

Reconciling Orthodox and Heterodox Views On Money and Banking

DAVID ANDOLFATTO

*Federal Reserve Bank of St. Louis and Simon Fraser University**

A wide range of heterodox theories claim that banks are special because they create money in the act of lending. Put another way, banks can create the funding they need ex nihilo, whereas all other agencies must first acquire the funding they need from other parties. Mainstream economic theory largely agrees with this assessment, but questions its theoretical and empirical relevance, preferring to view banks as one of many potentially important actors in the financial market. In this paper, I develop a formal economic model in an attempt to make these ideas precise. The model lends some support to both views on banking.

Keywords: loanable funds, endogenous money, banking.

JEL Classifications: E4, E5

1 Introduction

While mainstream macroeconomic models today are more attentive to the workings of the financial market, they mostly ignore the fact that bank deposit liabilities constitute money and that the broad money supply is largely determined by the lending practices of the bank sector. Given that these mainstream models purportedly help inform central bank policy, and given the prominent role that banks often play in financial crises, is this omission a mistake? A chorus of heterodox voices in macroeconomics are adamant that it is. This paper constitutes an attempt to evaluate their claim.

Let me review briefly, to the best of my ability, the key elements of two broad schools of thought on the matter. The mainstream view is that modeling the operational details of banking—the financing of assets with deposit liabilities—can be safely ignored for questions related to the business cycle. According to this view, the degree of liquidity associated with liabilities issued by the private sector lies on a continuum, so it makes little sense to draw a

* The paper is based on a set of lectures I delivered to my Monetary Policy and Theory class in the summer of 2018 at Simon Fraser University. I want to thank my students first, for paying attention through that long hot summer and second, for their comments and questions which led me to sharpen my thinking on the subject.

© 2018 David Andolfatto. Licensed under the Creative Commons Attribution - Noncommercial 3.0 Licence (<http://creativecommons.org/licenses/by-nc/3.0/>). Available at <http://rofea.org>.

sharp distinction between which of these liabilities constitute “money” and which do not (Tobin, 1963). Banks are just one of many actors in the market for loanable funds. Admittedly, they may play an important role in the cycle, but only to the extent that they constitute highly-levered entities (Krugman, 2012a, 2012b). The fact that bank lending creates money—i.e., that bank liabilities are more liquid than those of other financial agencies—is not a critical consideration. In any case, central bank policy, whether through reserve management or interest rate policy, ultimately controls the total (central and commercial) money supply (Rowe, 2012a).

The heterodox view is that banks are critically different from other financial market participants. In particular, while non-bank agencies wanting to finance investments first need to acquire the requisite funding, this is not the case for banks. In contrast to the non-bank sector, banks can *create* the money they lend. Consolidating the banking sector with other actors in a market for loanable funds stems from the misguided notion that banks lend out reserves that they must first acquire. Banks do not in fact lend reserves—they lend their deposit liabilities (which are incidentally made redeemable for cash). While banks make use of reserves for settlement purposes, reserves do not constrain bank lending (Fullwiler, 2012).¹ First, central banks typically supply the reserves demanded by the banking sector. Second, even with a limited supply of reserves, the banking sector possesses a wide latitude in determining the money supply. This is important because it implies that the macroeconomic impact of bank-financed investment is quantitatively more significant than market-financed investment. A hedge fund, for example, must first acquire the money it needs to fund an investment. The money it acquires reduces the spending power of those who have released it.² Because a bank does not require an outside source of funds, bank-financed investments have a greater impact on aggregate demand. Autonomous changes in bank-lending standards can therefore have an outsized effect on aggregate demand (Mason, 2012). A failure to understand the special nature of banking can lead to ill-designed financial market regulations that promote undiagnosed/undesirable credit market booms and crashes (Pettifor, 2014).

Sorting through the substantive differences in these two views and why it matters is not an easy task. You can judge for yourself by wading through a debate on the subject that surfaced not too long ago on the blogosphere when the mainstream Krugman (2012a, 2012b) replied to

¹ Fullwiler (2012) argues that banks are not constrained by reserves when the central bank targets an interest rate. This is true, but it only implies that the central bank is fully accommodating the demand for reserves in this case. Interest rate policy still affects the opportunity cost of lending. See also Rowe (2012a).

² It is generally agreed that the reallocation of purchasing power from savers to borrowers can have consequences for aggregate demand to the extent that savers and borrowers have (say) different propensities to consume. The heterodox view appears to consider such effects of second-order importance.

an opening salvo from the heterodox Keen (2012). The discussion that ensued did not lead to a satisfactory conclusion.³ I am sympathetic to the observation made by Krugman (2012a):

My point is that there seems to be a lot of implicit theorizing going on here — and at least at first glance, the implicit theorizing doesn't make a lot of sense. I could be wrong, but that's the whole point of simple models: to lay bare what you're assuming, and make it clear what, specifically, is driving your conclusions.

The purpose of this paper is make the theory more explicit. The model I develop below delivers conclusions that are consistent with either view depending on the parameter values describing the economic environment. The model therefore provides a way to identify which assumptions are critical to delivering the results favored by proponents of each view.

The framework I propose to use to help us organize our thinking on this matter is a simple overlapping generations model. At any given point in time, the population is divided into two age cohorts: “young” and “old.” There is heterogeneity within age cohorts as well. The young consist of “workers” (savers) and “investors” (borrowers). Naturally, this implies that the old will be characterized by their respective trading histories. Old savers will have accumulated financial wealth. Old investors will have accumulated financial obligations, which they honor from the proceeds of their maturing investments.

The analysis proceeds in stages. To begin, I study the outcome associated with a well-functioning market for private debt. In this economy, young investors borrow young labor to construct new capital goods, which produce a future output that can be divided between investor and creditor. Young labor is paid with private IOUs representing direct claims against future inventory. This action takes place on the proverbial “loanable funds market” that matches savers with borrowers.

Next, I introduce a fixed supply of nominal interest-bearing government debt, where the interest-expense is financed through taxes. I view the debt as being held entirely as an asset in a narrow bank, so that money takes the form of interest-bearing deposit liabilities issued by the banking sector. Workers can now allocate their labor to the services sector (producing consumption for the old in exchange for money) or, as before, to the capital goods sector (producing capital goods for young investors in exchange for private debt). The nominal interest rate is determined by a Taylor rule. A parameter in the Taylor rule dictates the prevailing real rate of interest, which determines investment demand and the demand for real money balances.

I then impose a credit-market friction that limits how much funding investors can obtain from the market for loanable funds. For simplicity, I consider the extreme case where no

³ Though some welcome clarity was provided by Rowe (2012b), Waldman (2012), Koning (2013) and Roche (2013).

credit is granted. In this economy, investment collapses as all savings are diverted to money (government debt). The price-level deflates. The only possible winners in this scenario are those holding the nominal debt during the deflation. To an econometrician, what transpires looks very much like a collapse in aggregate demand. Desired saving in the form of private debt does not match investment demand at the prevailing real rate of interest.

Finally, I permit the banking system to hold assets other than government debt. I make the following critical assumption: the liabilities issued by banks are more credible than those issued by private investors.⁴ With this power, banks can create and lend money to investors in exchange for their illiquid collateral. In the model, banks are unconstrained by reserves in the act of creating money (though the policy interest rate ultimately dictates loan demand and hence the size of the banking sector's balance sheet). The added supply of bank money stimulates real economic activity because it relaxes borrowing constraints. Exogenous variation in the ability/willingness of the banking sector to finance investment generates a business cycle with a procyclical price-level. The price-level rises in a boom not because of an expanding broad money supply, but because additional bank lending stimulates investment, which leads to a decline in the demand for real money balances through a portfolio substitution effect. The long-run rate of inflation is determined by the growth rate of the supply of outside money relative to the growth rate of money demand (in the model, both are normalized to zero).

What to make of all this? In the model, banks and financial markets are competing mechanisms for allocating credit. Banks are "special" only to the extent they are better than markets at funding investment. This specialness is not (in the model) logically rooted in their ability to create money. In particular, bond-finance in the model is "special" if it is the lower cost way to fund investment. Variations in the parameter that governs the willingness/ability of non-bank creditors to extend credit generates business cycles in the exact same way it would in a banking economy.

Because bank-finance coexists with other forms of financing, it is not entirely clear why monetary/fiscal policies should make special consideration of the banking sector *per se*.⁵ The fact that commercial banks create liabilities that are used widely in payments is not likely a first-order concern. We saw in the financial crisis of 2008, in particular, how highly-leveraged shadow banks were perfectly capable of disrupting the financial market without the capacity to create money.⁶ A case can be made that the ingredient missing from conventional

⁴ In general, we can think of this being true for a subset of the population and the opposite being true for others, implying the coexistence of bank deposits with other forms of financing.

⁵ The matter is different for policymakers in charge of regulatory design since in this case institutional details clearly matter.

⁶ In the context of the present discussion it is somewhat ironic to note that these non-bank entities behaved in the way traditional textbooks describe conventional banks as lending out reserves. Instead

macroeconomic models was leverage in general, not banking in particular. This is the sense in which the analysis below, while sympathetic to the heterodox perspective, largely ends up supporting the mainstream view.

2 The environment

Let us consider an overlapping generations model.⁷ At each date $t = 1, 2, \dots, \infty$, $2N$ young individuals enter the economy. The young live (have a planning horizon over) two periods. At the end of the second period of their planning horizon, the young exit the economy. At date $t = 1$, there are $2N$ old individuals (the initial old) who are the only ones in the economy with a planning horizon of one period.

The population of young individuals is divided evenly between *workers* and *investors*, so that there is a mass N of each type. In what follows, I make the harmless normalization $N = 1$. Workers are endowed with y_t units of output when young that can take the form of consumption or investment.⁸ Investors are endowed with an investment project that transforms k units of output at date t into $xf(k)$ units of output at date $t + 1$. Assume that $f'' < 0 < f'$ with $f(0) = 0$ and the standard Inada conditions. The parameter x indexes the expected productivity of capital spending.⁹ Assume $f'(k)k = \alpha f(k)$, where $0 < \alpha < 1$ indexes the importance of capital in production (*vis-à-vis* the implicit fixed factor).¹⁰

An individual who enters the economy at date $t = 0, 1, 2, \dots, \infty$ has preferences given by $U_t = c_{t+1}$. That is, all individuals are assumed to care only for consumption in the second period of life. Moreover, individuals are risk-neutral. Note that at any given point in time $t > 1$, the economy is populated by four types of individuals: young workers, young investors, old workers, and old investors—this is not a representative agent economy!

of reserves, entities like Lehman Brothers lent out (rehypothecated) U.S. Treasuries to finance higher yielding tranches of purportedly safe mortgage-backed securities.

⁷ I hope that readers not familiar with the OLG model will not be put off too much by the level of abstraction. Where certain assumptions are deemed excessively unpalatable, keep in mind that we can return later and add the pieces we feel are missing and check the robustness of our conclusions.

⁸ Alternatively, think of the young as being endowed with a unit of labor that can be transformed into y_t units of output. Labor employed in the consumption sector produces consumer goods. Labor employed in the investment sector produces new capital goods.

⁹ Nothing in the analysis below requires expectations to be “rational.” All that is important here is that individuals are forward-looking.

¹⁰ That is, each investor is assumed to possess e units of “entrepreneurial effort” and an investment technology $F(k, e)$ where F is standard neoclassical. Normalize the fixed factor $e = 1$ and define $f(k) \equiv F(k, 1)$.

2.1 A market for loanable funds

The gains to trade in this economy are self-evident: young investors want young labor in exchange for future consumption. There are a variety of ways in which we can think of these individuals trading. Here, I choose to model exchange as occurring in a centralized location where all are free to bid for the goods and services they desire and where the terms of trade are ultimately determined by “supply and demand” in a competitive market.¹¹ One could alternatively have modeled exchange as a decentralized market where savers and borrowers meet bilaterally or through a broker; see, for example, Lagos, Rocheteau and Weill (2009).

The economy can be thought of working as follows. Young investors need construction labor (supplied by young workers) to build the capital goods k_t that will generate the future expected return $x_{t+1}f(k_t)$. How do investors pay for the resources they need? They need to borrow them. One way to borrow labor is to pay for it using a privately-issued IOU entitling the bearer to a specific quantity of future output. The type of debt instrument I have in mind here resembles the store coupons issued by *Canadian Tire* (a popular Canadian retailer); see Figure 1. The note in Figure 1 promises that it can be redeemed for merchandise at any *Canadian Tire* store.

Imagine that our young investors operate like *Canadian Tire* and that our young workers get paid in *Canadian Tire* coupons. Each coupon promises the bearer r_t units of output (store merchandise) in the following period. Then, if workers sell k_t^s units of their labor for $k_t^s \leq y_t$ coupons, they get to consume $c_{t+1}^w = r_t k_t^s$ in period $t + 1$. Given that workers here have no other use for their labor, their optimal choice is to set $k_t^s = y_t$.

Figure 1: Canadian Tire Coupon



¹¹ A competitive market simply means that no individual can affect the market-determined terms of trade. Individuals are of course free to sell their goods at above or below market prices, but doing so would be pointless. In the former case, they'd make no sales. In the latter case, they'd sell just as much at higher prices by selling at the prevailing market price.

Now consider the problem faced by young investors. Given the market rate of return on coupons r_t , an investor's expected total return is equal to $x_{t+1}f(k_t) - r_t k_t$, where $r_t k_t$ is effectively the principal and interest due on a loan of size k_t . The profit-maximizing level of investment is given by a level of investment demand k_t^d that satisfies the condition

$$x_{t+1}f'(k_t^d) = r_t \quad (1)$$

The left-hand-side of (1) represents the expected marginal return to capital and the right-hand-side represents the marginal cost of borrowing. Condition (1) implicitly defines an investment demand function $k_t^d = k^d(r_t, x_{t+1})$ that is decreasing in r_t and increasing in x_{t+1} (the parameter indexing investor sentiment over the prospective return to investment).

Consistency of economic decisions in the aggregate translates here into the equation of supply and demand $k_t^s = k_t^d$. This market-clearing condition determines the equilibrium real interest rate r_t . Since $k_t^s = y_t = k_t^d$, it follows from condition (1) that the equilibrium real rate of interest is equal to $r_t = x_{t+1}f'(y_t)$. The associated equilibrium consumption allocations are not central to the analysis below, so I relegate their derivation to the appendix.

2.2 Government debt

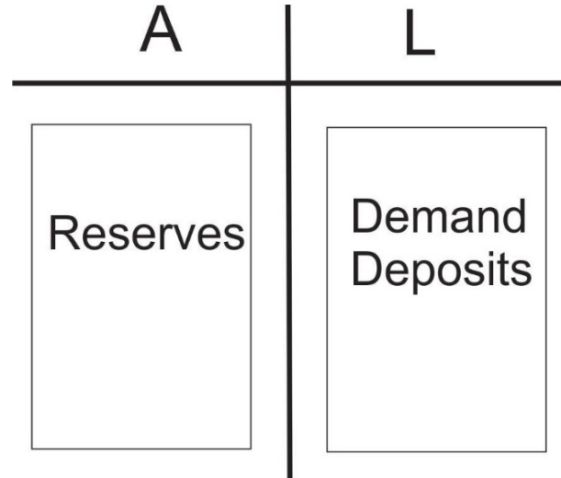
In what follows, I consider a consolidated monetary/fiscal authority and make no distinction between nominal government debt and central bank liabilities. In what follows, I assume a fixed stock of nominal debt M . Assume that M is injected into the economy at date $t = 1$ as a social security payment to the initial old. Because the debt M will end up being used as a means of payment, I will use the terms *government debt* and *government money* interchangeably.

There are several ways to map M into objects we (may potentially see) in reality. First, we could think of M as representing physical currency that passes from hand-to-hand in a peer-to-peer (P2P) fashion. Second, we could think of M as a central bank digital currency (CBDC). Imagine, for example, that individuals are permitted to open money accounts with the government or the central bank.¹² In this case, instead of P2P transactions involving cash, one could make payments from account-to-account by instructing the central bank to make the appropriate debit/credit operation on a central ledger. Third, we could imagine M as private bank digital currency (PBDC). In this configuration, the government delegates the responsibility of processing payments to the private bank sector, but retains control over the total money supply M . This latter institutional arrangement is sometimes referred to as "narrow banking" or "100% reserve banking" since the liabilities issued by private banks

¹² In fact, in the U.S. people can open electronic accounts with the U.S. Treasury; see www.treasurydirect.gov.

must be backed fully with government debt.¹³ Figure 2 depicts the balance sheet of a narrow bank.

Figure 2: Narrow Bank



Let p_t denote the price of output at date t measured in units of government money. If prices are expressed in this form, then we say that government money is a unit of account (UoA). If individuals accept government money for payment, then it is also a medium of exchange (MoE). Note that the UoA and MoE properties are conceptually distinct: one can imagine worlds where government debt is a UoA but not a MoE and vice-versa.¹⁴

I assume that the government chooses the nominal interest rate it is willing to pay on its debt (we can think of this choice as “monetary policy”). Let R_t denote the gross nominal interest rate. I assume further that the interest expense of the debt is financed entirely with tax revenue T_t , so that the government budget constraint in each period is given by

$$T_t = (R_{t-1} - 1)M \tag{2}$$

In general, the interest expense can also be financed in whole or in part by issuing new debt. I make the simplifying and important assumption that taxes are lump-sum. I also make the

¹³ “Backed fully” means that private bank money is convertible one-for-one with government money.

¹⁴ In periods of high inflation, merchants have been known to specify prices in units of a stable foreign currency, while accepting payment in the domestic currency. In the shadow banking sector, private securities are used as exchange media with their values expressed in units of currency.

simplifying (and not so important) assumption that the lump-sum tax falls entirely on old savers.

A young worker wants to sell his goods y_t for purchasing power that he can carry into the future. This purchasing power can take the form of government money, which can be spent on the goods supplied by future young workers y_{t+1} . It can also take the form of direct claims—like the *Canadian Tire* coupons described above—claims against the output produced by old investors $x_{t+1}f(k_t)$.

If a young workers sells q_t units of goods for government money, he earns $m_t = p_t q_t$ dollars. These dollars are debited from the accounts of the buyers and credited to the accounts of the sellers in the narrow banking system I described above. Note that since old workers are in possession of the entire money supply, these latter exchanges involve transactions between young and old workers. If a young worker sells k_t units of goods for coupons, the nominal value of this coupon revenue is $p_t k_t$. Note that $p_t k_t$ is only *measured* in dollars—it does not consist of actual dollars (government money). A young worker faces the budget constraint:

$$p_t q_t + p_t k_t = p_t y_t \quad (3)$$

That is, the worker faces a simple portfolio choice problem: how much of his $p_t y_t$ dollars of savings to hold in the form of $p_t q_t$ dollars of money and $p_t k_t$ dollars of coupons.

If a saver is to willingly hold both money and coupons, the two securities will have to earn the same expected rate of return.¹⁵ This no-arbitrage-condition can be expressed as

$$r_t = R_t(p_t/p_{t+1}) \quad (4)$$

where r_t represents the real rate of return on private debt and $R_t(p_t/p_{t+1})$ represents the inflation-adjusted return on government debt.¹⁶ The nominal rate of return on money is equal to R_t and is set by policy. Selling q_t goods for $p_t q_t$ dollars today becomes $R_t p_t q_t$ dollars tomorrow. Selling k_t goods today for $p_t k_t$ dollars *worth* of revenue must (by arbitrage) become $R_t p_t k_t$ dollars worth of coupons tomorrow. An old worker's budget constraint is therefore given by $p_{t+1} c_{t+1}^w = R_t p_t [q_t + k_t] - T_{t+1} = R_t p_t y_t - T_{t+1}$, or

$$c_{t+1}^w = r_t y_t - \tau_{t+1} \quad (5)$$

where $\tau_t \equiv T_t/p_t$. Because money and coupons earn the same expected rate of return, the composition of a saver's wealth portfolio is indeterminate. The composition of saving

¹⁵ Recall that savers are risk-neutral here.

¹⁶ Condition (4) is also sometimes referred to as the Fisher equation.

between money and coupons at the aggregate level is, however, determined by monetary policy.

Consider now the young investor, who wishes to acquire resources k_t . An investor who acquires one unit of the good today promises to repay r_t units of goods (or the purchasing power equivalent) tomorrow. In effect, the investor is selling $p_t k_t$ dollars worth of coupons to savers, who then expect to sell (or exchange) this (dollar equivalent) value for $R_t p_t k_t / p_{t+1}$ goods tomorrow. Alternatively, it is as if the investor borrows $p_t k_t$ dollars and promises to repay $R_t p_t k_t$ dollars. Note, however, that the investor is not borrowing dollars—he is issuing debt to finance the acquisition of real resources. In any case, the old investor's budget constraint is given by,

$$c_{t+1}^i = x_{t+1} f(k_t) - r_t k_t \quad (6)$$

If the investor is unconstrained in his borrowing, he will choose a level of investment that equates the marginal benefit of investment to its opportunity cost as before, i.e., investment demand k_t^d will satisfy condition (1).

Let me now describe the monetary policy rule that governs the determination of the policy rate R_t . In what follows, I assume that the central bank follows a Taylor rule which threatens to raise the nominal interest rate one-for-one with the expected rate of inflation, i.e.,

$$R_t = R_t^*(p_{t+1}/p_t) \quad (7)$$

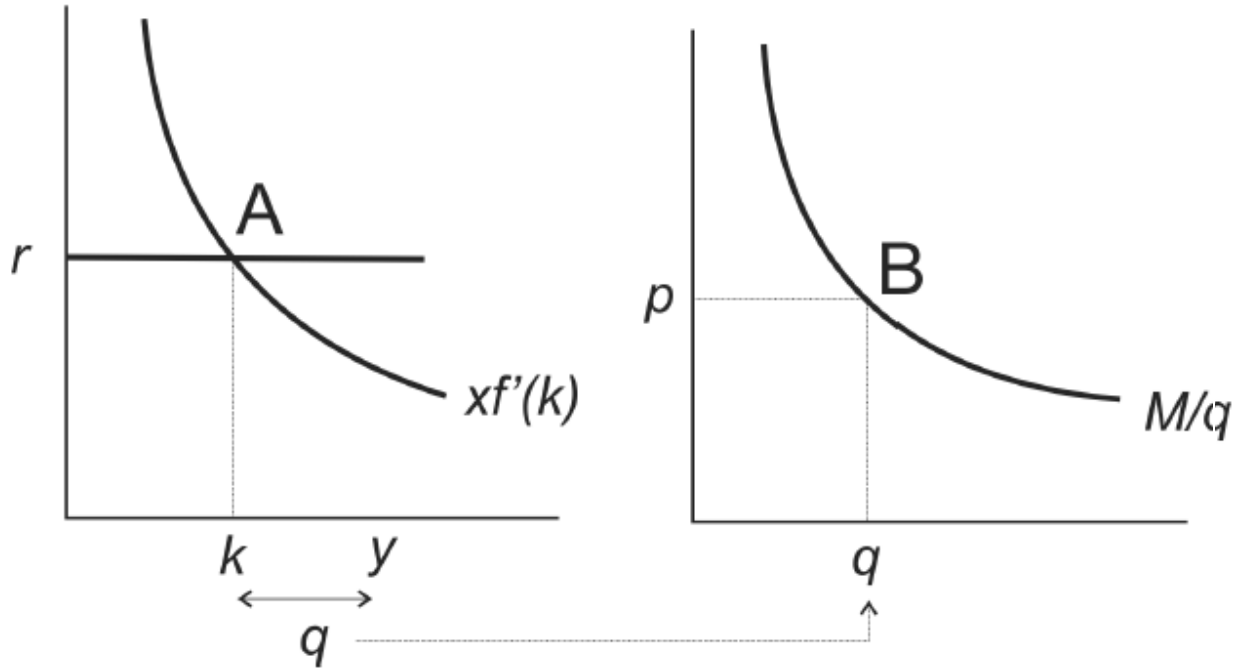
where R_t^* is a policy parameter. I have two reasons for choosing a monetary policy of this form. First, it is empirically realistic: central banks do tend to increase their policy rates in the face of rising inflation expectations. Second, this particular policy rule renders the monetary equilibrium unique and dynamically stable.¹⁷

If we combine the monetary policy rule (7) with the no-arbitrage-condition (4), we derive the result $r_t = R_t^*$. That is, in this (highly stylized) model, the government has direct control over the real rate of interest through its choice of R_t^* . The model has a nice recursive structure (see Figure 3):

1. Monetary policy determines the rule (7) with the choice of R_t^* determining the real rate of interest r_t .
2. Given the interest rate r_t and investor sentiment x_{t+1} , condition (1) determines the level of real capital spending k_t .
3. Given y_t and k_t (the former is exogenous and the latter is determined in step 2), condition (3) determines the demand for real money balances $q_t = y_t - k_t$.

¹⁷ I discuss this in more detail in Andolfatto (2015).

Figure 3: Model Structure



4. With q_t so determined, the equilibrium price-level must satisfy $M = p_t q_t$ (money supply = money demand) or $p_t = M/q_t$.

The equilibrium consumption allocation can be determined by combining the restrictions above with the individual budget constraints. The real GDP at date t can similarly be determined: $RGDP_t = y_t + x_t f(k_{t-1})$.

Let me discuss the determination of the price-level in a little more detail. In this model economy, a young worker's labor produces output that may take the form of either consumer goods or investment goods. Remember that young workers freely accept both government money and private coupons as payment for services rendered. Because the old workers pay in government money (or, recall, with the liabilities of a 100% reserve bank), young workers deliver consumer goods in exchange for government money and deliver investment goods in exchange for private coupons. In the aggregate, the quantity of consumer goods delivered to old workers must equal q_t (step 3 above). Since old workers have a marginal propensity to consume equal to one, they spend their entire financial wealth (both money and coupons). The quantity of money sold is M . The quantity of goods purchased for that money is q_t . The exchange rate (price-level) is therefore given by $p_t = M/q_t$.

As before, I relegate the determination of the consumption allocation to the appendix. The properties of the model resemble those of the textbook model with an IS curve and a policy-

determined real interest rate. While the real GDP in this model is fixed at $y_t + x_t f(k_{t-1})$ as of date t , interest rate policy influences the level of investment k_t and hence the level of future real GDP.¹⁸ The determination of the price-level is very much in the spirit of the *Quantity Theory of Money*. Lump-sum injections of money are generically non-neutral in OLG models because individuals have different propensities to consume. Money is generically not superneutral as well.¹⁹ The appendix contains a more detailed description of the model's properties.

3 Credit constraints and banking

If the financial market that links savers and borrowers worked seamlessly in the manner modeled above, then the services of a financial intermediary (like a bank) would be superfluous. The financial market modeled above works well because I have assumed an absence of financial market frictions, in particular: (i) free communication channels between savers and borrowers; (ii) full commitment on the part of debtors; and (iii) the absence of private information. To the extent that banks are designed to overcome such obstacles, one would expect banking arrangements to emerge to help solve these problems.

The problem from the perspective of young investors is that savers may be sceptical about the promises investors make. This is perhaps not surprising, given that promises (e.g., the coupons above) are made "out of thin air" and hence, exceptionally easy to fabricate. Investors claiming to have profitable investment opportunities may misrepresent prospective returns. They may not honor their promises to repay their creditors in a timely manner, or even at all. One way to model these considerations is to suppose that investors face a credit constraint that limits the amount of debt they are permitted to accumulate to a level less than what they could, in principle, repay. For example, suppose that the interest and principal repayment that savers expect to secure is determined by a "trust" parameter θ_t so that,

$$r_t k_t \leq \theta_t \quad (8)$$

¹⁸ One might object that since contemporaneous output is predetermined, the model presented here does not permit output to be "demand determined." I should like to point out a few things. First, note that there is nothing in my setup which guarantees that the equilibrium is socially optimal. There is a sense in which (say) construction labor can be underemployed if business sentiment is unduly negative, for example. Second, one could, with a little more effort, endogenize the labor market along the lines of Kocherlakota (2012) to model output that is "demand determined." Whether such an extension alters the conclusions that follow remains an open question, but my prior is that it does not.

¹⁹ Money is only superneutral if the interest on money is financed entirely through money creation.

Investment demand is now given by (1) if the debt constraint (8) is slack, or by $k_t = \theta_t/r_t$ if the debt constraint (8) binds.²⁰

Let us suppose that the debt constraint (8) binds. Then the equilibrium level of investment is given by $k_t = \theta_t/r_t$ so that the marginal return to capital spending is higher than its opportunity cost, i.e., $x_{t+1}f'(\theta_t/r_t) > r_t$. The equilibrium level of investment is now lower than before, which implies that the equilibrium demand for real money balances is higher. It follows that the price-level in this debt-constrained economy is lower than what would prevail in an economy where investors are not debt-constrained. Exogenous variation in creditor sentiment θ_t would in this case generate a business cycle with a procyclical price-level.

Let θ_t^b denote the "trust" parameter associated with a bank. The assumption that makes banks essential in this model is $\theta_t^b > \theta_t$. That is, suppose that bank debt is more marketable than private debt. There are a variety of reasons for why this may be the case, especially for some segments of the market. For example, bank loan officers are specialists in assessing which small business proposals make sense. Relative to the small investor, banks may also have better access to specialized lawyers, making it easier to write and enforce legal claims against collateralized debt. If this is the case, then there are potentially some gains to trade between banks and investors. In particular, banks could offer to buy investor coupons in exchange for newly-created banknotes. Or, equivalently, banks could lend investors an amount of electronic deposit liabilities against collateral in the form of investor coupons. The bank is, in effect, transforming illiquid private debt into a liquid payment instrument.

Note that banks do not require reserves to perform the liquidity transformation described above. To put things another way, banks do not "lend out reserves." But if this is the case, then where do banks get the money to lend? They create the money they need. As is the case with all promises, they are created *ex nihilo*. It's just that some promises are more credible than others.²¹

Let's consider the following scenario. Imagine that it is prohibitively costly to market private debt.²² That is, assume $\theta_t = 0$ for private debt, so that $k_t = 0$ (investment shuts down) in an economy where investors can only finance capital spending through bond-

²⁰ Note that it would be easy to generalize the model by assuming that θ is distributed in some manner across the population of investors. In this case, a set of investors would be unconstrained. The set of investors who are constrained would receive funding in proportion to their specific θ .

²¹ There is the question of what prevents banks from issuing unlimited quantities of money. In the model here, the over-issuance of bank money would serve to dilute the purchasing power of bank money. In a model that included bank capital, this loss of purchasing power would be absorbed by shareholders. Finally, in reality, bank money is made redeemable at par for reserves. Excessive bank money creation would ultimately lead individuals to convert their bank deposits into cash.

²² To evaluate the realism of this assumption, try buying your next cup of coffee with a personal IOU.

finance.²³ Now, imagine that a banks appear on the scene and that banks are fully credible ($\theta_t^b = \infty$). Young investors now turn to banks, who are willing to lend each investor $p_t k_t$ dollars worth of deposits, where k_t satisfies (1). Investors can now pay workers with bank money. Investors also agree to repay the loan with interest, $R_t p_t k_t$. Of course, to acquire the money to make good on their bank loan, investors need to sell some of their future output $x_{t+1} f(k_t)$ to those in possession of money (in this model, the old workers).

In this way, banks expand the money supply by the level of their lending so that the total money supply (as measured at the end of each period) is given by $M1_t = M + p_t k_t$, where k_t satisfies condition (1).²⁴ The banking sector's end-of-period balance sheet is transformed from Figure 2 to Figure 4.

What determines the price-level in this economy where part of the money supply is now created by banks through the act of lending? On the one hand, we have the total money supply $M1_t$. On the other hand, the total nominal demand for money is now given by $p_t y_t$. That is, when young workers look into their bank accounts, they see $p_t y_t$ dollars sitting there after selling their goods to old workers and young investors. The market-clearing condition is therefore given by,

$$M1_t = p_t y_t \quad (9)$$

This latter condition can be restated as $M + p_t k_t = p_t q_t + p_t k_t$ or, more simply, as $M = p_t q_t$. But this is exactly the same market-clearing condition we derived earlier for the narrow bank economy. Hence banking, in this model at least, only has consequences for the price-level to the extent it influences the demand for government money.

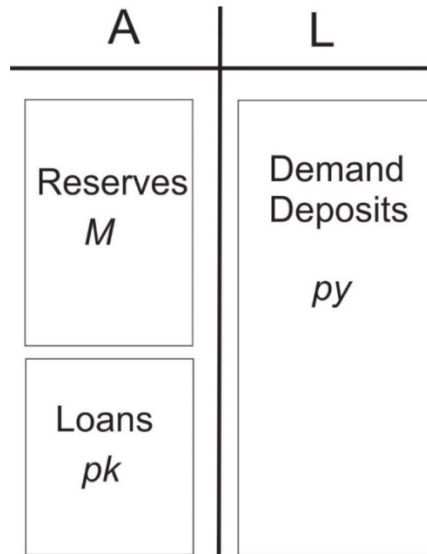
To drive this latter point home, note the price-level in the frictionless credit-market economy is the same as it is in the fractional reserve banking economy (that overcomes all credit frictions). An econometrician studying these two economies sharing the same k_t would measure the money supply to be M in the former economy and $M1_t > M$ in the latter. How is it possible that a larger (measured) money supply is not associated with a higher price-level? I can think of two ways of answering this question. First, the econometrician is mismeasuring the true money supply in the frictionless credit-market economy, since private debt is in effect a form of money (a close money substitute, at least). Second, if one insists on restricting the

²³ More realistically, we can imagine that this is true for a subset of the population.

²⁴ To this total one could add the stock of bank loans created in the previous period $p_{t-1} k_{t-1}$ that is now being used by old workers to purchase goods from old investors. But as old investors collect this revenue, they use it to pay back their bank loans. By the end of the period, this old stock of bank money is "destroyed" as loans are repaid. Repaid bank loans are excluded as part of the date t money supply.

definition of money to bank deposits, then the increase in the supply of deposit money is matched here with an increase in the demand for deposit money.

Figure 4: Fractional Reserve Bank



Let us now compare the price-level in the fractional reserve banking economy (that overcomes all credit frictions) with the price-level in the narrow bank economy without a credit-market. In the former economy, the demand for real government money balances satisfies $0 < q_t < y_t$. In the latter economy, the demand for real government money balances is given by $q_t = y_t$. Hence, the price-level $p_t = M/q_t$ is higher in the fractional reserve bank economy. In this case, the econometrician sees a positive correlation between the price-level and broad money supply (M1). While it would be natural to conclude that the higher price-level was caused by a larger money supply created by banks in the act of lending, the direction of causality does not run from $M1_t$ to p_t .

The direction of causality is as follows. An expansion in bank lending (say, brought about by an increase in θ_t^b) permits an expansion in capital investment, the effect of which is to draw savings away from government debt through a portfolio substitution effect. For a given interest rate, this implies a decline in the demand for government money. Since the supply of government money remains unaffected, the price-level must increase to equate the supply and demand for government money. The additional money created by the banking sector is incidental to this process (the exact same thing happens for an increase in θ_t).

4 Conclusion

Both mainstream and heterodox theories claim that the business cycle is largely attributable to variation in “aggregate demand.” Both theories accept the central role of debt and the desirability of stabilization policies. While the mainstream view accepts the fact that banks create money, the heterodox view emphasizes it. According to the former, bank lending only adds to aggregate demand to the extent it is accommodated by the central bank. According to the latter, this is largely the case but even if it was not, bank lending would still have considerable latitude to influence the money supply and hence, aggregate demand.

The model developed above supports the heterodox view if banks are better than financial markets in the business of relaxing debt constraints. Given the coexistence of bank and non-bank financing, it is not likely that banks have a universal comparative advantage in this regard. But even if one accepts the heterodox proposition, it’s not clear what the implications are for monetary policy. If financial stability is a concern, then a monetary authority must monitor economy-wide leverage, and not just the leverage emanating from the banking sector.²⁵

5 Appendix

5.1 Consumption allocation for loanable funds model

Since young workers only value future consumption, they save their entire income y_t . The only outlet for their saving here is in the form of capital investment, which is desired by investors. Young workers lend their resources to young investors who promise to repay $r_t y_t$ in the future, where r_t is the gross real rate of interest. Thus, workers consume $r_t y_t$ in the second period of their life and investors consume $x_{t+1} f(y_t) - r_t y_t$ in the second period of their life.

Since we know the equilibrium interest rate and level of investment, we can characterize the equilibrium level of consumption for each of the four types of individuals at a given date $t > 1$,

$$\begin{aligned} c_t^w &= x_t \alpha f(y_{t-1}) \\ c_{t+1}^w &= x_{t+1} \alpha f(y_t) \\ c_t^i &= (1 - \alpha) x_t f(y_{t-1}) \\ c_{t+1}^i &= (1 - \alpha) x_{t+1} f(y_t) \end{aligned}$$

²⁵ It should go without saying (but obviously does not) that I do not consider this paper to constitute the final word on the matter. What I hope to have established is a framework that can help organize our thinking on the matter. The purpose of the mathematical model is to help make explicit the assumptions driving various conclusions. This explicitness is meant to invite constructive criticism and avoid semantic debates.

where the expressions above make use of the fact that $\alpha f(k) = f'(k)k$. The consumption allocations (c_t^w, c_t^i) can be thought of as measuring *ex post* (realized) utility while the allocation (c_{t+1}^w, c_{t+1}^i) can be thought of as measuring *ex ante* (expected) utility. Suppose, for example, that $\{x_t\}$ follows a stochastic process. Then an increase/decrease in x_{t+1} can be interpreted as a bullish/bearish change in sentiment. The actual realization may, of course, differ from what is expected.

5.2 Consumption allocation for model with government debt

To complete the characterization of the equilibrium with government debt, combine (6) with (1) and the fact that $f'(k)k = \alpha f(k)$ to derive,

$$c_{t+1}^i = (1 - \alpha)x_{t+1}f(k_t) \quad (10)$$

and

$$c_t^i = (1 - \alpha)x_t f(k_{t-1}) \quad (11)$$

Next, update the government budget constraint (2) by one period and divide both sides by p_{t+1} to form $\tau_{t+1} = (R_t - 1)(M/p_{t+1})$ or $\tau_{t+1} = (R_t - 1)(p_t/p_{t+1})(M/p_t)$. Using the market-clearing condition $M = p_t q_t$, this latter expression becomes $\tau_{t+1} = (r_t - p_t/p_{t+1})q_t$ which, when combined with (5) becomes $c_{t+1}^w = r_t y_t - (r_t - p_t/p_{t+1})q_t$, or

$$c_{t+1}^w = r_t k_t + (p_t/p_{t+1})q_t \quad (12)$$

and

$$c_t^w = r_{t-1} k_{t-1} + (p_{t-1}/p_t)q_{t-1} \quad (13)$$

using the fact that $y_t = q_t + k_t$. This payoff structure across old workers and investors obviously depends on who bears the tax needed to finance the interest expense of the public debt. In the setup here, I have assumed that old workers pay the requisite tax, but given the linearity of preferences and the lump-sum nature of taxes, nothing important (apart from distributional consequences) hinges on who bears the incidence of the tax.

5.3 Properties of the model

In what follows, I describe some of the properties of the model when the economy is shocked in some manner (i.e., when a model parameter changes unexpectedly). I assume that the shocks are unexpected and transitory (so people are surprised, but they view the shock as a one-off event).

1. An increase in the interest rate R_t^* (tighter monetary policy) leads to a decrease in investment and a decline in the price-level. A higher real rate of interest contracts investment demand, inducing young workers to divert their savings to government money. The consequent increase in demand for real money balances is deflationary (individuals are willing to sell goods at lower prices in order to acquire the money they want). The surprise deflation increases welfare for old workers (13), who are in possession of the nominal debt when the interest rate is raised.²⁶ The change in welfare for young workers (12) is ambiguous. The welfare of old investors (11) is unaffected. The welfare of young investors (10) decreases, since their borrowing costs have risen.
2. A decrease in y_t (temporary negative supply shock) implies a contraction in desired saving. The temporary shortfall in saving is absorbed entirely by a decrease in the demand for real money balances—the level of investment (which hinges on the expectation of future economic conditions) remains unaffected. A decline in the demand for real money balances is inflationary. The surprise inflation decreases welfare for old workers (13), who are in possession of the money supply when the interest rate is raised. The change in welfare for young workers (12) is ambiguous. The welfare of young and old investors (11) is unaffected.
3. A decrease in x_{t+1} (bad news shock) leads to decline in investment, since the expected return to investment is now lower. Savers substitute out private sector financial instruments into government debt. The resulting increase in the demand for government money is deflationary—the price-level declines. The surprise deflation increases welfare for old workers (13), who are in possession of the money supply when the interest rate is raised. The change in welfare for young workers (12) is ambiguous. The welfare of old investors (11) is unaffected. The welfare of young investors (10) decreases, since the return to their investments is now expected to be lower.

If we shut down all time variation in the model parameters $\{R_t^*, y_t, x_{t+1}\}$, then the economy settles into a steady-state equilibrium with $r^* = R^* = R = xf'(k^*)$ and $q^* = y - k^*$. The equilibrium price-level is then given by the market-clearing condition $p^* = M/q^*$. The equilibrium rate of inflation is zero.

The transitory shocks considered above generate surprise changes in the price-level, but no change in the path of future price-levels. There is the question of how to properly label a one-time shock to the price-level. If p_t denotes the initial equilibrium price-level and $p'_t < p_t$ denotes the post-shock equilibrium price-level, then $p'_t/p_t < 1$ could properly be thought of as a *transitory deflation* even if theorists tend to label this a “price-level jump.” Note that this price-level jump down portends an *increase* in the expected rate of inflation since $(p_{t+1}/$

²⁶ If I had modeled debt with different maturities, a surprise increase in the interest rate would result in capital losses on long-maturity debt.

$p_t) > (p_{t+1}/p'_t)$. This is simply a property of mean-reverting processes. Consider, for example, a transitory decline in y_t . If y_{t+1} remains unaffected, then the recession induced by the decline in y_t implies an *increase* in the *expected* rate of economic growth.

Note that the Taylor rule (7) has implications for how the nominal interest rate reacts to the shocks considered above. Remember that our specification of the Taylor rule assumes that the monetary authority increases the policy rate one-for-one with the *expected* rate of inflation. So, in the first experiment above, a decline in R_t^* obviously calls for a cut in the policy rate R_t . However, the resulting decline in the price-level leads to an increase in the expected rate of inflation, which means that the policy rate rises by more than R_t^* . In the second experiment, the decline in y_t leads to an increase in the price-level, which implies a reduction in the expected rate of inflation. In this case, the monetary authority lowers its policy rate.²⁷ Finally, in the third experiment above, the negative news shock leads to a decline in the price-level and an increase in the expected rate of inflation. The policy rule here implies that the policy rate is increased. This is not, of course, a description of optimal monetary policy. I am only describing what is implied by a policy that operates according to a rule like (7).

There are ways to use the model to think about the appropriate conduct of monetary actions on an *ad hoc* basis (which is basically how policy works in the real world anyway). Suppose for example that a decline in x_{t+1} is the product of a growing irrational pessimism (a brooding animal spirit). The policy rule (7) keeps the real rate of interest stable. The implication is a large decline in capital spending, a flight-to-quality (government debt) and a transitory deflation. There would be some rationale, in this case, for the central bank to lower R_t^* in its policy rule—the effect of which would be to temporarily lower the policy rate to keep investment demand stable. If the pessimistic outlook over the return to capital spending turns out to be wrong (inconsistent with fundamentals), then investors will be glad that the central bank intervened the way it did when it did. Of course, policy mistakes are also possible.

6 References

- Andolfatto, David (2015). A model of U.S. monetary policy before and after the great recession, *Federal Reserve Bank of St. Louis Review*, Third Quarter, pp. 233–256.
- Fullwiler, Scott (2012). Krugman’s flashing neon sign, Naked Capitalism blog, April 2, 2012; <https://www.nakedcapitalism.com/2012/04/scott-fullwiler-krugmans-flashing-neon-sign.html>

²⁷ It is commonly asserted (with some justification) that central banks often make the mistake of tightening monetary policy in the face of a transitory negative supply shock (e.g., an increase in oil prices). The model here suggests that the appropriate course of action is to lower the policy rate.

- Keen, Steve (2012). Instability in financial markets: sources and remedies, INET Conference Berlin; <https://www.ineteconomics.org/uploads/papers/keen-steve-berlin-paper.pdf>
- Kocherlakota, Narayana (2012). Incomplete labor markets, Federal Reserve Bank of Minneapolis Working Paper.
- Koning, J.P. (2013). Do banks have a widow's cruse? Moneyness Blog, August 28, 2013; <http://jpkoning.blogspot.com/2013/08/do-banks-have-widows-cruse.html>
- Krugman, Paul (2012a). Minsky and methodology (wonkish), *The New York Times*, March 27, 2012.
- Krugman, Paul (2012b). Banking mysticism, *The New York Times*, March 27, 2012.
- Lagos, Ricardo, Rocheteau, Guillaume and Pierre-Olivier Weill (2009). Crises and liquidity in over-the-counter markets, NBER working paper no. 15414.
- Mason, J.W. (2012). The case of Keen, The Slack Wire blog, April 1, 2012; <http://slackwire.blogspot.com/2012/04/case-of-keen.html>
- Pettifor, Ann (2014). The Keen-Krugman debate, Open Democracy U.K., February 26, 2014; <https://www.opendemocracy.net/ourkingdom/steve-keen/keen-krugman-debate>
- Roche, Cullen (2013). Economics and banking—James Tobin 1963 (Redux), Pragmatic Capitalism, September 23, 2013; <https://www.pragcap.com/economics-and-banking-james-tobin-1963-redux/>
- Rowe, Nick (2012a). The supply of money is demand-determined, Worthwhile Canadian Initiative blog, April 1, 2012; http://worthwhile.typepad.com/worthwhile_canadian_initi/2012/04/the-supply-of-money-is-demand-determined.html
- Rowe, Nick (2012b). Banking 'mysticism' and the hot potato, Worthwhile Canadian Initiative blog, March 31, 2012, http://worthwhile.typepad.com/worthwhile_canadian_initi/2012/03/banking-mysticism-and-the-hot-potato.html
- Tobin, James (1963). Commercial banks as creators of money, Cowles Foundation Discussion Paper No. 159.
- Waldman, Steve Randy (2012). Banks and macroeconomic models, Interfluidity blog, August 25, 2013, <https://www.interfluidity.com/v2/4522.html>