and characterize its properties in a simple version of the model with two maturities. Section 4 shows how the main results of Section 3 can be generalized in an environment with any number of maturities. Section 5 offers some concluding comments.

2 The Model

Before presenting a detailed description of the economy, we first describe the basic setup in order to highlight the critical features of the model. Specifically, we employ an infinite horizon,

[3] To our knowledge, the most closely related theoretical analysis which examines the role of asset market frictions on bond yields is that of He and Milbradt (2012). Their analysis is quite different from ours, however, in that the focus is on default decisions by firms and how this impacts (and is impacted by) trade frictions in the corporate bond market. Also, our definition of liquidity is primarily related to an asset's possible role as a medium of exchange as opposed to their characterization in terms of the frictions in over-the-counter asset markets.

discrete-time economy in which each period is divided into three subperiods. The subperiods are identified by their markets. At the beginning of the period, a financial market meets in which assets with different maturities are traded. This market resembles the over-the-counter market of Duffie, Gârleanu, and Pedersen (2005), and we refer to it as the OTC market. In the second subperiod, agents meet in decentralized markets characterized by anonymous, bilateral trades, as in Lagos and Wright (2005). We refer to this market as the LW market. In the final subperiod, trade of newly available assets and goods takes place in a centralized (i.e. Walrasian) market, which we refer to as the CM. More details of these markets are given below.

The economy contains two types of infinitely-lived agents, buyers and sellers, defined by their actions in the LW market. The agents' types are permanent. Buyers are distributed over the unit interval, i.e. have unit measure. At the beginning of each period, a fraction $\ell < 1$ of buyers learn that they have an opportunity to purchase the good sold in the LW market; these are denoted as the C-type buyers. The remaining measure of buyers, $1-\ell$, denoted N-types, do not purchase goods in the LW market of the current period (i.e. they will be inactive buyers in the current period). To keep the analysis simple, we assume that all C-types match with a seller in the LW market, and we choose the measure of sellers to equal ℓ.

The role of the markets is as follows. Buyers, who are the agents that make all the interesting decisions in our model, leave the CM in the previous period and then find out whether they will be consuming in the LW market. Since trade is anonymous in that market, C-type buyers need liquid assets (defined below) to finance their purchases. The OTC market is strategically placed before the LW market opens, but after the uncertainty regarding consumption in the LW market has been resolved, in order to allow agents who might be short of liquidity to exchange illiquid assets for liquid assets.

Returning to the structure of the model, all agents discount the future between periods (but not subperiods) at rate $\beta \in (0,1)$. Buyers consume in the second and third subperiods and supply labor in the third subperiod. Their preferences for consumption and labor within a period are given by $\mathcal{U}(X,H,q)$, where X, H represent consumption and labor in the CM, respectively, and q consumption in the LW market. Sellers consume only in the CM and produce in both the CM and the LW market. Their preferences are given by $\mathcal{V}(X,H,q)$, where X, H are as above, and q stands for hours worked in the LW market (implicitly we assume that sellers receive disutility from working (h), and that the technology of producing q is linear, i.e. $q = h$). Following Lagos and Wright (2005), we adopt the functional forms

$$\mathcal{U}(X,H,q) = U(X) - H + u(q),$$
$$\mathcal{V}(X,H,h) = U(X) - H - c(h).$$

Assume that u, U are twice continuously differentiable with $u(0) = 0$, $u' > 0$, $u'(0) = \infty$, $u'(\infty) = 0$, $U' > 0$, $u'' < 0$, and $U'' \leq 0$. For simplicity, we set $c(h) = h$, but this is not cru-

cial for any results. Let q^* denote the optimal level of output in any LW market meeting, i.e. $q^* \equiv \{q : u'(q^*) = 1\}$. Also, there exists $X^* \in (0, \infty)$ such that $U'(X^*) = 1$, with $U(X^*) > X^*$.

In the third subperiod, all agents consume and produce a general good or fruit. The supply of this good comes from two sources: labor supplied by agents, and the output (i.e. dividend) of assets maturing that period. Agents have access to a technology that transforms one unit of labor into one unit of the fruit. Each period, the economy is endowed with a set of trees, as in Lucas (1978), that deliver a real dividend (i.e. fruit) at different dates (maturities). Each share of a tree of maturity $i \in \{1, ..., N\}$ purchased in period t, delivers 1 unit of fruit in period $t + i$. For reasons that will become clear later, we assume that the fruit is delivered before the LW market opens. Agents can store the fruit at no cost between the second and the third subperiod (when they consume it), but the fruit is perishable between time periods. Agents can purchase any amount of shares of a tree of maturity i at the ongoing market price $\psi_{i,t}$. The supply of trees that mature in i periods is denoted by $A_i > 0$, and it is fixed over time. This supply includes newly issued i-period trees and older trees that mature at the same date.

In addition to trees, agents also trade fiat money in the CM. The market price of money is denoted by φ_t. Its supply is controlled by a monetary authority, and it evolves according to $M_{t+1} = (1 + \mu)M_t$, with $\mu > \beta - 1$. New money is introduced, or withdrawn if $\mu < 0$, via lump-sum transfers to buyers in the CM. Money has no intrinsic value, but it is storable, divisible, and recognizable by all agents. Hence, it can serve as a medium of exchange in the LW market, and help bypass the frictions created by anonymity and the lack of a double coincidence of wants.

The anonymous, bilateral exchanges in the LW market are characterized by take-it-or-leave-it offers made by the buyer to the seller. Due to the anonymity in that market, exchange has to be *quid pro quo*, and the only objects that can serve as means of payment are money and fruit that has already been delivered; critically, claims to trees that mature in future periods can not be used for payment. These assumptions are discussed in detail in Section 2.1 below.

If a C-type agent finds herself in need of additional liquidity, she can visit the OTC market and search for a trading partner (an N-type) who might hold some liquid assets (i.e. money or trees that pay out in the current period) that she will not use in this period's LW market. Hence, gains from trade can be generated by C-types selling a portfolio of long term assets (i.e. assets that do not mature in the current period) in exchange for money and assets that mature in the current period. We assume that a matching function, $f(\ell, 1 - \ell) \leq \min\{\ell, 1 - \ell\}$, brings together C-types and N-types. The function f is homogeneous of degree one and increasing in both arguments. Within each match, the terms of trade are determined through proportional bargaining, following Kalai (1977), and the C-type's bargaining power is given by $\lambda \in (0, 1)$.

Throughout the paper we focus on steady-state equilibria, and most of the equilibrium analysis is carried out with respect to the asset prices ψ_i, $i = 1, ..., N$. When we wish to make statements regarding the interest rate of the various assets, we use the standard (textbook) formula

that links the price and interest rate of an asset. In particular, we have

$$\psi_i = \frac{1}{(1+r_i)^i}, \quad \textit{for all } i = 1, ..., N. \tag{1}$$

2.1 Discussion of the Physical Environment

The assumption according to which fruit that has already been delivered can serve as a means of payment in the LW market is a very important and, we believe, reasonable one. This assumption aims to capture the idea that assets that have already matured are "as good (liquid) as money". At the same time, it allows us to work with real rather than nominal assets, which is standard in the money-search and the theoretical finance literature. Employing real assets is desirable for a number of reasons. First, it simplifies the analysis and allows us to sharply characterize equilibrium for the various values of A_i, $i = 1, ..., N$. Second, it allows us to remain agnostic with respect to the identity of the issuer of these assets and her objectives.[4]

In this paper, we take as given that assets with long maturities cannot be used as media of exchange, and we study asset prices subject to this restriction on the physical environment. However, there might be deeper reasons why agents prefer to use money versus financial assets in order to carry out transactions. For example, Rocheteau (2011) and Lester, Postlewaite, and Wright (2012) both consider environments that do not place any restrictions on which objects can serve as media of exchange. They show that, if there is asymmetric information regarding the future returns of financial assets, fiat money (or, in our case, assets that have already matured) will arise endogenously as a superior medium of exchange in bilateral meetings.

In our model, all assets are first traded (issued) in Walrasian markets. This is a methodological innovation, due to Lagos and Wright (2005), which (together with quasi linearity of preferences) gives rise to degenerate asset holding distributions and ensures tractability. We believe that this setup is not only convenient, but also realistic. Many assets that are eventually traded in OTC secondary markets are indeed issued in primary markets with competitive characteristics. For instance, US Treasury Bills are issued through single-price auctions, in order "to minimize the government's costs [...] by promoting broad, competitive bidding" (Garbade and Ingber (2005)).[5] It should be pointed out that *only* newly issued assets are traded in the CM. Agents who wish to sell assets that mature in future periods can only do so in the OTC market. In other words, all secondary asset trade is over-the-counter.

[4] Of course, taking the supply of assets seriously by introducing, say, a government who issues assets in order to finance its debt would be extremely interesting. However, this is beyond the scope of the present paper.

[5] Also, recently, many large corporations have been using an internet-based Dutch auction in order to sell their shares to the general public for the first time. In particular, San Francisco based investment bank W.R. Hambrecht & Co. persuaded companies such as Google, Overstock.com, Clean Energy Fuels, Boston Beer Company, and others, to use this auction-based initial public offering, a process now known as an OpenIPO.

3 Equilibrium in the Model with $N = 2$

In this section, we focus on the case of two maturities. This version of the model conveys all the economic ideas that we wish to highlight, and delivers the most important results of the paper. Subsequently, Section 4 shows that an upward sloping yield curve will also be obtained in the general model with $N > 2$.

3.1 Value Functions

We begin with the description of the value functions in the CM. For a typical buyer, the state variables are the following. First, the units of money, m, that she brings into the CM. Second, the units of assets of maturity $N = 2$, a_2, that she bought in the previous period, and which will mature in the forthcoming period. Third, the units of fruit, d, that she received as dividend earlier in the period, i.e. before the LW market opened. The latter could have been delivered either from long term assets issued two periods ago, or from short term assets issued in the last period. The Bellman's equation is given by

$$\begin{aligned} W(m,d,a_2) = &\max_{X,H,\hat{m},\hat{a}_1,\hat{a}_2} \left\{U(X) - H + \beta \mathbb{E}\left\{\Omega^i(\hat{m},\hat{a}_1,\hat{a}_2)\right\}\right\} \\ &\text{s.t. } X + \varphi\hat{m} + \psi_1(\hat{a}_1 - a_2) + \psi_2\hat{a}_2 = H + \varphi(m + \mu M) + d, \end{aligned}$$

and subject to $\hat{a}_1 - a_2 \geq 0$. In the expression above, variables with hats denote next period's choices, and the term $\mathbb{E}$ denotes the expectations operator. The function Ω^i represents the value function in the OTC market for a buyer of type $i = \{C, N\}$, to be described in more detail below. It is important to highlight that we have defined $\hat{a}_1$ as the amount of *all* assets that mature in the next period (which is analogous to our definition of the supply of assets that mature in the next period). Hence, the amount of newly issued short term assets purchased by the agent is $\hat{a}_1 - a_2$, and we require $\hat{a}_1 - a_2 \geq 0$. This constraint simply enforces the assumption that agents cannot sell off-the-run short term bonds in the CM (recall from Section 2 that only newly issued assets are traded in the CM).

Some observations are in order. First, it can be easily verified that, at the optimum, $X = X^*$. Using this fact and replacing H from the budget constraint into W yields

$$\begin{aligned} W(m,d,a_2) &= U(X^*) - X^* + \varphi(m + \mu M) + d + \psi_1 a_2 \\ &+ \max_{\hat{m},\hat{a}_1,\hat{a}_2} \left\{-\varphi\hat{m} - \psi_1\hat{a}_1 - \psi_2\hat{a}_2 + \beta\mathbb{E}\left\{\Omega^i(\hat{m},\hat{a}_1,\hat{a}_2)\right\}\right\}. \end{aligned} \tag{2}$$

A standard feature of models that build on Lagos and Wright (2005) is that the optimal choice of the agent does not depend on the current state (due to the quasi-linearity of $\mathcal{U}$). This is also true here, with the exception that the range of admissible choices for $\hat{a}_1$ is restricted by the state

variable a_2. Moreover, as is standard in this types of models, the CM value function is linear. In fact, W is linear in the variable $z \equiv \varphi m + d$, which captures the total real balances of the buyer. This property will greatly simplify the analysis in what follows. We collect all the terms in (2) that do not depend on the state variables, and we write

$$W(z, a_2) = \Lambda + z + \psi_1 a_2, \tag{3}$$

where the definition of Λ is obvious.

Next, consider a seller's value function in the CM. It is well-known that in monetary models where the identity of agents (as buyers or sellers) is fixed over time, sellers will typically not leave the CM with a positive amount of asset holdings. The intuition behind this result is simple. In monetary models, assets will, in general, be priced above the "fundamental value", reflecting liquidity premia. Agents who know with certainty that they will not have an opportunity to consume in the forthcoming LW market (i.e. sellers) will not be willing to pay this premium. Here we take this result as given (for a detailed discussion, see Rocheteau and Wright (2005)). Therefore, when a seller enters the CM, she will only hold money or fruit that she received during trade in the preceding LW market, and the CM value function is given by

$$\begin{aligned} W^S(z) = \max_{X,H} \left\{ U(X) - H + \beta V^S \right\} \\ \text{s.t. } X = H + z, \end{aligned}$$

where V^S denotes the seller's value function in next period's LW market, to be discussed below.[6] Sellers also choose $X = X^*$, and W^S will also be linear. More precisely,

$$W^S(z) = U(X^*) - X^* + z + V^S \equiv \Lambda^S + z. \tag{4}$$

Consider now the value functions in the LW market. Let q denote the quantity of special good produced, and π the real value of *money and fruit* that change hands during trade in the LW market. These terms will be determined in Section 3.1.1. The LW value function for a buyer who enters that market with portfolio (z, a_2) is given by

$$V(z, a_2) = u(q) + W(z - \pi, a_2), \tag{5}$$

and the LW value function for a seller (who enters with no assets) is given by

$$V^S = -q + W^S(\pi).$$

[6] Since the seller leaves the CM with no assets, she will never visit the OTC market.

Finally, consider the value functions in the OTC market. After leaving the CM, and before the OTC market opens, buyers learn whether they will have a chance to access this period's LW market (C-types) or not (N-types). This chance will occur with probability $\ell \in (0,1)$. The expected value for the typical buyer, before she enters the OTC market, is given by

$$\mathbb{E}\left\{\Omega^i(m,a_1,a_2)\right\} = \ell\,\Omega^C(m,a_1,a_2) + (1-\ell)\,\Omega^N(m,a_1,a_2). \tag{6}$$

In the OTC market, C-type buyers, who may want additional liquid assets, are matched with N-type buyers, who may hold liquid assets that they will not use in the current period. Hence, trade in the OTC involves C-types giving up long term assets for short term assets and cash. Given the matching function $f(\ell, 1-\ell)$, define the matching probabilities for C-types and N-types as $\alpha_{_C} \equiv f(\ell,1-\ell)/\ell$ and $\alpha_{_N} \equiv f(\ell,1-\ell)/(1-\ell)$, respectively. Let χ denote the units of long term assets that the C-type transfers to the N-type, and ζ the real value of liquid assets that the C-type receives in return. These terms will be determined in Section 3.1.2. Then,

$$\Omega^C(m,a_1,a_2) = \alpha_{_C}\,V\left(z+\zeta, a_2-\chi\right) + (1-\alpha_{_C})V(z,a_2), \tag{7}$$

$$\Omega^N(m,a_1,a_2) = \alpha_{_N}\,W\left(z-\zeta, a_2+\chi\right) + (1-\alpha_{_N})\,W(z,a_2). \tag{8}$$

Notice that N-type buyers proceed directly to the CM. Also, notice that our definition $z \equiv \varphi m + d$ allows us to write V as a function of (z, a_2) (recall that each unit of a_1 will deliver one unit of fruit between the OTC and LW subperiods).

We now proceed to the description of the terms of trade in the LW and the OTC markets.

3.1.1 Bargaining in the LW Market

Consider a meeting between a C-type buyer with real balances z and long term assets a_2, and a seller who, as we have argued, holds no real balances or assets as she enters the LW subperiod. The two parties bargain over a quantity q, to be produced by the seller, and a real payment π, to be made to the seller. The buyer makes a take-it-or-leave-it offer, maximizing her surplus subject to the seller's participation constraint. The bargaining problem can be described by

$$\begin{aligned} &\max_{\pi,q}\left\{u(q) + W(z-\pi,a_2) - W(z,a_2)\right\} \\ &\quad \text{s.t.} \;\; -q + W^S(\pi) - W^S(0) = 0, \end{aligned}$$

and the constraint $\pi \leq z$. Taking advantage of the linearity of W, W^S (equations (3) and (4)), allows us to simplify the bargaining problem to

$$\max_{\pi,q} \{u(q) - \pi\}$$
$$\text{s.t. } q = \pi,$$

and $\pi \leq z$. The solution to the bargaining problem is described in the following lemma.

Lemma 1. *The solution to the bargaining problem is given by*

$$q(z) = \pi(z) = \min\{q^*, z\}$$

Proof. This result is very standard in the literature, hence, the proof is omitted. For a detailed proof see Geromichalos *et al* or Lester *et al*. □

The solution to the bargaining problem is very intuitive. The only variable that affects the solution is the buyer's real balances. As long as the buyer carries q^* or more, the first-best quantity q^* will always be exchanged. On the other hand, if $z < q^*$, the buyer does not have enough liquidity to induce the seller to produce q^*. In this case, the buyer will give up all her real balances, $\pi(z) = z$, and the seller will produce the quantity of good that satisfies her participation constraint, that is $q = \pi(z) = z$.

3.1.2 Bargaining in the OTC market

We now study the terms of trade in the OTC market. Although buyers are ex ante identical, only C-types will get an opportunity to consume in the forthcoming LW market, and may wish to acquire more liquid assets, i.e. assets that can be used as means of payment in that market. In a sense, money and short term assets (that are about to mature) are more valuable in the hands of C-types rather than N-types, because only the former types can take advantage of these assets' property to serve as media of exchange. The role of the OTC market is to allow agents to take advantage of the gains from trade generated as C-types increase their liquidity.

Consider a meeting in the OTC between a C-type with portfolio (z, a_2), and an N-type with portfolio $(\tilde{z}, \tilde{a_2})$. Let χ denote the units of long term assets that the C-type transfers to the N-type, and let ζ represent the real value of liquid assets received by the C-type. Also, let S^i, $i = \{C, N\}$, denote the surplus of type i, and $\lambda \in [0, 1]$ the bargaining power of the C-type. With proportional bargaining, the objective is to choose χ, ζ in order to maximize S^C, subject to: a) the constraint that the ratio S^C/S^N should be equal to the ratio $\lambda/(1-\lambda)$, and b) the feasibility

constraints $\chi \leq a_2$ and $\zeta \leq \tilde{z}$. The terms S^i are given by[7]

$$S^C \equiv V(z+\zeta, a_2-\chi) - V(z, a_2),$$
$$S^N \equiv W(\tilde{z}-\zeta, \tilde{a}_2+\chi) - W(\tilde{z}, \tilde{a}_2).$$

Substituting for W, V from (3) and (5) in the expression above, and exploiting Lemma 1 (the LW bargaining solution) allows us to write

$$S^C = u\left(\min\{q^*, z+\zeta\}\right) - u\left(\min\{q^*, z\}\right) + \zeta + \min\{q^*, z\} - \min\{q^*, z+\zeta\} - \psi_1\chi,$$
$$S^N = \psi_1\chi - \zeta.$$

For the rest of Section 3.1.2, we focus on the case in which $z < q^*$, since, if the opposite is true, the C-type is carrying the maximum possible liquidity, and there is no trade in the OTC market that can generate a positive surplus. Similarly, we know that all OTC trades will involve a transfer of real balances ζ, such that $z+\zeta \leq q^*$.[8] These observations imply that $\min\{q^*, z+\zeta\} = z+\zeta$ and $\min\{q^*, z\} = z$, and allow us to write

$$S^C = u\left(z+\zeta\right) - u\left(z\right) - \psi_1\chi.$$

Therefore, the OTC bargaining problem can be written as

$$\max_{\chi,\zeta} \left\{ u\left(z+\zeta\right) - u\left(z\right) - \psi_1\chi \right\} \tag{9}$$
$$\text{s.t. } u\left(z+\zeta\right) - u\left(z\right) - \psi_1\chi = \frac{\lambda}{1-\lambda}\left(\psi_1\chi - \zeta\right), \tag{10}$$

and $\chi \leq a_2$, $\zeta \leq \tilde{z}$.

Notice that if we solve equation (10) with respect to $\psi_1\chi$, we obtain

$$\psi_1\chi = \zeta + (1-\lambda)\left[u(z+\zeta) - u(z) - \zeta\right]. \tag{11}$$

This expression states that the real value of assets that the N-type receives as payment equals the value of real balances she is giving up, ζ, plus a fraction $1-\lambda$ (her bargaining power) of the surplus generated by the OTC transaction, i.e. the term $u(z+\zeta) - u(z) - \zeta$. Substituting for $\psi_1\chi$

[7] Since the C-type has a consumption opportunity, she will proceed to the LW market with an additional ζ units of real balances, but also with her long term asset holdings reduced by the amount χ. The N-type will proceed directly to the CM with less money and short term assets, but with more long term assets.

[8] Two comments are in order. First, as we shall see in Section 3.4, equilibria where agents carry $z \geq q^*$ do exist. However, in these equilibria no OTC trade takes place. Since here the objective is to describe the terms of trade in the OTC market, we focus on the interesting case where $z < q^*$. Second, restricting attention to trades such that $z+\zeta \leq q^*$ simply means that the C-type will never acquire more real balances than she needs in order to attain q^*.

from (11) into (9), simplifies the bargaining problem to

$$\max_{\chi,\zeta} \lambda \left\{ u\left(z+\zeta\right) - u\left(z\right) - \zeta \right\},$$

subject to the constraint in (10), and the feasibility constraints $\chi \leq a_2$, $\zeta \leq \tilde{z}$. As is standard in proportional bargaining, the C-type's surplus equals a fraction λ of the total surplus generated by OTC trade. The solution to this bargaining problem is described in the following lemma.

Lemma 2. *Consider a meeting in the OTC market between a C-type and an N-type with portfolios* (z, a_2) *and* $(\tilde{z}, \tilde{a}_2)$*, respectively, and define the cutoff level of long term asset holdings*

$$\bar{a}(z,\tilde{z}) \equiv \frac{1}{\psi_1}\Big\{ (1-\lambda)\left[u\left(\min\{z+\tilde{z}, q^*\}\right) - u(z)\right] + \lambda \min\{q^* - z, \tilde{z}\}\Big\}. \tag{12}$$

Then, the solution to the bargaining problem is given by

$$\chi(z,\tilde{z},a_2) = \begin{cases} \bar{a}(z,\tilde{z}), & \text{if } a_2 \geq \bar{a}(z,\tilde{z}), \\ a_2, & \text{if } a_2 < \bar{a}(z,\tilde{z}). \end{cases} \tag{13}$$

$$\zeta(z,\tilde{z},a_2) = \begin{cases} \min\{q^* - z, \tilde{z}\}, & \text{if } a_2 \geq \bar{a}(z,\tilde{z}), \\ \zeta^a(z,a_2), & \text{if } a_2 < \bar{a}(z,\tilde{z}), \end{cases} \tag{14}$$

where we have defined

$$\zeta^a(z,a_2) \equiv \Big\{ \zeta : (1-\lambda)\left[u\left(z+\zeta\right) - u(z)\right] + \lambda\zeta = \psi_1 a_2 \Big\}. \tag{15}$$

Proof. □

If $z + \tilde{z} \geq q^*$, the C-type should receive exactly as many real balances as she lacks in order to purchase q^* in the forthcoming LW market, i.e. $\zeta = q^* - z$. In contrast, if $z + \tilde{z} < q^*$, even if the two types pull together all their real balances, these will not allow the C-type to attain q^*. The second best, requires the N-type to give all her real balances to the C-type, $\zeta = \tilde{z}$. However, one should also ask whether the C-type has sufficient amounts of a_2 to compensate the N-type for the transfer of liquidity. This critical level of assets is defined in (12), and it depends on whether $z + \tilde{z}$ exceeds q^* or not. If $a_2 \geq \bar{a}(z,\tilde{z})$, the C-type is not constrained, and $\zeta = \min\{q^* - z, \tilde{z}\}$, as described above. In this case, the C-type gives up exactly $\bar{a}(z,\tilde{z})$ units of long term assets. When $a_2 < \bar{a}(z,\tilde{z})$, the C-type will not be able to purchase the desired amount of liquid assets, given by $\min\{q^* - z, \tilde{z}\}$. In that case, she will give away all her long maturity assets, $\chi = a_2$, and the transfer of real balances will be determined such that the sharing rule of the surplus between the two parties (equation (11)) is satisfied. Notice that the N-type's long term asset holdings do not affect the bargaining solution.

Having established the bargaining solutions in the OTC and LW markets, we now proceed to the derivation of the buyer's objective function and the description of her optimal behavior.

3.2 Objective Function and Optimal Behavior

In this sub-section, we wish to characterize the optimal portfolio choice of the representative buyer. We will do so by deriving the buyer's objective function, i.e. a function that summarizes the buyer's cost and benefit from choosing any particular portfolio $(\hat{m}, \hat{a}_1, \hat{a}_2)$. Substitute (7) and (8) into (6), and lead the resulting expression by one period to obtain

$$\begin{aligned}\mathbb{E}\left\{\Omega^i(\hat{m}, \hat{a}_1, \hat{a}_2)\right\} = f\, V\left(\hat{z} + \zeta, \hat{a}_2 - \chi\right) + (\ell - f)\, V\left(\hat{z}, \hat{a}_2\right) \\ + fW(\hat{z} - \tilde{\zeta}, \hat{a}_2 + \tilde{\chi}) + (1 - \ell - f)\, W(\hat{z}, \hat{a}_2),\end{aligned} \tag{16}$$

where f is a shortcut for $f(\ell, 1-\ell)$. Since each unit of asset that matures in the next period pays one unit of fruit before the LW market opens, it is understood that $\hat{z} = \hat{\varphi}\hat{m} + \hat{a}_1 = \hat{\varphi}\hat{m} + d$.

The four terms in (16) represent the benefit for a buyer who holds a portfolio $(\hat{m}, \hat{a}_1, \hat{a}_2)$ and turns out to be a matched C-type (with probability f), an unmatched C-type (with probability $\ell - f$), a matched N-type (with probability f), or an unmatched N-type (with probability $1 - \ell - f$), respectively. The expressions χ, ζ, and $\tilde{\chi}, \tilde{\zeta}$ are implicitly described by the solution to the OTC bargaining problem. In particular,

$$\begin{aligned}\chi &= \chi(\hat{z}, \tilde{z}, \hat{a}_2), \qquad \zeta = \zeta(\hat{z}, \tilde{z}, \hat{a}_2), \\ \tilde{\chi} &= \chi(\tilde{z}, \hat{z}, \tilde{a}_2), \qquad \tilde{\zeta} = \zeta(\tilde{z}, \hat{z}, \tilde{a}_2).\end{aligned}$$

In these expressions, the first argument represents the C-type's real balances, the second argument represents the N-type's real balances, and the third argument stands for the C-type's long term asset holdings (recall from Lemma 2 that the N-type's long term asset holdings do not affect the bargaining solution). Terms with tildes stand for the representative buyer's beliefs about her potential counterparty's real balances and long term asset holdings in the OTC.[9]

Next, we substitute W and V from (3) and (5), respectively, into (16). We insert the term $\mathbb{E}\{\Omega^i(\hat{m}, \hat{a}_1, \hat{a}_2)\}$ into (2), and we focus on the terms inside the maximum operator of (2). We define the resulting expression as $J(\hat{m}, \hat{a}_1, \hat{a}_2)$, and we refer to it as the buyer's objective function. The objective function is further separated into a cost component and an expected-benefit component of carrying assets. We denote this expected benefit function by $G(\hat{z}, \hat{a}_2)$, recognizing that money and short term assets are perfect substitutes and combining them into a choice of

[9] For instance, $\tilde{\zeta} = \zeta(\tilde{z}, \hat{z}, \tilde{a}_2)$ stands for the amount of real balances that the agent will give away if she is a matched N-type. This term depends on her own real balances ($\hat{z}$), and the real balances ($\tilde{z}$) and long term asset holdings ($\tilde{a}_2$) of her trading partner (a C-type). The terms χ, ζ, and $\tilde{\chi}$ admit similar interpretations.

real balances, $z \equiv \hat{\varphi}\hat{m} + \hat{a}_1$. After some manipulations, one can verify that

$$\begin{aligned} J(\hat{m}, \hat{a}_1, \hat{a}_2) &= -\varphi\hat{m} - \psi_1\hat{a}_1 - \psi_2\hat{a}_2 + \beta\, G(\hat{\varphi}\hat{m} + \hat{a}_1, \hat{a}_2), \\ G(\hat{z}, \hat{a}_2) &= f\,\left[u\left(\hat{z} + \zeta\right) + \hat{\psi}_1\left(\hat{a}_2 - \chi\right)\right] + (\ell - f)\left[u\left(\hat{z}\right) + \hat{\psi}_1\hat{a}_2\right] \\ &\quad + f\,\left[\hat{z} - \tilde{\zeta} + \hat{\psi}_1\left(\hat{a}_2 + \tilde{\chi}\right)\right] + (1 - \ell - f)\left(\hat{z} + \hat{\psi}_1\hat{a}_2\right). \end{aligned}$$

The negative terms in the definition of J represent the cost of purchasing various amounts of the three assets available in the economy.[10] The four terms in the definition of G admit similar interpretations as their counterparts in equation (16). For instance, the first term represents the expected benefit of a C-type buyer who matches in the OTC market. This agent will increase her LW consumption by an amount equal to ζ, but she will also go to next period's CM with her long term assets reduced by χ. In this event, the terms ζ, χ will depend on her own choices $\hat{z}, \hat{a}_2$, and on her trading partner's (who is an N-type) real balances, $\tilde{z}$.

We can now proceed with the examination of the buyer's optimal choice of $(\hat{z}, \hat{a}_2)$. We will do so for any possible money and asset prices, and for any given beliefs about other agents' money and asset holdings. We focus on prices that satisfy $\varphi > \beta\hat{\varphi}$, since we know that this will be always true in steady-state monetary equilibria with $\mu > \beta - 1$ (unless $\varphi = \hat{\varphi} = 0$, interpreted as a non-monetary equilibrium). Also, the asset prices have to satisfy $\psi_1 \geq \beta$ and $\psi_2 \geq \beta\hat{\psi}_1$, since violation of these conditions would generate an infinite demand for the assets. The optimal behavior of the buyer is described formally in Lemma 3 below. Here, we provide an intuitive explanation of the buyer's optimal portfolio choice.

The objective function of the buyer depends on the terms $\chi, \zeta, \tilde{\chi}$, and $\tilde{\zeta}$, which, in turn, depend on the bargaining protocol in the OTC market. Given the buyer's beliefs $(\tilde{z}, \tilde{a}_2)$, she can end up in different branches of the bargaining solution, depending on her own choices of $(\hat{z}, \hat{a}_2)$. In general, the domain of the objective function can be divided into five regions in $(\hat{z}, \hat{a}_2)$-space, arising from three questions: (*i*) When the C-type and the N-type pool their real balances in the OTC market, can they achieve the first-best in the LW market? (*ii*) If I am a C-type, do I carry enough assets to compensate the N-type? (*iii*) If I am an N-type, do I expect a C-type to carry enough assets to compensate me? These regions are illustrated in Figure 1, and are described in detail as follows (for this discussion it is important to recall the definition of the asset cutoff term $\bar{a}(\,\cdot\,,\,\cdot\,)$ from Lemma 2).

1. $\hat{z} \in (q^* - \tilde{z}, q^*)$ and $\hat{a}_2 > \bar{a}(\hat{z}, \tilde{z})$.

   In this region, the real balance holdings of the C-type and the N-type together allow the C-

[10] In the objective function, the term $-\psi_1\hat{a}_1$ appears as the cost of purchasing assets that mature in the next period. However, we know that the term $\psi_1 a_2$ is also present in the agent's value function (see equation (2)), so that, practically, the cost of leaving the CM with $\hat{a}_1$ units of assets that mature tomorrow is $-\psi_1(\hat{a}_1 - a_2)$. However, the term $\psi_1 a_2$ only has a level effect, and it does not change the optimal choice of $\hat{a}_1$, with the exception that any choice of the agent should respect the restriction $\hat{a}_1 - a_2 \geq 0$.

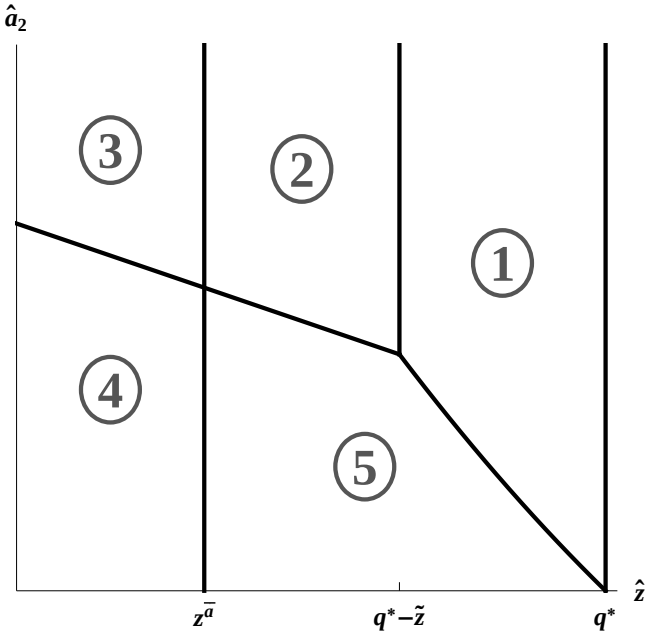

Figure 1: Regions of the individual choice problem.

type to purchase q^* in the LW market. If the agent is a C-type, her long term asset holdings are enough to compensate an N-type for her real balances. If the agent is an N-type, the potential counterparty may or may not carry enough long term assets to purchase the first-best level of real balances, $q^* - \tilde{z}$, but that is a level effect on $G(\hat{z}, \hat{a}_2)$ and does not affect the optimal choice.[11]

2. $\hat{z} < q^* - \tilde{z}$, $\hat{a}_2 > \bar{a}(\hat{z}, \tilde{z})$, but $\tilde{a}_2 < \bar{a}(\tilde{z}, \hat{z})$.
   Here there are not enough real balances in an OTC match to allow the C-type to purchase q^* in the LW market. If a C-type, the agent carries enough long term assets to buy all the real balances of the N-type, but if an N-type, the agent does not expect the C-type counterparty to carry enough long term assets to buy all of the agent's real balances.

3. $\hat{z} < q^* - \tilde{z}$, $\hat{a}_2 > \bar{a}(\hat{z}, \tilde{z})$, and $\tilde{a}_2 > \bar{a}(\tilde{z}, \hat{z})$.
   There are not enough real balances in an OTC match to allow the C-type to purchase q^* in the LW market. In an OTC match, the agent expects all of the real balances of the N-type to be traded for less than all of the long term assets of the C-type (regardless of whether the buyer in question is the C or the N-type).

4. $\hat{z} < q^* - \tilde{z}$, $\hat{a}_2 < \bar{a}(\hat{z}, \tilde{z})$, but $\tilde{a}_2 > \bar{a}(\tilde{z}, \hat{z})$.
   There are not enough real balances in an OTC match to allow the C-type to purchase q^* in

[11] Since here the objective is to describe the buyer's optimal behavior, we focus on how different choices of $(\hat{z}, \hat{a}_2)$ lead to different branches of the OTC bargaining protocol. In Region 1, the buyer is not certain whether her C-type counterparty is asset constrained or not, but she also does not care. What determines Region 1 is that, conditional on being an N-type, the buyer's real balances never affect the terms of trade.

the LW market. If a C-type, the agent does not carry enough long term assets to buy all the real balances of the N-type, but if an N-type, the agent expects the C-type counterparty to carry enough long term assets to buy all her real balances.

5. $\hat{a}_2 < \bar{a}(\hat{z}, \tilde{z})$, and either $\tilde{a}_2 < \bar{a}(\tilde{z}, \hat{z})$ or $\hat{z} \in (q^* - \tilde{z}, q^*)$.
   If a C-type, the agent does not carry enough long term assets to buy all the real balances of the N-type. If an N-type, the agent expects not to give away all of her real balances, either because the C-type counterparty does not carry enough long term assets to afford it, or because she does not need all of those real balances. This distinction does not affect the buyer's optimal choice.

We can now state the most important facts about the optimal choice of the representative buyer:

Lemma 3. *Taking prices, $(\varphi, \hat{\varphi}, \psi_1, \hat{\psi}_1, \psi_2)$, and beliefs, $(\tilde{z}, \tilde{a})$, as given, and assuming that $\mu > \beta - 1$ and $\varphi > 0$, then the optimal choice of the representative agent, $(\hat{m}, \hat{a}_1, \hat{a}_2)$, satisfies:*

a) Money and short term assets are perfect substitutes. If $\psi_1 > \varphi/\hat{\varphi}$, then $\hat{a}_1 = 0$, and if $\psi_1 < \varphi/\hat{\varphi}$, then $\hat{m} = 0$.

b) If the optimal choice $(\hat{z}, \hat{a}_2)$ is strictly within any region, or on the boundary of Region 1 with any other region, and if $\psi_1 = \varphi/\hat{\varphi}$, it satisfies the first-order condition $\nabla J = \mathbf{0}$, or equivalently, $\beta \nabla G = (\psi_1, \psi_2)$.

c) If $\varphi > \beta\hat{\varphi}$ and $\psi_2 = \beta\hat{\psi}_1$, the optimal $\hat{z}$ is unique, and any $\hat{a}_2$ is optimal as long as $(\hat{m}, \hat{a})$ is in Regions 1, 2, or 3 (or on their boundaries).

d) If $\varphi > \beta\hat{\varphi}$ and $\psi_2 > \beta\hat{\psi}_1$, the optimal choice is unique, and it lies in Regions 4 or 5 or on their boundaries with Regions 2 and 3.

Moreover, let $G^i(\hat{z}, \hat{a}_2)$, $i = 1, ..., 5$, denote the expected benefit function in Region i, and $G^i_k(\hat{z}, \hat{a}_2)$, $k = 1, 2$, its derivative with respect to the k-th argument. Then, we have:

$$G^1_1(\hat{z}, \hat{a}_2) = 1 + (\ell - \lambda f)\left[u'(\hat{z}) - 1\right], \tag{17}$$

$$G^2_1(\hat{z}, \hat{a}_2) = 1 + (\ell - \lambda f)\left[u'(\hat{z}) - 1\right] + \lambda f\left[u'(\hat{z} + \tilde{z}) - 1\right], \tag{18}$$

$$G^3_1(\hat{z}, \hat{a}_2) = 1 + (\ell - \lambda f)\left[u'(\hat{z}) - 1\right] + f\left[u'(\hat{z} + \tilde{z}) - 1\right], \tag{19}$$

$$G^4_1(\hat{z}, \hat{a}_2) = 1 + \ell\left[u'(\hat{z}) - 1\right] + (1 - \lambda) f\left[u'(\hat{z} + \tilde{z}) - 1\right] + \lambda f \frac{u'\left[\hat{z} + \zeta^a(\hat{z}, \hat{a}_2)\right] - u'(\hat{z})}{(1 - \lambda) u'\left[\hat{z} + \zeta^a(\hat{z}, \hat{a}_2)\right] + \lambda}, \tag{20}$$

$$G^5_1(\hat{z}, \hat{a}_2) = 1 + \ell\left[u'(\hat{z}) - 1\right] + \lambda f \frac{u'\left[\hat{z} + \zeta^a(\hat{z}, \hat{a}_2)\right] - u'(\hat{z})}{(1 - \lambda) u'\left[\hat{z} + \zeta^a(\hat{z}, \hat{a}_2)\right] + \lambda}, \tag{21}$$

$$G^1_2(\hat{z}, \hat{a}_2) = G^2_2(\hat{z}, \hat{a}_2) = G^3_2(\hat{z}, \hat{a}_2) = \hat{\psi}_1, \tag{22}$$

$$G_2^4(\hat{z},\hat{a}_2) = G_2^5(\hat{z},\hat{a}_2) = \hat{\psi}_1 \left\{ 1 - f + f \, \frac{u'\left[\hat{z} + \zeta^a(\hat{z},\hat{a}_2)\right]}{(1-\lambda)u'\left[\hat{z} + \zeta^a(\hat{z},\hat{a}_2)\right] + \lambda} \right\}, \tag{23}$$

where $\zeta^a(\,\cdot\,,\,\cdot\,)$ *was defined in* (15).

Proof. See the appendix. □

If the price of long term assets satisfies $\psi_2 = \beta\hat{\psi}_1$, the cost of carrying long term assets is zero and, therefore, it would be suboptimal for the buyer to be in a region where her long term assets would not allow her to afford the optimal quantity of liquid assets, when a C-type. As a result, when $\psi_2 = \beta\hat{\psi}_1$, the buyer never chooses a portfolio in the interior of Regions 4 and 5. If $\psi_2 > \beta\hat{\psi}_1$, carrying long term assets is costly. The optimal choice of the buyer is characterized by the first-order conditions and, graphically, it lies within Regions 4 or 5. For any set of prices which satisfy $\psi_1 = \varphi/\hat{\varphi} > \beta$, the optimal choice of real balances is uniquely characterized by the first-order condition with respect to either $\hat{m}$ or $\hat{a}_1$.

Next, we demonstrate the determination of the demand for real balances. This demand, D_z, is plotted in Figure 2 against the ratio $\varphi/(\beta\hat{\varphi})$, which captures the holding cost of real balances.[12] The level of long term asset holdings is kept fixed at $\hat{a}_2 = a_2'$ indicated in the lower panel of the figure. Notice that the lower panel of Figure 2 is identical to Figure 1. Aligning the two plots vertically, allows the reader to easily indicate which region of Figure 1 the buyer will find herself in, for any choice of $\hat{z}$, and for a given value of $\hat{a}_2$. For $\hat{a}_2 = a_2'$, any $\hat{z} > \bar{z}_{1,5}$ implies that the buyer is in Region 1, and in this region one additional unit of real balances has the following benefits: a) it serves as a store of value, if the buyer is an N-type; b) it allows the buyer to purchase more goods in the LW market, if she is an unmatched C-type; and c) it allows the buyer to reduce her demand for the N-type's real balances, if she is a matched C-type.

As $\hat{z}$ decreases below $\bar{z}_{1,5}$, the buyer finds herself in Region 5. The function D_z is continuous and exhibits a kink at $\bar{z}_{1,5}$, and the slope of D_z is steeper to the left of $\bar{z}_{1,5}$. To illustrate this property, consider how the marginal benefit of carrying one additional unit of real balances changes as the buyer moves from Region 1 to Region 5. Recall that, in Region 1, an additional unit of real balances has three effects. The effects indicated by (a) (store of value when N-type) and (b) (higher marginal utility when unmatched C-type) are still valid as we enter Region 5. What differs is the marginal benefit of real balances when the buyer is a matched C-type: in this event, an additional unit of $\hat{z}$ does not only allow her to reduce her demand for the N-type's real balances (effect (c) above), but it allows her to acquire extra purchasing power in the forthcoming LW market.[13] Hence, the slope of the demand function is higher (in absolute

[12] More precisely, $\varphi/(\beta\hat{\varphi})$ captures the holding cost of money. However, in any equilibrium where $\hat{m}, \hat{a}_1 > 0$, the holding cost of the two liquid assets will necessarily be the same.

[13] Put simply, in the event that the buyer is a matched C-type, if she is in Region 1, she will be able to buy q^* anyway. Bringing more $\hat{z}$ will not change the quantity of LW consumption (it will still be equal to q^*), but it will allow her to rely less heavily on the N-type's liquid assets (which could be quite important, especially if the terms

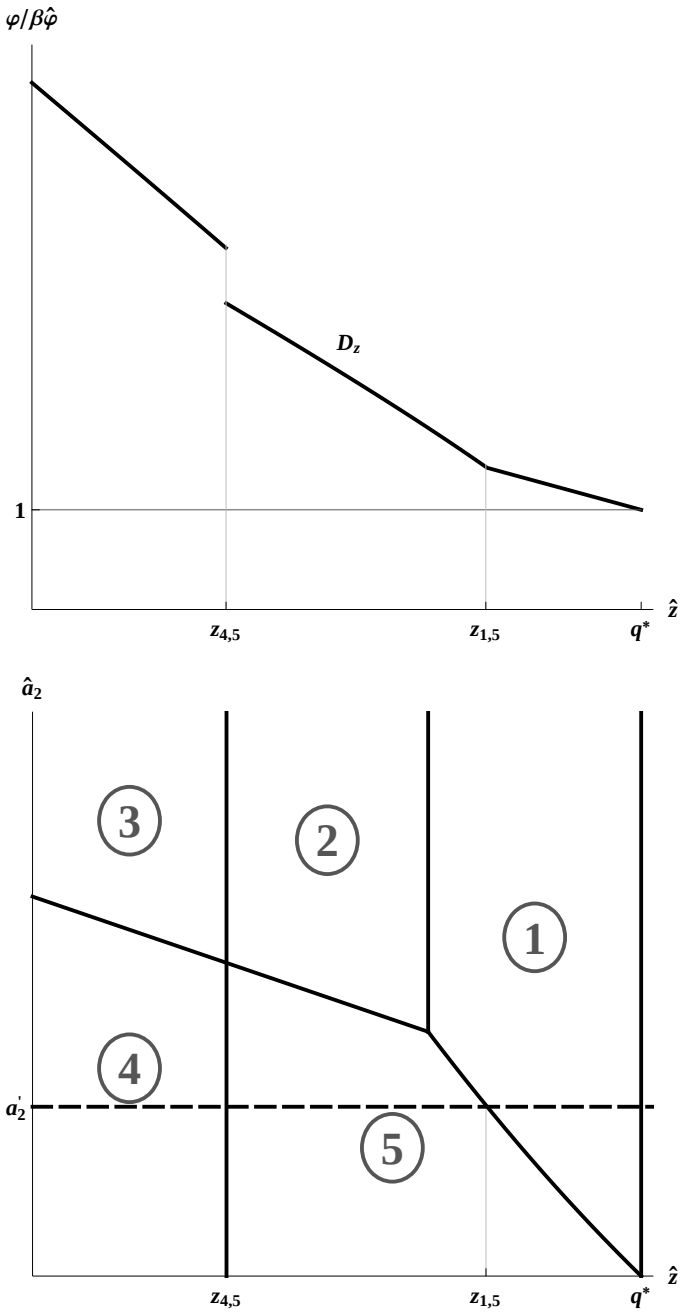

Figure 2: Demand for real balances given long term asset holdings a_2'.

value) for $\hat{z}$ in the range $[\bar{z}_{4,5}, \bar{z}_{1,5})$ compared to $[\bar{z}_{1,5}, q^*)$. Also, from (17) and (21), we have

$$G_1^5 - G_1^1 = \lambda f \left\{ u'(\hat{z}) - 1 + \frac{u'\left[\hat{z} + \zeta^a(\hat{z}, \hat{a}_2)\right] - u'\left(\hat{z}\right)}{(1-\lambda)u'\left[\hat{z} + \zeta^a(\hat{z}, \hat{a}_2)\right] + \lambda} \right\},$$

of trade are against her in the OTC market, i.e. if λ is low). On the other hand, in Region 5, the matched C-type cannot buy q^* even after purchasing all the real balances of the N-type that she can afford. In this case, bringing more $\hat{z}$ strictly increases her LW consumption.

which is what differentiates D_z on the two sides of $\bar{z}_{1,5}$. When $\hat{z} = \bar{z}_{1,5}$, we have $\hat{z} + \zeta^a = q^*$, and it is easy to verify that $G_1^5 - G_1^1 = 0$. As a result, D_z exhibits a kink but is continuous at $\bar{z}_{1,5}$.

Finally, D_z exhibits a jump, at $\bar{z}_{4,5}$, the value of $\hat{z}$ that, given $\hat{a}_2 = a_2'$, brings the agent on the boundary of Regions 4 and 5 (and in the interior of Region 5 if and only if $\hat{z} > \bar{z}_{4,5}$). Consider the behavior of D_z in a neighborhood of this point. In Region 5, an additional unit of real balances serves as a store of value, if the buyer is an N-type, and it allows the buyer to purchase more goods in the LW, if she is a C-type (matched or unmatched). These effects remain valid as we enter into Region 4. However, in Region 4 a new effect arises, which is relevant when the buyer is a matched N-type. In this region, the C-type counterparty can afford to buy all of the buyer's real balances, hence the buyer's choice of $\hat{z}$ affects the OTC terms of trade even when she is an N-type (assuming that $\lambda < 1$). Specifically, the less real balances the buyer brings, the more desperate the C-type will be for those real balances, and the more long term assets she will be willing to give up in order acquire them. Formally, (21) and (20) imply that

$$G_1^4 - G_1^5 = (1 - \lambda) f \left[u' \left(\hat{z} + \tilde{z} \right) - 1 \right].$$

Since $\bar{z}_{4,5} + \tilde{z} < q^*$, this term is strictly positive when $\hat{z} = \bar{z}_{4,5}$, provided that $\lambda < 1$. This gap between the values of G_1^4 and G_1^5 reflects the discontinuity of D_z at $\bar{z}_{4,5}$.

Having established the optimal behavior of the representative buyer, we are now ready to discuss equilibrium.

3.3 Definition of Equilibrium and Preliminary Results

We restrict attention to symmetric steady-state equilibria, where all agents choose the same portfolios, and the real variables of the model remain constant over time. Since, in steady-state, the real money balances do not change over time, we have $\varphi/\hat{\varphi} = 1 + \mu$ in any monetary equilibrium where $\hat{\varphi} > 0$. In such an equilibrium, we must also have $\psi_1 = \hat{\psi}_1 = 1 + \mu$, since money and short term assets are perfect substitutes. To see how this simple relationship emerges, one just needs to equate the rate of return on money, $\hat{\varphi}/\varphi - 1 = (1+\mu)^{-1} - 1$, with the rate of return on the short term assets, $\hat{\psi}^{-1} - 1$. Before stating the definition of a steady-state equilibrium, it is important to notice that symmetry rules out Regions 2 and 4 in Figure 1, since a C-type and an N-type buyer are ex ante identical.

In order to characterize equilibrium sharply, we use three restrictions. First, aggregate real balances Z are the combination of real money (φM) and maturing short term bonds (A_1), so $Z \geq A_1$. Second, recall the constraint that agents cannot sell off-the-run short term assets in the CM; at most, they can refrain from buying newly issued short term bonds. So the post-CM holdings of short term bonds (equal to A_1 in symmetric equilibrium) must exceed the pre-CM holdings for every agent, including those of asset buyers in the preceding OTC market (equal to

$A_2 + \chi(Z, Z, A_2)$ in symmetric equilibrium), thus $A_1 \geq A_2 + \chi(Z, Z, A_2)$. Together, these restrictions rule out combinations of low Z and high A_2. With the following restriction on structural parameters, Region 3 is ruled out altogether:

$$\frac{1 + (1-\lambda)\left[\frac{u(q^*) - u(q^*/2)}{q^*/2} - 1\right]}{1 + (\ell - \lambda f)\left[u'(q^*/2) - 1\right]} > \frac{\beta}{2}. \tag{24}$$

Graphically, this restriction guarantees that, in Figure 3, the line $Z = 2A_2$ lies below the boundary of Regions 1, 3, and 5 (or equivalently $\bar{A}_1 \geq q^*/2$). While it is possible to construct a counterexample, the restriction is satisfied for a wide range of utility functions if f is close to ℓ (C-types have a high probability of matching). Henceforth, we assume that the model's parameters satisfy the inequality stated in (24).[14]

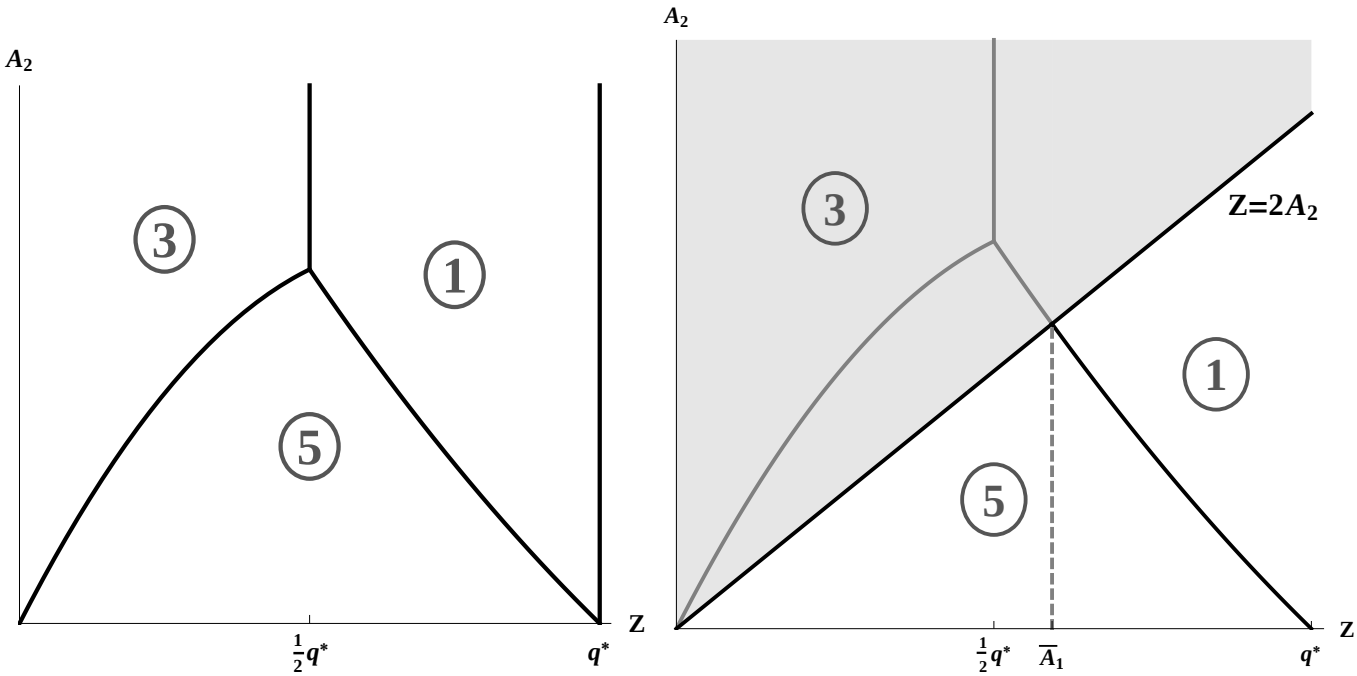

Figure 3: Aggregate regions of equilibrium, in terms of real balances.

With the above constraints satisfied, only two regions remain on aggregate:

1. Agents carry enough real balances and long term assets so that, when matched in the OTC market, the C-type can acquire sufficient liquidity in order to achieve the first-best in the LW market.

5. Agents carry so few long term assets that, when matched in the OTC, the C-type will sell all of her long term assets but not obtain enough of the N-type's real balances in order to achieve the first-best in the LW market.

[14] More importantly, this restriction simplifies the analysis without ruling out any interesting results. In particular, Results 1-3 stated in Section 3.4 would still go through without this additional assumption.

These regions are described in Figure 4, and we will refer to them as the "aggregate regions", as opposed to the "individual regions" described in Figure 1. In general, Region 1 represents the region of abundance of the long maturity asset, and Region 5 represents the region of scarcity.

Definition 1. A symmetric steady-state equilibrium is a list $\{\varphi, \psi_1, \psi_2, \chi, \zeta, Z, q_1, q_2\}$, where $Z = \varphi M + A_1$ represents the real balances, which are equal to the amount of good exchanged in the LW market when the buyer was not matched in the preceding OTC market, i.e. q_1. The term q_2 is the amount of good exchanged in the LW market when the buyer *was* matched. The equilibrium objects satisfy:

i. The representative buyer behaves optimally under the equilibrium prices ψ_1, ψ_2, φ, and, moreover, $\psi_1 = \hat{\psi}_1 = \varphi/\hat{\varphi} = 1 + \mu$ if $\hat{\varphi} > 0$.

ii. The equilibrium quantity q_2 is defined as the following function of Z:

$$q_2(Z) = \begin{cases} q^*, & \text{in Region 1,} \\ \tilde{q}(Z), & \text{in Region 5,} \end{cases}$$

where $\tilde{q}$ solves $(1-\lambda)\left[u(\tilde{q}) - u(Z)\right] + \lambda\left(\tilde{q} - Z\right) = \psi_1 A_2$.

iii. The terms of OTC trade (χ, ζ) satisfy (13) and (14) evaluated at the aggregate quantities Z and A_2.

iv. Markets clear at symmetric choices, and expectations are rational: $\hat{m} = (1+\mu)M$, $\hat{z} = \tilde{z} = Z$, $\hat{a}_1 = A_1$, and $\hat{a}_2 = \tilde{a}_2 = A_2$.

Lemma 4. *Define the function* $Z(\mu, A_1) \equiv \max\{A_1, \{Z : (1+\mu)/\beta = 1 + (\ell - \lambda f)(u'(Z) - 1)\}\}$. *If* $\mu > \beta - 1$ *and* $A_1 \geq A_2 + \chi[Z(\mu, A_1), Z(\mu, A_1), A_2]$ *are satisfied, then a symmetric steady-state equilibrium exists and is unique.*

Proof. See the appendix. □

Having formally described the definition of a steady-state equilibrium and guaranteed its existence and uniqueness, the next task is to characterize such equilibria. Ultimately, we wish to describe the equilibrium variables as functions of the exogenous supply parameters A_1, A_2 and the policy parameter μ. Thus, before we state the main results, it is useful to describe the aggregate regions in terms of the parameter μ rather than Z. This task becomes easier with the help of Figure 4. A formal description of the various curves that appear in this figure will follow (equations (25)-(28) below). First, we provide an intuitive interpretation. The following three observations are crucial.

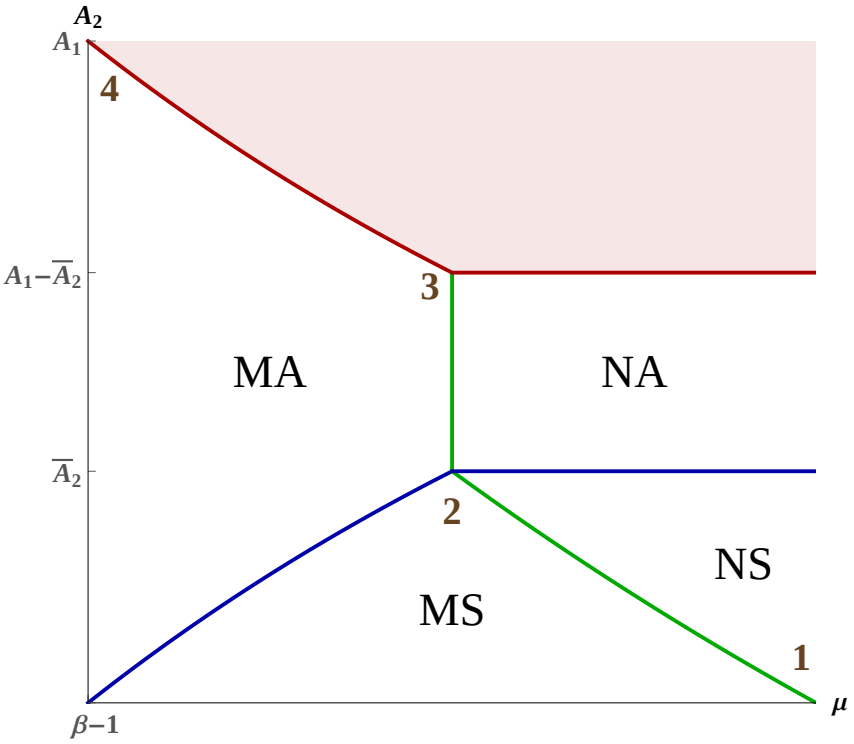

Figure 4: Aggregate regions of equilibrium, in terms of inflation.

a) The real balances $Z = \varphi M + A_1$ are decreasing in μ, but also bounded below by A_1. Consequently, if inflation exceeds a certain level $\bar{\mu}(A_1, A_2)$, then $\varphi = 0$, i.e. no monetary equilibrium exists. This critical level is a decreasing function of A_1. When A_2 is relatively plentiful, it does not affect the demand for real balances and the line $\mu = \bar{\mu}(A_1, A_2)$ is vertical (between points 2 and 3). However, when A_2 is relatively scarce, it does affect the terms of trade in the OTC market and hence the demand for real balances. As A_2 decreases, C-types must increasingly rely on their own real balances, so that, despite an increasing cost of holding money (as μ increases), a monetary equilibrium still exits. Thus, the line $\mu = \bar{\mu}(A_1, A_2)$ is downward sloping for low A_2 (between points 1 and 2). In summary, for any given (A_1, A_2), increasing inflation beyond $\bar{\mu}$ has no effect on real balances.

b) The line between the origin and point 2 is the inverted image of the boundary of Regions 1 and 5 in Figure 3, i.e. it separates the parameter space in a way that for any A_2 north of the line, long term assets are abundant in the OTC market. This line slopes upwards because higher inflation both reduces the amount of real balances and increases the need to trade in the OTC market, hence making A_2 more likely to be scarce. As we move east of point 2, we enter the non-monetary region, and the real balances are independent of μ. Hence, the line that separates the space into the region of abundance or scarcity of the long term assets (in the OTC market) becomes a horizontal line (i.e. it depends only on the relative values of A_1, A_2, but not on μ).

c) Furthermore, we need to consider the constraint that agents are unable to sell off-the-run short term assets in the CM. In other words, we need to guarantee that every agent enters the

CM with an amount of off-the-run assets smaller than the amount of short term assets they leave the CM with. At the Friedman rule, i.e. for $\mu = \beta - 1$, no OTC trade will take place, and the relevant constraint is simply $A_2 \leq A_1$. Away from the Friedman rule, N-type agents will leave the OTC with an additional amount of assets, χ, which increases with inflation, so the constraint $A_2 \leq A_1 - \chi$ becomes more binding (between points 3 and 4). For $\mu \geq \bar{\mu}$ (east of point 3), real balances and χ are unaffected by μ and the constraint becomes a horizontal line.

In summary, combinations of parameters (A_2, μ) that lie in the shaded region in Figure 4 are ruled out. In the remaining parameter space, every point that lies on the west (east) of the green piece-wise curve is associated with monetary (non-monetary) equilibrium. Similarly, every point that lies on the north (south) of the blue piece-wise curve is associated with equilibria where the long term assets are abundant in the OTC market. Thus, every equilibrium necessarily lies in one of four distinct regions clearly marked in Figure 4: i) MA stands for *monetary* equilibrium where long term assets are *abundant* in the OTC market, ii) MS stands for *monetary* equilibrium where long term assets are *scarce* in the OTC market, iii) NA stands for *non-monetary* equilibrium where long term assets are *abundant* in the OTC market, and iv) NS stands for *non-monetary* equilibrium where long term assets are *scarce* in the OTC market.

In the last part of this subsection, we define a few equilibrium objects that will allow us to provide a sharp characterization of equilibrium in what follows. First, we define the cutoff level of long term asset supply for the non-monetary region (the mathematical representation of the horizontal segment of the blue line in Figure 4):

$$\bar{A}_2(A_1) \equiv \frac{1}{\beta} \frac{1 + (1-\lambda)\left[u(q^*) - u(A_1)\right]}{1 + (\ell - \lambda f)\left[u'(A_1) - 1\right]}. \tag{25}$$

Next, we define the upper bound of inflation consistent with monetary equilibrium (the mathematical representation of the green piece-wise curve in Figure 4). If $A_2 \geq \bar{A}_2(A_1)$, we have

$$\bar{\mu}(A_1, A_2) = \beta - 1 + \beta\left(\ell - \lambda f\right)\left[u'(A_1) - 1\right] \tag{26}$$

On the other hand, if $A_2 < \bar{A}_2(A_1)$, we have

$$\bar{\mu}(A_1, A_2) = \beta - 1 + \beta\left[\ell - \frac{\lambda f}{(1-\lambda)u'(A_1 + \zeta^N) + \lambda}\right]\left[u'(A_1) - 1\right] + \frac{\beta\lambda f[u'(A_1 + \zeta^N) - 1]}{(1-\lambda)u'(A_1 + \zeta^N) + \lambda}, \tag{27}$$

where ζ^N denotes the short term asset trading volume in the OTC market in the case of a non-monetary equilibrium, and it solves

$$\begin{aligned}(1-\lambda)\left[u(A_1 + \zeta^N) - u(A_1)\right] + \lambda\zeta^N = \\ \beta A_2\left\{1 + f\lambda\tfrac{u'(A_1+\zeta^N)-1}{(1-\lambda)u'(A_1+\zeta^N)+\lambda} + \left[\ell - f\lambda\tfrac{1}{(1-\lambda)u'(A_1+\zeta^N)+\lambda}\right]\left[u'(A_1) - 1\right]\right\}.\end{aligned} \tag{28}$$

Finally, it is straightforward to show that equilibrium real balances, Z, satisfy the following equations. In Region MA, we have

$$\frac{1+\mu}{\beta} = 1 + (\ell - \lambda f)\left[u'(Z) - 1\right], \tag{29}$$

and in Region MS, we have

$$\frac{1+\mu}{\beta} = 1 + \lambda f \frac{u'(Z+\zeta^M) - 1}{(1-\lambda)u'(Z+\zeta^M) + \lambda} + \left(\ell - \lambda f \frac{1}{(1-\lambda)u'(Z+\zeta^M) + \lambda}\right)\left[u'(Z) - 1\right], \tag{30}$$

where ζ^M denotes the short term asset trading volume in the OTC market in the case of a monetary equilibrium, and it solves

$$(1-\lambda)\left[u(Z+\zeta^M) - u(Z)\right] + \lambda\zeta^M = (1+\mu)A_2. \tag{31}$$

3.4 Characterization of Equilibrium

We are now ready to characterize equilibrium. We begin this subsection with an intuitive description of the results presented in Propositions 1, 2, and 3. The prevailing parameter in the analysis is the supply of maturing assets A_1. If this supply is plentiful, in a way to be made precise in Proposition 1, short term assets alone are enough to satisfy the liquidity needs of the economy (for trade in the LW market). In this case, there is no room for money and no role for OTC trade. On the other hand, if A_1 is not enough to satisfy the liquidity needs of the economy (which we consider the interesting case), a role for money arises (the lower the value of A_1, the bigger that role). By no-arbitrage, the short term asset price will be fully determined by the policy parameter μ, in particular $\psi_1 = 1+\mu$. Away from the Friedman rule, the equilibrium real balances will always be suboptimal ($Z < q^*$), and this has two important implications for asset prices. First, ψ_1 will carry a *liquidity premium* (i.e. $\psi_1 > \beta$), because the marginal unit of short term assets is not only a good store of value, but it can also increase consumption in the LW market. Second, with $Z < q^*$, trade in the OTC market becomes crucial. In this case, the long term assets can potentially also carry a liquidity premium, not because they can facilitate trade in the LW market, but because they can be used in the OTC market in order to purchase liquid assets. Naturally, ψ_2 will include a liquidity premium if the supply A_2 is relatively scarce, in the precise sense that the equilibrium falls in Regions MS or NS in Figure 4.

We now describe these results in a formal way.

Proposition 1. *If $A_1 \geq q^*$, the equilibrium is always non-monetary regardless of μ, no trade occurs in the OTC market, and asset prices always equal their fundamentals: $\psi_i = \beta^i$ for $i = 1, 2$.*

Proof. See the appendix. □

This result is similar in spirit to Geromichalos, Licari, and Suarez-Lledo (2007) and Lester, Postlewaite, and Wright (2012). When $A_1 \geq q^*$, the supply of short term assets suffices to cover the liquidity needs of the economy (i.e. the need for trade in the anonymous LW market). This has the following consequences. First, it is clear that in this economy there is no role for money: every LW meeting will always involve the exchange of the optimal amount of good, q^*. Second, since agents already bring with them sufficient liquidity in order to purchase q^*, there is no role for trade in the OTC market. Third, since short term assets are issued in a competitive market, ψ_1 will reflect the benefit of holding one additional unit of these assets. But since here $A_1 \geq q^*$, the marginal unit of short term assets is good only as a store of value, and not as a facilitator of trade in the LW market. Thus, the unique equilibrium price must be $\psi_1 = \beta$. Finally, with no trade in the OTC market, long term assets cannot possibly be valued for any (direct or indirect) liquidity properties, which simply means that $\psi_2 = \beta^2$.

Henceforth, we maintain the assumption $A_1 < q^*$. Proposition 2 describes equilibrium prices and how they are affected by monetary policy. Proposition 3 does the same for the equilibrium value of production in the LW market. For this discussion, it is important to recall the definitions in equations (25)-(31).

Proposition 2. *The equilibrium price of short term assets is given by* $\psi_1 = \min\{1+\mu, 1+\bar{\mu}(A_1, A_2)\}$. *The equilibrium price of long term assets depends on the value of* A_2. *We have two cases:*

Case 1: *If* $A_2 \geq \bar{A}_2(A_1)$, *then* $\psi_2 = \beta\psi_1$.

Case 2: *If* $A_2 < \bar{A}_2(A_1)$, *then there exists a cutoff* $\tilde{\mu}(A_2)$ *such that:*
a) For all $\mu \in (\beta - 1, \tilde{\mu}(A_2)]$, *we have* $\psi_2 = \beta\psi_1$;
b) For all $\mu \in (\tilde{\mu}(A_2), \bar{\mu}(A_1, A_2))$, *we have* $\psi_2 = \beta\rho(\mu, A_2)\psi_1$, *where* $\rho(\mu, A_2) \in (1, (1+\mu)/\beta)$ *is a strictly increasing function of* μ *and a strictly decreasing function of* A_2;
c) For all $\mu \geq \bar{\mu}(A_1, A_2)$, *we have* $\psi_2 = \beta\rho(\bar{\mu}, A_2)\psi_1$.

The term ρ *is given by*

$$\rho(\mu, A_2) = 1 + \lambda f \frac{u'(Z + \zeta^M) - 1}{(1-\lambda)u'(Z + \zeta^M) + \lambda}, \tag{32}$$

where ζ^M *is defined in (31).*

Proof. See the appendix. □

The results reported in Proposition 2 are highlighted in Figure 5. As pointed out earlier, in any monetary equilibrium (for $\mu < \bar{\mu}(A_1, A_2)$), by no-arbitrage, the rate of return on money and the short term asset has to be equal, implying that $\psi_1 = 1 + \mu$. An increase in μ makes the cost of holding money higher, and induces agents to replace money with the relatively cheaper short term asset, which is a perfect substitute. In equilibrium, this leads to an increase in the demand

for short maturities and their price ψ_1. However, if the monetary authority increases μ beyond the threshold $\bar{\mu}(A_1, A_2)$, the equilibrium becomes non-monetary, and any further increase in μ has no effect on asset prices (or any other equilibrium variables). For any $\mu > \beta - 1$, the price of short term assets carries a liquidity premium (i.e. $\psi_1 > \beta$), which reflects the assets' property to mature in time to take advantage of consumption opportunities in the LW market.

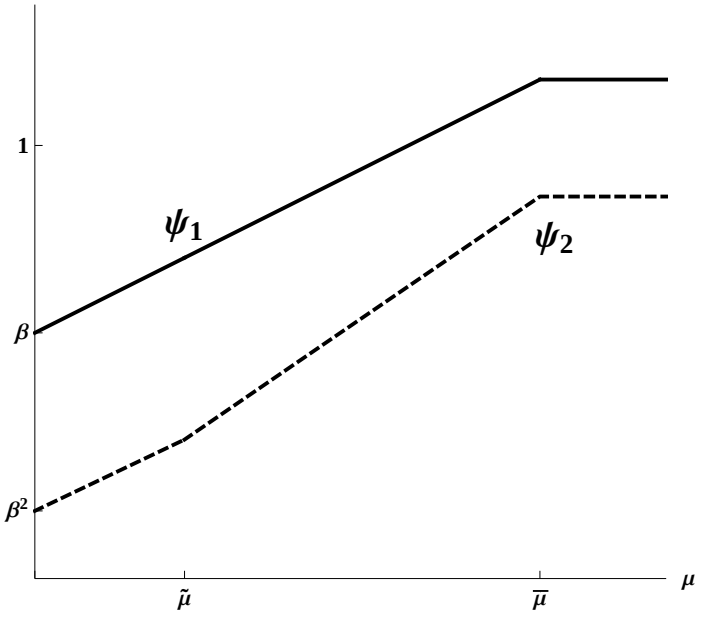

Figure 5: Equilibrium prices as functions of inflation.

The results that concern the equilibrium price of long term assets are even richer. Long term assets can be priced at a (liquidity) premium for two reasons:[15] first, because long term assets will become short term assets in the next period; second, because long term assets can be used in the OTC market in order to purchase liquid assets. In other words, the assets that do not mature today have *indirect* liquidity properties because they help agents bypass the cost of holding liquid assets (which is strictly positive when $A_1 < q^*$). If equilibrium lies in Regions MA or NA (i.e. the regions of abundance of long term assets in OTC trade), $\psi_2 = \beta\psi_1 > \beta^2$, and long term assets sell at a premium, but only because they will become short term assets in the next period. In contrast, if equilibrium lies in the regions of relative scarcity of A_2 (Regions MS or NS), an additional unit of long term assets can help agents purchase essential liquidity in the OTC (i.e. liquidity that allows them to boost LW consumption). This property is valued by agents, who are now willing to buy long maturities at a price greater than $\beta\psi_1$. Thus, the term $\rho > 1$ represents a premium that reflects the aforementioned indirect liquidity properties of long term assets. It is increasing in μ (within the regions of monetary equilibrium), precisely because the inflation tax that agents can avoid by holding long term assets is itself increasing

[15] To be clear, the long term asset price will include a liquidity premium, whenever ψ_2 exceeds the price that it would obtain if we were to close down the LW market (and, therefore, shut off any liquidity channel in the model). Clearly, this price would be the so-called fundamental value $\psi_2 = \beta^2$.

in μ. Similarly, ρ is decreasing in A_2, because the service that long term assets provide (helping agents avoid the cost of holding liquid assets) becomes more valuable when A_2 is more scarce.

Consider now the equilibrium values of the quantity of good in the LW market.

Proposition 3. *The equilibrium value of q_1 is always equal to Z. When $\mu < \bar{\mu}(A_1, A_2)$, then $\partial q_1/\partial \mu < 0$, and when $\mu > \bar{\mu}(A_1, A_2)$, then $\partial q_1/\partial \mu = 0$. Regarding the equilibrium value of q_2:*

Case 1: *If $A_2 \geq \bar{A}_2(A_1)$, then $q_2 = q^*$ for any $\mu > \beta - 1$.*

Case 2: *If $A_2 < \bar{A}_2(A_1)$, then for the same cutoff $\tilde{\mu}(A_2)$ as in Proposition 2:*

a) For all $\mu \in (\beta - 1, \tilde{\mu}(A_2)]$, $q_2 = q^*$;

b) For all $\mu \in (\tilde{\mu}(A_2), \bar{\mu}(A_1, A_2))$, $q_2 = Z + \zeta^M < q^*$ and q_2 is a strictly decreasing function of μ;

c) For all $\mu \geq \bar{\mu}(A_1, A_2)$, $q_2 = A_1 + \zeta^N < q^*$ which does not depend on μ.

Proof. See the appendix. □

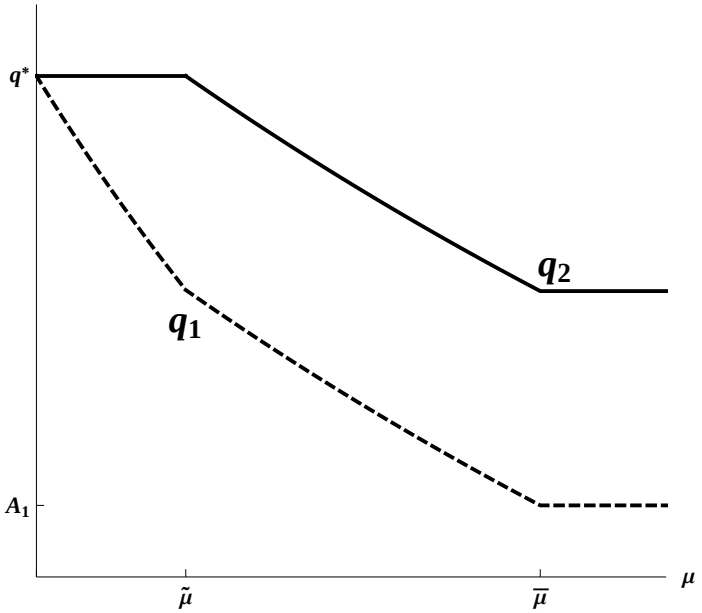

Figure 6: Equilibrium LW quantities as functions of inflation.

The results demonstrated in Proposition 3, and illustrated in Figure 6, are also very intuitive. Agents who did not match in the OTC have to rely exclusively on their own real balances. Hence, q_1 will always coincide with Z, and it will be a decreasing function of μ, for $\mu < \bar{\mu}$. The equilibrium quantity q_2 represents the amount of good that the buyer can afford to purchase in the LW market, when she has previously traded in the OTC market. Hence, whenever equilibrium lies in the Regions MA or NA, we have $q_2 = q^*$. In contrast, if equilibrium lies in the regions of scarcity of A_2 in OTC trade (Regions MS or NS), the buyer will not be able to afford the first-best, and $q_2 < q^*$. In this case, q_2 is a decreasing function of (not affected by) μ if and only if equilibrium is monetary (non-monetary).

3.5 Empirically Relevant Predictions of the Model

In Proposition 2, we described the equilibrium prices for the model with two maturities. In this section, we highlight and discuss three predictions of the model, which are consistent with empirical observations.

3.5.1 The Term Premium of Long Term Assets

In order to discuss the predictions of the model for the term structure of interest rates, we first define the term premium between long term and short term assets simply in terms of the return differential:[16]

$$\tau_{1,2} \equiv r_2 - r_1,$$

where r_i, $i = 1, 2$, was defined by equation (1). Note that, given the restriction to steady-state equilibria (so that r_i is constant), in the absence of any liquidity considerations, the expectations hypothesis would imply a flat yield curve. However, the model predicts the existence of a positively-sloped yield curve, which is a well-established feature of bond yields.

Result 1. *Assume that $A_1 < q^*$ and that the conditions for existence of equilibrium (Lemma 4) are satisfied. Then the term premium is strictly positive; formally, $\tau_{1,2} > 0$.*

Result 1 reveals that our model delivers a positive term premium, as long as the supply of long term maturities is relatively scarce.[17] To see why this result is true, consider first the case in which long term assets are relatively plentiful (Regions MA or NA). Here, we have $\psi_2 = \beta\psi_1$, and $\tau_{1,2} > 0$ requires

$$r_2 > r_1 \Leftrightarrow \left(\frac{1}{\beta\psi_1}\right)^{\frac{1}{2}} > \frac{1}{\psi_1},$$

which is always true, since $A_1 < q^*$ implies $\psi_1 > \beta$ (Proposition 2). If equilibrium lies in the regions of scarcity of long term assets (Regions MS or NS), we have $\psi_2 = \beta\rho\psi_1$, where $\rho \in (1, (1+\mu)/\beta)$ represents the indirect liquidity premium. In this case, $\tau_{1,2} > 0$ requires

$$r_2 > r_1 \Leftrightarrow \left(\frac{1}{\beta\rho\psi_1}\right)^{\frac{1}{2}} > \frac{1}{\psi_1} \Leftrightarrow \psi_1 > \beta\rho.$$

[16] In models with aggregate uncertainty (i.e. stochastic interest rates), the term premium is often defined as $\tilde{\tau}_{1,2} = \mathbb{E}_t[(1+r_{2,t})^2/(1+r_{1,t+1})] - (1+r_{1,t})$. That is, the term premium is the excess return from selling a two-period bond after one period. With no uncertainty and constant interest rates it is easy to verify that $\tau_{1,2} > 0$ if and only if $\tilde{\tau}_{1,2} > 0$, so that the qualitative behavior of our measure is the same. Also, notice that, given our definition of $\tau_{1,2}$, a positive term premium is equivalent to an upward sloping yield curve.

[17] If $A_1 \geq q^*$, we know from Proposition 1 that assets will always be priced at their fundamental value, i.e. $\psi_i = \beta^i$, for $i = 1, 2$. This, in turn, implies (by (1)) that $r_i = 1/\beta - 1$, for $i = 1, 2$, so that $\tau_{1,2} = 0$.

Consider for example an equilibrium in Region MS (the argument for Region NS is similar). We have $\psi = 1 + \mu$, but we also know that $\rho < (1 + \mu)/\beta$. Thus, it is clear that $\psi_1 > \beta\rho$, and we conclude that $\tau_{1,2} > 0$.

The analysis above indicates that the positive term premium is crucially linked to the existence of liquidity premia in asset prices. Short term assets are a close substitute to money, and, under the assumption that $A_1 < q^*$ (and $\mu > \beta - 1$), they always carry a liquidity premium that reflects the ability of the marginal unit to increase LW consumption. On the other hand, long term assets cannot substitute money in the LW market, so that agents who hold long maturities must be compensated for their relatively illiquid in the form of a positive term premium.

The term premium $\tau_{1,2}$ is positive even if the long term assets carry indirect liquidity properties (Regions MS or NS) due to their ability to help agents acquire liquid assets in the OTC market. However, if ρ is close to its upper bound, i.e. $\rho \approx (1 + \mu)/\beta$, we have $\tau_{1,2} \approx 0$. Recalling the definition of ρ (equation (32)), and noticing that this expression coincides with the first two terms on the right-hand side of (30), implies that there are two ways to obtain a zero term premium: a) if $Z = q^*$ (which is not possible under the assumptions $A_1 < q^*$ and $\mu > \beta - 1$), or b) if the multiplier of the term $u'(Z) - 1$ in (30) equals zero. This multiplier is given by $\ell - \lambda f[(1 - \lambda)u'(Z + \zeta^M) + \lambda]^{-1}$, and it will equal zero only if $\ell = f$ (C-types match with probability 1) *and* $\lambda = 1$ (C-types have all the bargaining power). This result is very intuitive. A buyer will be willing to hold long term assets at yield $r_2 = r_1$ only if they are as liquid as short term assets, and this will be true only if the C-type (the type of agent who needs liquidity) is guaranteed to match in the OTC market and is able to extract the whole surplus of that match.

Finally, it is interesting to study the effect of inflation on the term premium. Hence, focus on the regions of monetary equilibrium, and, for simplicity, consider the case of plentiful A_2, i.e. Region MA (the argument for Region MS is slightly more complicated since it involves the derivative of ρ with respect to μ). In this region, we know that

$$\tau_{1,2} = r_2 - r_1 = \left[\frac{1}{\beta(1+\mu)}\right]^{\frac{1}{2}} - \frac{1}{1+\mu}.$$

First, as $\mu \to \beta - 1$, we have $Z \to q^*$, and, consistent with the discussion above, $\tau_{1,2} \to 0$. Moreover, one can easily verify that

$$\frac{\partial \tau_{1,2}}{\partial \mu} = \frac{1}{(1+\mu)^{\frac{3}{2}}}\left[\frac{1}{(1+\mu)^{\frac{1}{2}}} - \frac{1}{2\beta^{\frac{1}{2}}}\right],$$

which is positive if and only if $\mu < 4\beta - 1$. For reasonable (not too small) values of β, we have $\bar{\mu} < 4\beta - 1$. Thus, in monetary equilibrium, $\tau_{1,2}$ is increasing in μ. This result is quite intuitive. Inflation increases the prices of (and reduces the interest rates on) both types of assets, which means that the sign of $\partial\tau_{1,2}/\partial\mu$ might be ambiguous. However, the effect of inflation on short

term assets (which are closer substitutes to money) is stronger, implying that $\partial\tau_{1,2}/\partial\mu > 0$.

3.5.2 The Effect of Secondary Market Liquidity on Asset Returns

One of the key insights of our model is that the issue price of long maturity assets is crucially affected by the liquidity of the secondary asset market, i.e. how easy it is for agents to liquidate these long maturity assets. To highlight the importance of this liquidity mechanism for equilibrium asset returns, we conduct the following experiment: we extend the baseline model (with $N = 2$) to include a second set of assets whose only difference from the original assets studied in previous sections is that they *cannot be traded* in secondary markets (the new assets are present *only* in Section 3.5.2). In any other aspect, the new assets are identical to the original ones. The new assets come in fixed supplies denoted by B_1, B_2, with $B_i = A_i$, $i = 1, 2$, and each unit of asset of maturity $i = 1, 2$ purchased in period t delivers one unit of (the same) fruit before the LW market of period $t+i$ opens. Agents can purchase new assets of maturity $i = 1, 2$ at the ongoing market price p_i (in the CM). However, once an agent buys some claims to the new trees, she has to hold them to maturity. For convenience, we will refer to the newly introduced assets as type-B assets, and we will let type-A assets denote the original ones.

We show that, as long as the supply of short term assets is not so large as to satisfy the liquidity needs of the economy, the issue price of long maturities will be higher for the assets that can be traded in secondary markets, thus reflecting a liquidity premium. Subsequently, we provide some evidence from recent empirical finance literature in support of this result.

Result 2. *Suppose that $A_1 + B_1 < q^*$, and the parameters are such that equilibrium lies in Regions MS or NS. Then, $\psi_1 = p_1$, $\psi_2 = \beta\rho(\mu, A_2)\psi_1$, $p_2 = \beta p_1$, and $\rho(\mu, A_2) > 1$, so that $\psi_2 > p_2$. Moreover, the indirect liquidity premium $\rho(\mu, A_2)$ is decreasing in A_2.*

Proof. See the appendix. □

Result 2 is straightforward. Since short term claims to both types of trees are perfect substitutes to money (and to each other), the existence of "interesting equilibria" (i.e. equilibria with liquidity premia) requires $A_1 + B_1 < q^*$. If this condition is satisfied, the prices ψ_1 and p_1 will include a liquidity premium and will be equal.[18] The price p_2 will include a liquidity premium, but only to the extend that long term assets (of type-B) will become short term assets in the following period, i.e. $p_2 = \beta p_1$. On the other hand, ψ_2 can include an additional indirect liquidity premium, indicated by $\rho(\mu, A_2)$, which reflects the assets' property to help agents avoid the cost of holding liquid assets. Thus, if A_2 is relatively scarce, we have $\psi_2 > p_2$. In terms of

[18] More precisely, one can show that $\psi_1 = p_1 = \min\{1 + \mu, 1 + \bar{\mu}\}$, where $\bar{\mu}$ is the upper bound of admissible monetary policies, and it is analogous to the term $\bar{\mu}(A_1, A_2)$ defined in Section 3.4.

asset yields (rather than prices), letting r_2^B denote the interest rate on long term assets of type-B, Result 2 indicates that $r_2^B > r_2$. Finally, restricting attention to values of A_1, B_1, A_2 such that the equilibrium lies in Regions MS or NS, the difference $\psi_2 - p_2$ is decreasing in A_2. This is true because the service that long term assets of type-A provide (helping agents avoid the holding cost of liquid assets) becomes more valuable when A_2 is more scarce (see also Proposition 2).

Krishnamurthy and Vissing-Jorgensen (2012) provide direct evidence in support of these findings. The authors compare the yields on 6-month FDIC-insured certificates of deposit (CDs) and 6-month treasury bills over the 1984-2008 period. Both assets are default-free, but unlike T-bills, CDs have to be held to maturity. Consequently, the authors argue, any spread reflects the higher liquidity of T-bills. They report that, over the sample period, the spread was 2.3 percentage points on average, and was negatively related to the supply of T-bills. As Result 2 reveals, the model is consistent with both of these findings.

More generally, our model predicts that, *ceteris paribus*, equilibrium prices (yields) are increasing (decreasing) in the ease with which agents can trade assets in the secondary OTC market (a formal way to state this result is that $\partial\psi_2/\partial f > 0$ or $\partial r_2/\partial f < 0$, either of which follows immediately from Proposition 2). This finding is consistent with Gürkaynak, Sack, and Wright's (2010) analysis of the yield curve for inflation-indexed Treasury debt (i.e. TIPS). In particular, the authors demonstrate that, over the period from 1999 to 2005, the TIPS yields have, in general, fallen as market liquidity (measured by trading volume) in the TIPS market has increased.

3.5.3 The On-the-run Phenomenon

One interesting feature of our model is that N-type agents who, in the OTC market of period t, purchase assets issued at $t-1$ and maturing at $t+1$, could also obtain identical assets (maturing at $t+1$) in the forthcoming CM (of period t). Therefore, our model provides a framework in which one can compare the price of on-the-run short term assets with the price of older assets (off-the-run) which mature on the same date. Warga (1992) documents that the return of an off-the-run portfolio exceeds, on average, the return of an on-the-run portfolio with similar duration. Our model is consistent with this observation.

Result 3. *Assume that $A_1 < q^*$ and that the conditions for existence of equilibrium (Lemma 4) are satisfied. Define $\psi_o \equiv \zeta/\chi$ to be the (real) price of long term bonds (issued in the previous period and maturing in the next one) in the OTC market. Comparing ψ_o with the issue price of assets that mature in the next period, ψ_1, we obtain:*

$$\psi_1 = \psi_o \left[(1-\lambda) \frac{u(Z+\zeta) - u(Z)}{\zeta} + \lambda \right]. \tag{33}$$

In any equilibrium, $\psi_1 > \psi_o$.

Proof. See the appendix. □

The assets that are sold by N-types in period t's OTC market (issued at $t-1$ and maturing at $t+1$) have the same maturity structure as the short term assets issued in period t's CM. Hence, one might expect that their prices should be equal. This argument fails to recognize two important facts. First, the seller of off-the-run assets is not the same agent as the seller (issuer) of on-the-run assets.[19] Second, the very structure of the markets in which these two types of assets are traded is different. With respect to the first point, a seller of off-the-run assets is an agent who received a consumption opportunity (a C-type) and who, typically, is short of liquidity. This agent will be desperate for the N-type's liquidity and more willing to sell assets at a low price. Moreover, ψ_o is determined in an OTC market characterized by search and bargaining. Hence, while ψ_1 reflects the fundamental properties of short term assets (the marginal benefit of holding one extra unit), ψ_o represents the terms of trade that implement the "correct" sharing rule of the surplus generated during OTC trade. As long as the N-type has some bargaining power ($\lambda < 1$), she will always extract a fraction of the surplus and purchase assets at price $\psi_o < \psi_1$. This point becomes clear by noticing that $\psi_o = \psi_1$ only if $\lambda = 1$ (equation (33)).

Vayanos and Weill (2008) also provide a theoretical explanation of the so-called on-the-run phenomenon. They build a model where on the on-the-run bonds are more valuable because they are more liquid than their off-the-run counterparts and because they constitute better collateral for borrowing in the repo market (a phenomenon known as "specialness").[20] Importantly, in their model, both of these advantages of on-the-run assets arise endogenously and simultaneously. In this paper, we highlight another possible explanation for the on-the-run phenomenon: the fact that sellers of off-the-run bonds could be more desperate for liquidity (i.e. more eager to sell). However, both models share a common, and very crucial, assumption: asset trade is characterized by search and bargaining frictions. This assumption is essential for the models' ability to capture the on-the-run phenomenon.

4 Equilibrium in the Model with N Maturities

Extending the baseline model to include longer-term assets is straightforward, as long as agents are always able to obtain the representative portfolio in the CM without selling off-the-run

[19] In fact, here we remain agnostic as to who is the issuer of these assets by treating them as "Lucas trees".

[20] It is important to highlight that Vayanos and Weill (2008) define liquidity in a slightly different way than we do. In that paper, liquidity is defined as the ease with which agents can find buyers for their assets. Here, assets are liquid primarily because they can help agents facilitate trade in the anonymous LW market (money and short term assets). However, long term assets also have indirect liquidity properties, since they can help agents acquire liquid assets in the OTC market. In fact, the latter notion of liquidity (that of long term assets) is quite close to the one employed by Vayanos and Weill: how liquid long term assets will be, is in direct relationship with how easily they can be traded (for money and maturing assets) in the OTC market.

assets. A simple sufficient condition would be that $A_1 \geq 2A_2 \geq \ldots \geq 2^{N-1}A_N$.

With $N > 2$, there are many combinations of long term asset portfolios that a C-type can sell in order to obtain additional liquidity in the OTC market. We choose to not place any restrictions on which assets can be traded for liquidity. That is, we assume that in any OTC meeting the C-type can exchange any portfolio of long term assets (assets that do not mature in the current period) for a portfolio of liquid assets (money and assets that do mature in the current period). In that sense, even though $N > 2$, the interesting distinction is still between assets that mature now (and are therefore liquid) and assets that do not mature now (but can be traded for liquid assets in the OTC). Hence, the spirit of the analysis of the previous section does not change.

We now generalize Proposition 2 for N maturities. Recall the definition of $\bar{A}_2$ (equation (25)). For the sake of brevity, we will focus on monetary equilibria. It can be shown, that there exists a $\bar{\mu}$ such that equilibrium will be monetary if and only if $\mu < \bar{\mu}$.

Proposition 4. *Assume that $\mu < \bar{\mu}$ and $2^{N-1}A_N \leq \ldots \leq 2A_2 \leq A_1 < q^*$. Define the supply of long term assets relevant for abundance vs scarcity in OTC trade as*

$$A_L \equiv \beta^{N-2}A_N + \ldots + A_2.$$

Then the equilibrium price of one-period assets is given by $\psi_1 = 1 + \mu$, and the equilibrium price of long term assets (i.e. ψ_i, $i \geq 2$) depends on the value of A_L. We have two cases:

Case 1: *If $A_L \geq \bar{A}_2(A_1)$, then $\psi_i = \beta^{i-1}\psi_1$.*

Case 2: *If $A_L < \bar{A}_2(A_1)$, then there exists a cutoff $\tilde{\mu}(A_L)$ such that:*
a) For all $\mu \in (\beta - 1, \tilde{\mu}(A_L)]$, we have $\psi_i = \beta^{i-1}\psi_1$;
b) For all $\mu \in (\tilde{\mu}(A_L), \bar{\mu})$, we have $\psi_i = \left(\beta\rho^L\right)^{i-1}\psi_1$, where $\rho^L \in (1, (1+\mu)/\beta)$ is a strictly increasing function of μ and a strictly decreasing function of A_i for all $i \geq 2$.

The term ρ^L is defined jointly with ζ^L (the real balance trading volume in the OTC market) as a function of equilibrium real balances Z:

$$\rho^L = 1 + \lambda f \frac{u'(Z + \zeta^L) - 1}{(1-\lambda)u'(Z + \zeta^L) + \lambda}$$

$$(1+\mu)\beta \sum_{i=2}^{N} \left(\beta\rho^L\right)^{i-2} A_i = (1-\lambda)\left[u(Z + \zeta^L) - u(Z)\right] + \lambda\zeta^L$$

Proof. See the appendix. □

Proposition 4 reveals that the results in the case of a general $N > 2$ are qualitatively very similar to the ones in the $N = 2$ case. In particular, one-period assets are "in a class of their

own", since they are the only assets that are (direct) substitutes to money. Hence, in equilibrium, we obtain $\psi_1 = 1 + \mu$, just like in Section 3. The price of longer term assets, ψ_i, $i \geq 2$, always carries a liquidity premium because these assets will eventually also become short term assets in future periods, i.e. $\psi_i > \beta^i$ in any equilibrium. Moreover, if the supply of longer term assets is relatively scarce (Case 2-b of the proposition), the price ψ_i, $i \geq 2$, will also contain an indirect liquidity premium, ρ, which reflects the assets' property to be traded for liquid assets in the OTC market. As in Section 3, the premium ρ is increasing in inflation (in monetary equilibrium) and decreasing in the supply of long term assets (in the regions of "scarcity"), because a high μ or a low A_L makes the service that long term assets provide more valuable.

It is srtaightforward to check that a positively sloped yield curve will also arise here regardless of the region of equilibrium. Consider for instance a relatively abundant supply A_L (the argument for the case of scarce supply is similar). In this case, we have $\psi_1 = 1 + \mu$ and $\psi_2 = \beta(1 + \mu)$ (like in Section 3), and we have already shown that $r_2 > r_1$. Thus, focus on $i \in \{2, ..., N-1\}$, and consider the term $r_{i+1} - r_i$. It can be easily verified that

$$r_{i+1} > r_i \Leftrightarrow \left[\frac{1}{\beta^i(1+\mu)}\right]^{\frac{1}{i+1}} > \left[\frac{1}{\beta^{i-1}(1+\mu)}\right]^{\frac{1}{i}} \Leftrightarrow \left(\frac{1}{\beta}\right)^{\frac{1}{i(i+1)}} > \left(\frac{1}{1+\mu}\right)^{\frac{1}{i(i+1)}},$$

which will be always satisfied, since in any monetary equilibrium $\mu > \beta - 1$.[21]

We conclude that the model with $N > 2$ maturities delivers an upward sloping yield curve throughout the domain $i = 1, ..., N$. It is important to highlight that this result emerges even though any two assets with life-time $i, j \geq 2$ are qualitatively similar, in that neither of them can serve as a direct substitute to money- a property that characterizes only one-period assets. Nevertheless, assets with maturity $i \geq 2$ are, in a sense, still more liquid than assets with maturity $i + 1$ (hence, $r_{i+1} > r_i$, for all $i \geq 2$) because the former will become one-period assets (and perfect substitutes to money) *earlier* than the latter.

5 Conclusions

Liquidity preference is often proposed as a resolution to the well-documented empirical failures of the expectations hypothesis of the term structure. This paper provides a theoretical basis for this preference. We demonstrate that a positive liquidity premium emerges because of three key features: a) agents are subject to stochastic consumption expenditures; b) the markets in

[21] Alternatively, this result can be characterized in terms of positive term premia. For instance, consider the term premium defined by the excess return from selling a 3-period bond after one period. This is given by $\tilde{\tau}_{2,3} = \psi_2/\psi_3 - 1/\psi_1$. Substituting the equilibrium prices from Proposition 4 (again focus on case 2-a; the argument for case 2-b is similar), we obtain $\tilde{\tau}_{2,3} = 1/\beta - 1/(1+\mu)$, which is positive in any monetary equilibrium ($\mu > \beta - 1$). It is also straightforward to establish that term premia defined in terms of the return differential of buying a k-period bond ($k \geq 2$) and holding it to maturity relative to buying a sequence of k one-period bonds are also positive.

which these expenditures take place are decentralized so that a medium of exchange (i.e. a liquid asset) is typically required; and c) secondary assets markets, that is, markets in which agents can sell assets to acquire liquidity, are characterized by search and bargaining frictions. In addition to implying the existence of positive term premia and a positively sloped yield curve, these features also help to explain the on-the-run bond pricing phenomena as well as the role that asset supply has on yields.

References

BACKUS, D. K., A. W. GREGORY, AND S. E. ZIN (1989): "Risk premiums in the term structure: Evidence from artificial economies," *Journal of Monetary Economics*, 24(3), 371–399.

BANSAL, R., AND A. YARON (2004): "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 59(4), 1481–1509.

BERENTSEN, A., G. CAMERA, AND C. WALLER (2007): "Money, credit and banking," *Journal of Economic Theory*, 135(1), 171–195.

BERENTSEN, A., S. HUBER, AND A. MARCHESIANI (2011): "Free-riding on Liquidity," *University of Zurich Department of Economics Working Paper*, (32).

BOEL, P., AND G. CAMERA (2006): "Efficient monetary allocations and the illiquidity of bonds," *Journal of Monetary Economics*, 53(7), 1693–1715.

DUFFIE, D., N. GÂRLEANU, AND L. H. PEDERSEN (2005): "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

GARBADE, K., AND J. INGBER (2005): "The Treasury auction process: Objectives, structure, and recent adaptations," *Current issues in economics and finance*, 11(2).

GEROMICHALOS, A., J. M. LICARI, AND J. SUAREZ-LLEDO (2007): "Monetary Policy and Asset Prices," *Review of Economic Dynamics*, 10(4), 761–779.

GEROMICHALOS, A., AND I. SIMONOVSKA (2011): "Asset Liquidity and International Portfolio Choice," Discussion paper, National Bureau of Economic Research.

GÜRKAYNAK, R. S., B. SACK, AND J. H. WRIGHT (2010): "The tips yield curve and inflation compensation," *American Economic Journal: Macroeconomics*, pp. 70–92.

GÜRKAYNAK, R. S., AND J. H. WRIGHT (2012): "Macroeconomics and the term structure," *Journal of Economic Literature*, 50(2), 331–367.

He, Z., and K. Milbradt (2012): "Endogenous liquidity and defaultable bonds," Discussion paper, National Bureau of Economic Research.

Jacquet, N. L., and S. Tan (2012): "Money and asset prices with uninsurable risks," *Journal of Monetary Economics*.

Kalai, E. (1977): "Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons," *Econometrica*, 45(7), 1623–30.

Krishnamurthy, A., and A. Vissing-Jorgensen (2012): "The aggregate demand for treasury debt," *Journal of Political Economy*, 120(2), 233–267.

Lagos, R. (2010): "Asset prices and liquidity in an exchange economy," *Journal of Monetary Economics*, 57(8), 913–930.

——— (2011): "Asset Prices, Liquidity, and Monetary Policy in an Exchange Economy," *Journal of Money, Credit and Banking*, 43, 521–552.

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

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

Lester, B., A. Postlewaite, and R. Wright (2012): "Information, Liquidity, Asset Prices, and Monetary Policy," *Review of Economic Studies*, forthcoming.

Lucas, Robert E, J. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46(6), 1429–45.

Mishkin, F. S. (2007): *The economics of money, banking, and financial markets*. Pearson education.

Piazzesi, M. (2010): "Affine term structure models," *Handbook of financial econometrics*, 1, 691–766.

Piazzesi, M., and M. Schneider (2007): "Equilibrium yield curves," in *NBER Macroeconomics Annual 2006, Volume 21*, pp. 389–472. MIT Press.

Rocheteau, G. (2011): "Payments and liquidity under adverse selection," *Journal of Monetary Economics*.

Rocheteau, G., and R. Wright (2005): "Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium," *Econometrica*, 73(1), 175–202.

SALYER, K. D. (1990): "The term structure and time series properties of nominal interest rates: Implications from theory," *Journal of Money, Credit and Banking*, 22(4), 478–490.

SINGLETON, K. J. (2009): *Empirical dynamic asset pricing: model specification and econometric assessment*. Princeton University Press.

VAYANOS, D., AND J.-L. VILA (2009): "A preferred-habitat model of the term structure of interest rates," Discussion paper, National Bureau of Economic Research.

VAYANOS, D., AND P.-O. WEILL (2008): "A Search-Based Theory of the On-the-Run Phenomenon," *The Journal of Finance*, 63(3), 1361–1398.

WARGA, A. (1992): "Bond returns, liquidity, and missing data," *Journal of Financial and Quantitative Analysis*, 27(4).

A Appendix

TO BE COMPLETED