Exogenous and endogenous monetary systems and excess deposit growth
Vincent Spreuwenberg

Abstract

The monetary system consists of two distinct systems, one that circulates existing money and one that creates and destroys money. This observation by itself is nothing new but by taking a “compare and contrast” approach several differences between these systems stand out. Having a firm grip on the distinction between these two systems then also helps to explain why certain “standard” economic theories like the quantity of money theory are so difficult to verify empirically.

Understanding the difference between exogenous systems and endogenous monetary systems also makes it clear why it is inevitable to have deposit growth independent of lending or asset purchasing activity by deposit issuing banks. The paper then goes on to explain how a second factor is re-monetization of already monetized assets. Both factors lead to deposit growth in excess of the net of pledged, real, assets. These excess deposits change the role of banks as liquefiers of assets to investors in risky assets. This weakening relationship between deposits and real assets may offer a partial explanation of various monetary phenomena like the global downward tendency in interest rates.

1. Introduction

Classifications like M0, M1, and M2 define the “moneyness,” the degree to which assets approximate cash, of certain assets but do not address what these moneys are, how they come into existence, and how they disappear. For example, money market funds are considered to have a different level of “moneyness” than bank deposits or cash. But how exactly did the money that was used to purchase the money market funds come into existence in the first place?

Concepts like high powered money, inside money, and outside money¹ are used by some to explain how the monetary system works and where money comes from but are less than satisfying. But they are good starting points from which to explore various explanations of how our monetary system works.

This paper begins by looking at traditional descriptions of the banking system and concluding that there are two separate and distinct banking systems which are interwoven. Taking these two systems each to their logical extremes lays the foundation for the rest of the paper. The next section delves into the nature of two different types of money that exist. This section explores how the two types of money shape the systems that they exist in and examines the interaction between the money and those systems. At this point we see that the validity and relevance of concepts like MV=PQ actually depends on an assumption with respect to the type of (unarticulated) monetary system. The final sections look at losses that can occur in the two

¹ The high-powered concept is a reference to the monetary base, which presupposes that there is a fractional-reserve lending system which then multiplies this base money. Inside money generally refers to money which is created by (private) banks. Outside money is money that doesn’t have a liability associated with it, such as gold-backed money or other assets that are not directly associated with the economy in which the money circulates. “Inside and Outside Money” (Lagos, 2008) goes into detail and is certainly worth reading.
different systems and the consequences different types of losses have. Losses have different
effects in the two systems, and one type of loss can lead to an important concept that I refer to as
“stranded money.” The concept of stranded money along with relending of financial assets leads
to insights that may help explain monetary events like the multi-decade global decline of interest
rates.

“Once you eliminate the impossible, whatever remains, no matter how improbable, must be the
truth.”
Arthur Conan Doyle

2. A Tale of Two Systems
Two types of monetary systems currently exist in parallel and different types of money exist in
those systems. One system, exogenous money, is a closed monetary loop and separate from the
productive economy where we make goods and services but is used by that productive economy.
The other system is created in and used by the real economy, which makes it endogenous, and it
is open, not closed.
The easiest way to think about these two monetary systems is through the historical context of
the traditional banking system classification that is used in any introductory money and banking
class. In addition, understanding two features of all banking systems—the nature of deposits and
negative carry—provides necessary context. It is then possible to examine the different types
of money that exist.

The three traditional theories of banking are:
1. Financial intermediary theory of banking: Banks transfer money from savers to investors.
2. Fractional reserve theory of banking: An individual bank has a certain amount of “real”
money and is allowed to create credit as a ratio of the reserves—the “real” money—that
it has.
3. The credit creation theory of banking: Banks create new money through lending.

After reviewing these three systems we will whittle them down to two: the financial intermediary
system and the credit creation system. We will then take these systems to their logical extremes
to create a framework. To be clear, these extreme versions do not exist in reality, or at least not
in any economically meaningful way, but help to clarify the theoretical boundaries of the system
that we actually have and live in.

2.1 Banks as financial intermediaries
In a financial intermediary banking system, banks distribute money between various parties.
Savers deposit money in a bank, and the bank in its capacity as an intermediary invests the
money, or enables the money to be invested by a borrower, in something that has a high enough
return to compensate the depositor as well as the bank for the use of the money.
The roots of this system are based in the times when commodity money was the basis of the
financial system. It is also the system that your average (non-bank) person operates in. When you
borrow a dollar from your neighbor, she or he no longer has access to that dollar. The dollar is
neither created nor destroyed; it just changes ownership, and at some future point the dollar
(hopefully) gets repaid to its owner, usually with interest as compensation.
Under the Glass-Steagall act, investment banking firms fit this role explicitly. They did not provide credit but allocated capital between savers and investors. The money circulating in this system was often issued by governments, but private banks also issued money. Depending on experience and sentiment, money issued by certain banks was more desirable than that of other banks.

2.2 The fractional reserve system

The concept of fractional reserve banking grew out of the days when goldsmiths issued more receipts for gold than the amount of physical gold they held. However, the meaning of the fractional reserve system changed at various points in history. The Promissory Notes act of 1704 in England was key to the development of fractional reserve banking because it made promissory notes legal as negotiable instruments. In other words, it made gold warehouse receipts into legal money (Encyclopedia of Money, 2011). In theory, those gold warehouse receipts were “as good as gold.”

Goldsmiths, however, charged interest on the receipts they issued. In order to earn more interest, they started issuing more gold receipts than there was gold in their vaults. As a result, the value of the physical gold in their vaults became literally a fraction of the value of the receipts in circulation. The question that came out of that decision was how many more liabilities in the form of deposits and banknotes versus assets were allowed to exist. Or, in slightly different terms, what should the reserve ratio be? A different way to look at this problem is to view the reserve ratio as a tool to make assets (gold) balance with liabilities (the money issued by the banking system).

The first order reason for a reserve ratio to exist, and to set a maximum quantity of money that could be issued, was to have a reasonable defense against bank runs, where more withdrawals of gold were demanded than there was physical gold in a vault. Taking a step back however, authorities were really solving for three variables, a triangle. On one side you had the physical quantity of gold that the private banking system had in its possession. A second side was the quantity of currency demanded by society. And on the third side was the probability of depositors wanting to cash in more receipts than there was gold. The latter was an issue because it is impossible to give depositors more gold than exists in the vault. The greater the ratio of money required versus quantity of gold available, the greater the probability of a bank run, where any particular bank would run out of gold.

When the Federal Reserve system came into existence in 1913, it solved two of the problems: that of bank runs and gold shortage, at least on a very short-term basis. When there were excess withdrawals at a bank, which were driven by more demand for (gold-based) money than there was in the vault, the Federal Reserve acted as lender of last resort to the banks by using its own gold stock. The third side of the triangle, the pricing of gold, however, was not yet managed.

2.2.1 Gold pricing
On the gold standard, whether it was a deposit or physical currency, money was essentially an IOU for gold. You could literally take, say, $32 in cash to the Fed, and they would give you an ounce of gold. However, the quantity of gold in the vaults of banks changed at a slower rate than the quantity of deposits that had been created and bank bills—physical currency—that were in existence. To manage the relationship between the quantity of gold and the quantity of dollars, the price of gold had to be adjusted again and again. One ounce of gold could be exchanged for more and more dollars not because the dollar was losing value, but because the quantity of gold was limited. To make the quantity of gold balance with the number of dollars in circulation the exchange ratio had to change.

In the 1920s, every time dollar/gold was devalued, authorities were essentially solving a quantity mismatch problem. The government owned a certain quantity of gold, which served as collateral, but society demanded an increasing quantity of dollars. That demand was met by money created by privately held commercial banks in the form of deposits and a sprinkling of physical currency issued by the Bureau of Engraving and Printing.

The only way to make dollars and gold balance was to modify the exchange rate, that is, increase the price of gold expressed in dollars. The price of gold—the exchange rate between physical gold and dollars—was simply the mathematical tool that tied the quantity of gold to the amount of money in circulation. In a way, the price of gold became a metric of how much money the economy required versus how much physical gold was available. From this point of view, it wasn’t a dollar devaluation/gold appreciation situation per se, and the dollar wasn’t necessarily going down in value as it related to the goods and services—other than gold—that it could be exchanged for. The demand for dollars increased faster than the quantity of gold held by the banking system. As a result, the equilibrium price required to make the quantity of gold balance with currency demand was below the market price of gold and had to be increased regularly. The divergence between the monetary gold price and the market gold price had as a consequence that the private sector could no longer be allowed to own physical gold. Without limiting ownership, one could exchange 32 one-dollar bills for, say, 1 ounce of gold, then turn around and sell that gold for more than $32 in the open market. This is why during those times gold had two prices—a free market price and an official, monetary price.

The convertibility of dollar into gold created a situation where everyone who wanted had a direct relationship, albeit not an account, with the Federal Reserve. Anyone who wanted could drain reserves from both private banks (by swapping deposits for cash, and from the Federal reserve (by swapping cash for gold). With the dollar becoming the global means of exchange, with the result of large quantities of dollars in the hands of potentially unfriendly powers, closing the gold window made a lot of sense.

The implicit equation behind the fractional reserve system was:

\[
\text{Quantity of gold owned by central banks} \times \text{Official price of gold} \times \frac{1}{\text{reserve ratio}} = \text{monetary stock}
\]
Over time, the reserve ratio gradually became something different. In 1970, gold was formally—supposedly temporarily—de-coupled from the dollar, and thereby indirectly from many other currencies. However, central banks maintained the reserve ratio concept. The gold side of the reserve ratio was first replaced by reserve assets. Eventually, bank shareholders’ equity, in effect, replaced the implied commodity-backed money versus the newer credit money. The change ended up implying that there was a difference between commodity-backed money and credit. Over time, the fractional reserve theory has lost credibility, and even the Federal Reserve removed reserve requirements from the banking system in March 2020 (Board of Governors of the Federal Reserve System, 2020), effectively killing the fractional reserve concept. When the Fed took this action, it was an implicit admission that the fractional reserve concept in modern banking no longer existed.

In the most recent iteration of regulations, the US Federal Reserve uses several explicit equity/deposit related ratios to assess the health of banks. The idea behind this is that if the bank makes bad loans and loans don’t get repaid, there will still be enough money in the form of owners’ equity to make depositors whole. The fractional reserve theory then was a bridge between banks acting as intermediaries and banks acting as creators of credit.

2.3 The credit creation theory of banking
A third theory of banking describes the credit creation system in which banks create money through lending. The sequence of events in the credit creation system is the opposite of the system in which banks act as intermediaries. The bank accepts collateral for a loan and simultaneously creates a deposit. The system is open ended, and money can be created and destroyed at will through the issuance and repayment of loans. The deposits created in this process are practically indistinguishable from existing—“real” (commodity based)—money.

2.4 Bank Deposits
A cursory look at the three theories of banking shows that bank deposits play an important role. Although significant, the observation that deposits can be created by a bank in the process of lending is nothing new.

From a purely economic perspective in a modern banking system, bank deposits are similar to store credits for returned items. Both give you the right to future purchasing power, but they differ in some important details.

A store credit is generated by returning an item or by something like a gift card, and those items were obtained by swapping existing bank deposits or cash for either the item or gift card. The credit can only be used in that store and often for a limited time only.

In contrast, bank deposits can be used to settle any debt at any time—during business hours, at least—and don’t have an expiration date.

But what are bank deposits? Bank deposits are unsecured IOUs from businesses that have a banking license. They are widely accepted as final settlement of debts. Although widely accepted

---

2 For a good summary of authors who have written about this you may want to read the appendix to Chapter 2 of Morgan Ricks’ (2016) *The Money Problem – Rethinking Financial Regulation*. Electronic copy available at: https://ssrn.com/abstract=4637150
as payment, bank deposits are not government liabilities. Even so, the government itself will accept bank deposits to pay tax liabilities and will transfer deposits to private banks for payments to the private sector.

Deposits can end up on the balance sheet of a bank in two ways. First, a bank deposit is created when a customer deposits physical currency, which is issued by an agency of the federal government, or when another bank sends a deposit to the bank in the name of the customer. The bank then creates a liability (to the bank), called a deposit that is an asset to the customer. The customer then can use the deposit to pay their liabilities, either by transferring some of the deposit or by receiving cash in exchange for a reduction in the deposit.

Note that a deposit of cash involves swapping a federally issued instrument—dollar bills—for a privately issued credit instrument. Therefore, making a cash deposit constitutes exchanging a U.S. government liability for a private business liability. Depositing cash in a bank entails an increase in credit risk while not getting compensated for that increase in risk. You receive only $1 dollar of bank credit for every dollar denominated bill you deposit.

In contrast, creating a deposit through a bank transfer involves the exchange of a credit instrument from another private bank. There is no change in credit risk.

This distinction may sound theoretical, but especially during times of financial stress the question of who issues the liability, or deposit, becomes very important. For example, during the financial crisis of 2008 it became extremely clear that not all banks were created equal.

Banks are businesses, and they can, and do on occasion, go out of business, in some cases leaving creditors short of what they are owed. Since 2009, over 500 banks in the U.S. have failed, so although bank failures are not everyday events, they happen with some regularity (Goldberg, 2022). Therefore, larger depositors, who are not covered by FDIC insurance, often purchase deposit insurance. More often, they may purchase treasuries (through reverse repos) at the end of the banking day to hold overnight. That way, if something happens to the bank overnight, the account owner will own treasuries—liabilities issued by the federal government—rather than unsecured deposit instruments from a private bank, which may be worth less than what the deposit balance was the previous day. Larger depositors tend to be more sophisticated and monitor the financial condition of their bank much more closely than the average retail depositor.

The moral of the story is that deposits are not a tally for money that was given to a bank for safekeeping. Banks are not custodians of deposits. They use deposits to invest in (hopefully) positively yielding investments in order to generate revenue and with, hopefully, low enough risk that the deposit is available when you want to have access to it.

It should be clear from the mechanics operating in the different banking systems that the money that circulates in them is different. The money in the intermediary system is static in quantity and determined by decree, or at least by some force that is only tangentially related to the real economy. In contrast, the quantity of money in the credit banking system is dynamic and determined by the loans that are made and their collateral.

The nature of these two different types of money has shaped the systems they exist in. Although money and banking systems are intimately intertwined and reflexive to a degree, I’d argue that the nature of a particular type of money drives the shape of the system, not the other way around.

2.5 Negative Carry
As useful as money is, there are only a few things you can do with it: spend it, save it, invest it, or hold it. Holding money has a cost, known as negative carry. Negative carry is an attribute of money in general and impacts exogenous and endogenous systems and their moneys equally. Negative carry seems negligible or even zero over a short holding period, but it is not. Over longer periods negative carry is significant, especially when you take into account that it applies to all currency balances that exist, not just one individual’s. Think of the size and cost of the financial sector to have a better sense of the size of the issue.
Traditionally, money is said to act as a store of value. Store of value tells you that if the current holder intends to keep the currency for a while before exchanging it, the currency is expected to hold its purchasing power. However, negative carry undermines the ability of money to act as a store of value. In this section we’ll look at various aspects of negative carry and its consequences for different types of money, including gold, deposits, fiat money, and Bitcoin.

2.5.1 Gold
Proponents of gold like to point out that it has maintained its value over time and how that value has never completely disappeared, unlike that of many fiat currencies. This may appear to be true, but it is a grossly incomplete statement of affairs. What is missing from that observation is the negative carry, particularly the storage cost, that comes with holding gold and how it impacts the economics of holding on an individual and aggregate level.

2.5.1.1 Individual Costs of Holding Gold
There are two reasons for storing gold: entropy and humans. Gold is a very soft metal. If gold coins were handled by people all the time they would literally wear away into very fine gold dust. Although technically the gold would still exist, it would be in a form that is unusable as currency and can be considered lost.
Humans are a vastly more pressing reason for safeguarding gold. Humans steal. Humans lose things. Humans die without telling their heirs where they stored the gold. The more gold gets handled, the greater the chances of loss. Some lost gold will be found—think of coins that are dug up or scooped out of the ocean. But some won’t. The latter losses are lost to the system.
If you believe that gold is a currency and not just an asset, you will use gold as a method of payment in your transactions as well as hold it. If you assume that gold is both a currency and an asset, the ultimate value of your gold holdings will be zero because you will pay for the storage of your gold asset with gold currency. Even if you don’t physically scrape a few grams off a gold bar to make your annual payment for storage, in effect that is what you do.
If you start with a 100oz gold bar and pay 1oz in storage costs per year (1% of your holdings is not an abnormally high storage fee), after 30 years, 70% of your holdings are left. After 100 years, your heirs have literally nothing. It has all gone to storage costs, or negative carry. You may be using gold that you earned through other sources to protect your gold. But using gold that flows into your pocket rather than from your stock to pay for storage costs is simply a matter of how you choose to account for your expenditure.
After you have died, if your heirs decide to keep the gold bar, they are quite literally throwing good money after bad because the cumulative storage costs are in excess of the value of what they are trying to protect.

2.5.1.2 Aggregate Costs of Holding Gold
At the aggregate level, resources are spent to keep physical gold from disappearing and to safeguard it. As more gold spreads around the globe for payments, more of it gets lost or stolen and more goods, services, and energy are dedicated to maintaining it. The resources expended are significant. The bigger the aggregate hoard of gold society has and the more widely it is distributed, the more society spends on keeping other people from having it. Protecting gold is not the most productive use of labor and resources.

The more gold that has been dug up and concentrated, the more that can be lost—for example, large hoards are more attractive to steal. So, the effective loss rate will increase over time unless more and more resources are consumed to protect the hoard. The presence of gold, then, increases entropy in the larger economic system without any corresponding increase in productivity. The opposite occurs because society will spend more resources protecting what should be a medium of exchange, something which enables transactions between productive goods and activities. Instead, more and more resources are devoted to keeping the medium of exchange secure.

2.5.2 Fiat currency
Fiat currency comes in two forms. The first form is the physical bills and coins you can put in your pocket. The second form is electronic. Both forms incur negative carry.

2.5.2.1 Individual Costs of Using Fiat Currency
Bills and coins can wear out, get lost, or be destroyed. The average $20 bill lasts about 8 years before it is shredded and a new one must be printed to replace it (Board of Governors of the Federal Reserve System, 2020). Coins last longer, but quite a few of them get lost (Picchi, 2016); no one knows how many are found. So physical fiat currency has negative carry: it wears out and needs to be replaced every so often.

Once you own currency you can lose it. It can drop out of our pocket, your kid can steal it, and so on. So, you need to protect it, which incurs negative carry.

Even though physical currency has negative carry, it also has a plus side in that it creates value as it is made. With the exception of pennies and nickels, coins and bills have face values higher than the cost to produce them. If you ignore inflation, you can argue that minting and printing money creates value. Physical money is a significant portion of the total money stock. As of December 2022, $2.307 trillion worth of federal reserve notes were floating around (Board of Governors of the Federal Reserve System, 2022), about 1/3 of the $20 trillion in M1 (FRED Economic Data, 2022).

The electronic form of fiat money requires inputs, just like physical money. You need a device like a computer or a smartphone to deposit, transfer, and access your money. To use the money, you need a distribution device like a credit card, debit card, or payment app. Those have all kinds of digits associated with them that you need to keep track of and keep secure. Keeping physical and electronic currency functional requires a continuous stream of inputs. The devices and other resources required add up to a significant negative carry which is hard to spot because the costs are spread among all the participants in the system.

2.5.2.2 Aggregate Cost of Fiat Currency
After physical currency leaves your pocket or electronic currency leaves your account, both require resources to maintain and keep them useful. Banks have huge administrative machines
to keep track of all the numbers. Those numbers are also tracked and checked by regulators, auditors, and the like.

Even though most transactions and all record-keeping are now digital, many people still receive paper statements. We are continually burning fossil fuels and hacking down trees to keep track of our currency.

Regardless of the form of currency or statements, the processing apparatus needs to be fed labor and inputs to keep it in existence and operating. In Q2 2022, the financial and insurance sector represented 6.9% of GDP or about $1.3 trillion (Bureau of Economic Analysis, 2022). The cost of running the financial system is not trivial.

2.5.3 Bitcoin

Experimental digital forms of currency, like cryptocurrency, have negative carry above and beyond electronic forms of fiat currency. I call cryptocurrencies experimental because they are not yet widely accepted as forms of payment. All cryptocurrencies are digital and require computer resources to function.

Bitcoin, for example, is generated through a computer process called mining. The computing that miners do validates and records Bitcoin transactions. The miner provides a service and in return gets a Bitcoin as compensation. To get a Bitcoin you can mine it yourself or exchange fiat currency (or gold, if you like) for it—i.e., buy it.

Holding a Bitcoin or using one requires a secure digital wallet. The digital wallet contains a private key and a matching public key, which are both long strings of numbers. To carry out transactions with Bitcoin, you share the public key. The Bitcoin itself resides on a distributed database called a blockchain.

2.5.3.1 Individual holding costs

As long as you own a Bitcoin there is a chance you will lose it. Once you lose it, it is highly unlikely that you can recover it. It is estimated that around 20% of all Bitcoins (that is quite a bit more than zero) that exist have been irretrievably lost (Reiff, 2019).

Losing Bitcoins is different from losing gold or fiat currency. Unless you write your public key on a piece of paper and lose the physical paper and someone finds it, once a Bitcoin is lost someone else cannot find it. Thus, there is an asymmetry that gold and fiat don’t have. Stolen gold or currency still exist but are owned by a different party. Bitcoin cannot be stolen without your public key, and once lost is permanently inaccessible.

To keep access to your Bitcoin completely secure, you need to store your public and private keys on an off-line computer or external storage device and put that in a secure place. In a way, the situation is similar to owning gold—a physical object that needs to be safeguarded.

If you believe Bitcoin is a currency, you will pay the safe-keepers and infrastructure providers with that currency, and they will end up owning all Bitcoin. Your computer, electric bill, and other infrastructure will all be paid for in Bitcoin. Your stash of Bitcoin will slowly but surely get smaller because of this negative carry.

2.5.3.2 Aggregate Holding Costs

Bitcoin and Bitcoin transactions depend on computer networks that continuously consume energy and over time need to be upgraded and replaced. The energy used is primarily generated with fossil fuels. To the extent fossil fuels are involved, the cost levied on society cannot be recovered.
Once all Bitcoins are mined and there is no longer a reward for people to dedicate computer resources to verify transactions, society must supply the resources. Either businesses will charge for Bitcoin ownership and usage, or government will implement some type of taxation. Usage fees or taxes will be levied in, of course, Bitcoin. There is a cap on the number of Bitcoins, so permanent losses by individual Bitcoin holders will cause the quantity of Bitcoins available to society to shrink every year. This can become an impediment to transactions and lead to currency shortages, imposing additional negative carry. The lesson here is that all current options for currencies have negative carry, which means they consume themselves. They exist in a system where slowly but surely their value erodes. All need external inputs to keep existing or to remain in your possession. They are not productive assets, so they cannot act as stores of value.

Money then is not and cannot be a store of value in any economic system as long as that same money is used to create, maintain, and safeguard itself. Entropy makes the combination of being a store of value and a medium of exchange incompatible. The more money there is in circulation, the less efficient the economy is because a greater part of the productive part of the economy needs to be directed to safeguard and manage the currency. The only roles money— whichever flavor you prefer— can serve is as a unit of account and as a way to settle transactions between economic players. The real question then becomes which currency requires the least external inputs to exist.

3. Two categories of money

The next question is whether there are differences in the moneys that exist in the intermediary and credit banking systems. The answer is system dependent and a conditional “yes.” Money can be categorized into two types, exogenous money and endogenous money. The concept that there are different types of money in existence is not new. For example, Cullen Roche in his paper Monetary Realism (Roche, 2015), Zoltan Pozsar, and various Modern Monetary Theory papers have used terms like inside money and outside money. In the papers written by these authors, however, the differentiation tends to revolve around the issuer of money (private or public) but not the nature of the money itself. In this paper the focus is on the nature of the money itself, not who the issuer is, although we touch on that at several points.

These two types of money are similar in their daily use but differ in their origin and final destination. The differences in their fundamental nature have far reaching consequences for how they work, how the systems in which they are used evolved, and how any new system must be designed.

The first type of money in this discussion, exogenous money, exists independently of the real, productive, economy. Exogenous money has a long history and was the prevalent type of money until relatively recently.

3.1 Exogenous money

Visualize an island where, for example, gold is used as money. A certain quantity of gold exists on the island and is unrelated to the population and unrelated to economic activity. Even if a disease wipes out all the people on the island, the quantity of gold will remain constant. The same thing is true the other way around: when the first habitant lands on the island all the gold is already there. If one person loses a coin somebody else will pick it up.
An example of such a system would be a 100% gold-based system with a fixed quantity of gold and where the only accepted medium of payment is gold. Prices in such a system would then be a function of two factors: first, the utility of a particular good or service; and second, the available quantity of money.

In a strict exogenous monetary system, the quantity of money and the real economy are unrelated. Think of the game Monopoly. There is a certain quantity of money in the box and that quantity is unrelated to how many players there are or how good they are at playing the game. The total quantity of money in an exogenous system is determined either by decree or by happenstance. This type of money does not have a counterparty associated with it, so it is always one-sided. An example, again, is gold. The owner simply owns it, and no other party is involved. The difference in how exogenous money is used in this essay versus outside money in other papers is that this paper separates the ability of an average economic participant to create it at will from money that is created by some central authority, whether that authority is private or public.

In other writings, there generally is a mechanism by which a central bank has the ability to create reserves, or inside money at will. In this description of exogenous money, we ignore for clarity’s sake who the issuer is or how the quantity of money is determined.

“Shadow banking” is what some call a part of this system, where firms that do not have the ability to create deposits will acquire deposits and then relend them or invest them on other types of assets. The operators in this ecosystem are part of the exogenous system but actually hold their deposits at commercial banks. Whether the deposits they own are created as the result of liquifying assets is not clear. In any case, the deposits they control and lend out very often don’t lead to more deposit creation but essentially transfer deposits between parties.

Exogenous money requires relatively little in terms of administration. The only thing that needs to be tracked is ownership. Traditionally, pure exogenous money was physical, which made tracking it relatively straightforward. Nowadays however, exogenous money is not necessarily physical, as we will see.

Whether or not exogenous money is physical, it is important to understand the source of its value. It is also important to understand how lending, interest rates, aggregate assets, banking, and central banking function with exogenous money.

### 3.1.1 Source of Value

The value of exogenous money is bestowed on the currency through its exchange value. It is desirable and valuable because it is widely accepted in a particular society as exchangeable for goods and services. This money has value equal to the utility of the goods and services it can be exchanged for. Any utility that the money may have as a commodity is incidental and should be less than the exchange value. Otherwise, the currency (for example gold) would be used for industrial purposes rather than as a monetary instrument, leading to currency shortages.

Widespread acceptance of exogenous money creates a one-to-many relationship with goods and services. Any gold coin, for example, can purchase any good but only one physical good can be exchanged for one (of many essentially identical) coin. This is different from barter where both parties must desire each other’s good (known as the double coincidence of wants), which makes it a one-to-one relationship.
Exogenous money is worth whatever you can exchange it for. It can function only if most economic participants use it. This feature then implicitly sets the relationships between other goods and services because their value is expressed in the same unit of exchange. All transactions in this system, including saving and lending, are then, in effect, asset swaps. In an exogenous monetary system, goods and services are created and then are exchanged for the currency (in this example, gold). In transactions, exogenous money is only exchanged, not created or destroyed.

The only question with exogenous money is who owns it. There are no questions about whether the money itself should exist or how much should exist.

The exogenous monetary system is akin to a water balloon: no matter how hard you squeeze it or how fast you squeeze it, the quantity of money is fixed and stays inside society, which is why several concepts from the physical world apply. Concepts like Savings=Investment, velocity of money, the quantity theory of money in all its incarnations, and the crowding out concept in borrowing and lending make sense because the monetary system is closed and the quantity of money in the system is constant while demand for money is variable. Money is recycled and is never created or destroyed.

If economic activity in this system expands, the quantity of currency required to enable economic activity increases. However, because the quantity of money is fixed, the quantity available per transaction decreases. The same quantity of money is divided by a larger number of transactions per time unit, and money becomes relatively scarcer.

Every price is, in essence, an exchange rate between the exogenous currency and a particular good or service. The price or exchange rate between the money and the good or service is influenced by two factors. The first factor is the value or utility of the good or service versus other goods and services. The second factor is the availability of the money, which is impacted by its overall scarcity. What appear to be price changes in goods and services, when expressed, for example, in grams of gold may in fact be just an artifact of the fixed quantity of money rather than a real change in utility, or even relative utility, of a good or service.

Savings which are not lent out, i.e., not available for circulation, can be problematic because this reduces the amount of currency available for transactions. If an economy has more currency units than transactional demand, savings is not an issue. But if economic participants save, at some point transactional currency demand will exceed currency in circulation. Once some participants consistently spend less than they earn and don’t make their savings available to be lent out, it becomes inevitable that the quantity of money in circulation goes down and its scarcity value goes up, increasing the incentive to save even more.

Absent periodic redistribution, a large portion of exogenous money tends to end up in the hands of a few. With too little money in the hands of others to enable transactions, it needs to be redistributed for society to function. Anyone who has played a game of Monopoly knows how this ends: one person ends up with all the assets, and the game has to be restarted by dividing up the money and the assets. With this in mind, it makes sense that societies that use a strict form of exogenous money need to undergo a monetary reset every so often where money gets redistributed to various economic participants. During biblical times, money was mainly
exogenous, and Deuteronomy mentions debt forgiveness every few years.³ The problem that redistribution of money is trying to solve is not one of productivity or societal injustice but lack of liquidity, which has the potential to curtail economic activity below where it otherwise could be.

3.1.2 Lending
If savings are lent out in an exogenous system, it is reasonable to charge interest because the lender temporarily loses use of the money and is exposed to a risk of loss. That interest however must come out of the available, existing fixed pool of currency, leaving less available for other purposes. The payment of interest in an exogenous system is problematic because it adds to the demand for currency while there is no actual supply for that currency. The currency that was lent out may be swapped for (or spent on) investments, which increases productivity, but increased productivity doesn’t increase the quantity of money that exists.

The fixed nature of exogenous money also implies that successful lending and borrowing, where money is repaid at an interest rate greater than zero, is deflationary. The only reason somebody will borrow money at an interest rate X is that it is valued at greater than X (for example if somebody really needs the liquidity) or they expect to earn more than X. In other words, the expected change in value of assets in society is greater than the growth in the currency stock, which is zero, so the ratio of assets to money stock keeps increasing. Every unit of currency relative to assets becomes more scare and therefore more and more valuable. If this trend is consistent and material enough, it can provide a disincentive to spend money either on consumption or investment. In other words, it can lead to a reduction of liquidity or even seizing up of the financial system.

Even in a system with a fixed amount of money in existence, however, the amount that can be borrowed and loaned out is effectively unlimited because money can theoretically be borrowed and then re-loaned. But that doesn’t change the total net quantity of money available.

3.1.3 Interest Rates
In addition to the usual factors determining interest rates, such as expected return and risk, other factors play a role in an exogenous system. The main issue is that the quantity of money that is available to be lent out is fixed. The interest rates at which transactions take place are a function of both demand and supply of money. But in an exogenous system, the aggregate quantity of money is fixed so the net lendable quantity of money has a hard upper limit.

In an exogenous system, where money is fixed and limited in quantity, the notion of the interest rate as a proxy for risk makes sense. A lender takes risk when making the loan, so higher risk projects should pay a higher interest rate.

The combination of a limited quantity of money with varying demand for capital creates a potentially non-linear interest rate-to-risk relationship due to the inability of the supply of money

---

³ Deuteronomy 15:1-23 “At the end of every seven years you shall grant a release. 2 And this is the manner of the release: every creditor shall release what he has lent to his neighbor. He shall not exact it of his neighbor, his brother, because the LORD’s release has been proclaimed. ³ Of a foreigner you may exact it, but whatever of yours is with your brother your hand shall release. ⁴ But there will be no poor among you; for the LORD will bless you in the land that the LORD your God is giving you for an inheritance to possess— ⁵ if only you will strictly obey the voice of the LORD your God, being careful to do all this commandment that I command you today. ⁶ For the LORD your God will bless you, as he promised you, and you shall lend to many nations, but you shall not borrow, and you shall rule over many nations, but they shall not rule over you.
to adjust to market forces past the available quantity of lendable funds. The supply curve for lendable money does have some elasticity but is still constrained by the total amount of money available in the system.

Demand beyond that upper limit will result in bidding up the price of money (the interest rate) without any increase in the quantity of money lent. This causes a larger piece of the economic pie to go to capital providers, the lenders, rather than the producers of value, as would be the case if the supply curve were elastic.

One consequence of the limited quantity of capital is that the viability of a project is driven by two forces rather than one. First, the cost of capital is driven by its scarcity because it is intentionally, or at least by design, supply limited. The second force is the productivity of the contemplated project. As a result, a project that, absent a constraint on the supply of money, has an otherwise acceptable expected rate of return versus other similar, competing projects, may not be executed if capital scarcity drives the cost of capital above the level of interest that would reflect only the risk of the project.

One place in particular where this is very visible is the real estate private lending space where rates for 1-year loans at 70% loan-to-value have gyrated between 12% and 6%. The risk/reward of the actual investments was relatively unchanged, but the quantity of money that individuals wanted to invest in the asset class was changing dramatically, causing interest rates to gyrate. The fixed quantity of money in an exogenous system makes concepts like crowding out effect, velocity of money, and other similar—often monetarist and Austrian—concepts that implicitly assume that money is supply constrained applicable. When the public sector (or any sector) borrows, there will be less money left for the private (or other) sector(s) to borrow. The aggregate capacity to lend has a hard limit and is unrelated to the viability of the projects that are in need of funding, leading to possible funding shortages.

The assumption that the quantity of money is fixed and that money circulates between savers and investors, and between revenue and expenses—in other words between sources and users of money—however is never mentioned, leaving the reader open to think that these concepts are universally true when this, in fact, is not the case.

3.1.4 Aggregate assets in an exogenous system

From a consolidated balance sheet point of view, any counterparty-free, exogenous currency exists in addition to other assets on the combined balance sheet of society.

There do not appear to be non-arbitrary mechanisms to modulate the quantity of an exogenous money like gold (or Bitcoin) to keep pace with demand for that currency. How should decisions to change the quantity of currency in existence be made? If society deemed it necessary to increase the quantity of currency, how would this currency be distributed? Should it go pro-rata or according to some schedule, and what would that distribution scheme look like? And in the case of the need for a reduced quantity of currency, how would it be decided who should give up some of their money and who would take custody of it?

Exogenous money proponents are highly likely to oppose such mechanisms because they would take away an aspect of exogenous money which is highly prized by some—the inability of bankers and politicians to alter the quantity of money available to society. The consequence is that it makes people hostages to money rather than have money enable economic activity.

3.1.5 Banking in an exogenous monetary system

Electronic copy available at: https://ssrn.com/abstract=4637150
Banking in exogenous systems is a largely Savings = Investment and custodial activity. Banks can earn money by acting as an agent in lending money to a third party and earn fees and an interest rate spread. Or they can act in a principal capacity, take customers’ deposits, invest them, and earn the difference between the yield on the investment and the cost of capital. When a bank acts as a custodian for a customer, it does not use the assets entrusted to generate income. Rather, the bank charges a fee for the service, which is similar to a safety deposit box service. You pay a fee, and the bank will safeguard whatever you put into the box. Banks can run into trouble for two reasons. First, they can run out of funds to operate and therefore stop performing their day-to-day functions. Second, their equity buffer can fall below a required level set by a regulator. In an exogenous system, a bank’s equity buffer can drop because of lending and investment losses. Then, although customers’ deposits are not endangered, the bank is not allowed to perform normal day-to-day functions. The money still exists, but it is not accessible to the owners. No matter how bankrupt a bank is, the money itself does not get destroyed. It simply moves to different owners. The question of repayment becomes one of priority: which party should get paid first and how much.

3.1.6 Central banking in an exogenous monetary system
The concept of a central bank in an exogenous monetary system makes little sense. Consider the main roles of a central bank: stimulating aggregate demand, decreasing aggregate demand, targeting interest rates, and acting as a transaction settlement agent. Nobody—not even the central bank—can create exogenous currency, so the central bank can’t stimulate demand (a la Keynes) by creating money through asset purchases. Similarly, the central bank cannot decrease aggregate demand because it has only a finite quantity of securities to sell and selling them would drain reserves from the system. If the central bank were funded with money raised, for example, through taxation, there would be a real, hard limit to the value of the assets it could buy, which would make it a perfect target for speculators. The finite resources of this central bank would make targeting interest rates impossible because it has a finite quantity of money that it can use to purchase assets in order to inject money into the economy.

3.2 Endogenous Money
The second type of money, which is created through the issuance of credit, is endogenous. It is created inside the economic system and is directly connected to the value of assets and the productive capacity of the economy. James Tobin wrote about it as far back as 1963 (Tobin, 1963), so the mechanism is not new or unknown. Even before Tobin’s paper, John Maynard Keynes in his 1930 Treatise on Money said that banks have two ways to create deposits.

Such a bank creates claims against itself for the delivery of money, i.e. what, hereafter, we shall call Deposits, in two ways. In the first place it creates them in favour of individual depositors against value received in the shape either of cash or of an order [i.e. a cheque] authorising the transfer of a deposit in some bank (either another bank or itself). A member of the public comes along with cash in his pocket or with a cheque drawn on a bank, which he hands in on the understanding that he is entitled in
return to a claim to cash (i.e. a deposit) which he can either exercise or transfer to someone else.

But there is a second way in which a bank may create a claim against itself. It may itself purchase assets, i.e. add to its investments, and pay for them, in the first instance at least, by establishing a claim against itself. Or the bank may create a claim against itself in favour of a borrower, in return for his promise of subsequent reimbursement; i.e. it may make loans or advances. (Keynes, 1930)

In 2014, the Bank of England published a pamphlet that neatly reiterated how money is created in the modern banking system.

In the modern economy, most money takes the form of deposits. But how those bank deposits are created is often misunderstood: the principal way is through commercial banks making loans. Whenever a bank makes a loan, it simultaneously creates a matching deposit in the borrower’s bank account, thereby creating new money. (Bank of England, 2014)

This description is good but slightly incomplete. The form of money creation described happens inside the economy and is a function of demand for money combined with the value of the assets pledged as collateral for loans. As such, it is endogenous because no outside force determines what the quantity of money in the economy is. Also, as we will see, the money that gets created this way is not in addition to other assets the way, for example, gold is in a strictly exogenous system.

The quantity of money in the system is determined by the economic participants, in particular the holders of assets, whether that asset is tangible or intellectual. How much money exists in a purely endogenous system, then, is a function of economic participants’ liquidity preferences and is limited by the aggregate value of assets and the maximum advance rate permitted against the pledged assets.

Under the theoretically pure form of this system, for every dollar created there must be a corresponding liability.

3.2.1 The Value of Endogenous money

Endogenous money derives its value from two sources. The first source is its value in exchange, its acceptance value, similar to exogenous money. Second, as deposits are created when a bank accepts collateral for a loan and therefore are the flip side of debt, it is directly tied to the value of the collateral that was posted when the debt and deposit were created. The pledged assets cannot be sold without destroying the same quantity of money that was created when the loan was made. This, we will see, makes all the difference.

3.2.2 Two different concepts of loans

Before we get to the mechanism of creating money through lending, we must take a closer look at the different forms of lending. The word “loan” is often used to refer to two separate and distinct concepts.
In exogenous systems a loan refers to the temporary transfer of money between two parties. However, in an endogenous monetary system, a loan refers to swapping obligations with money creation as a result (Sigurjónsson, 2015). The difference is that in exogenous lending, the quantity of money in the system remains constant, but in an endogenous system the quantity of money increases even though the net worth of society remains constant. Using the word “loan” without reference to the system in which the loan is made is almost certain to lead to incorrect conclusions.

If banks were to lend how the non-bank sector, for example private equity firms or investment banking firms, lends and how loans in an exogenous system work, the following would happen:

- A potential borrower goes to a bank with an asset—say, a house—worth $100,000.
- The bank would put a lien on the house so the house can’t be sold without first repaying the debt and so the rest of the world knows that there is a loan on this particular property.
- The bank would then give $80,000 (not the full $100,000 so that there is a sufficient equity buffer to incentivize the borrower to see the project through to completion) of its own, already existing money to the borrower to be repaid at a later date. This money could come from the bank’s equity capital, or it could be other existing money that the bank borrowed.

Once the bank has lent all the money it has available, it would have to close shop temporarily until it either raised more equity or debt capital or wait until loans it made are paid off.

Instead, in an endogenous system—and how bank lending is widely accepted to work in reality—the following happens:

- The bank takes a house worth $100,000 as collateral and puts a lien on the house.
- The bank simultaneously records an asset on the books for $80,000, the IOU that is signed (the loan), and an offsetting $80,000 deposit in your checking account as a liability to the bank.

That deposit comes out of nowhere—it is created ex nihilo. It is literally just a bookkeeping entry (debit an asset account for the $80,000 IOU from the customer and credit the customer’s checking account for $80,000), and the bank’s equity account is untouched.

What makes the Bank of England’s pamphlet, as well as other people’s description of this process, slightly, but materially, incomplete is that it doesn’t alert you to the fact that assets are liened or pledged; they are taken out of society’s unencumbered collateral pool. The lien ensures that the asset cannot be pledged a second time. Although the customer legally still owns the house, their net worth is reduced by the value of the loan and increased by the value of the deposit. The net change in the customer’s net worth therefore is zero. The change in the bank’s net worth is also zero. It has the IOU as an asset and the customer’s checking account for $80,000 as a liability.

The lien prevents one asset from being used as collateral for multiple loans, and therefore prevents excess deposit creation, and it helps keep the societal aggregate balance sheet in check, at least with respect to lending against real assets.

In effect, the bank did two things. It turned a relatively illiquid asset into an even more illiquid asset by putting a lien on it. At the same time, the bank created a very liquid deposit that can be spent on anything, anytime. In aggregate, not much changed. Although the bank created $80,000 out of nothing, it took $80,000 of collateral out of society’s unencumbered collateral pool, so in total no value was created or destroyed.
Although from a value, or net worth, point of view nothing happened, something important did happen. Something of value, but relatively illiquid in nature, was liquified. The fact that our hypothetical homeowner was willing to pay various fees and expenses, jump through hoops, and fill out lots of paperwork indicates that the liquification was valuable.

In a modern monetary system, only entities with banking licenses are allowed to liquify assets, that is, convert illiquid IOUs into deposits (which are highly liquid IOUs of the bank). Or, to phrase it differently, in a purely endogenous system banks create money through creating and swapping IOUs, and commercial bank lending is an exercise in asset liquification resulting in money creation.

The money created in this type of transaction is a liability of the bank, but that liability is offset by the pledged asset. This type of money creation, to a fair degree, has an aspect of liquidity arbitrage. It therefore makes sense that the rents charged on the illiquid leg of the transaction are more than the rents received on the liquid leg of the transaction; there is a positive spread and an upsloping yield curve. 4

3.2.3 Interest
On a fundamental level, the interest charged on a loan created in an endogenous system can be as low as the cost of liquification. There is no actual cost of capital because the capital is created in the liquification action. However, there is no hard upper limit to the cost of capital (hence no limit to the bank’s profitability) aside from the limit set by the productivity of the activity for which the capital is used. The maximum rate of interest that makes economic sense is the rate of return on the project for which the money is used minus transaction costs.

Interest on money in this system is not an issue as long as the economic activity for which the money is created is more productive than the cost of the money so that the economic activity in aggregate is additive to the combined balance sheet of society. The reason the absolute level of the economically viable interest rate is not important is that the required liquidity to make the interest payments can come from borrowing against, or liquefying, the improved asset, which now has much more value.

For example, if an entrepreneur buys a run-down house for $100,000, puts in $50,000 in construction costs, pays $10,000 in interest, and sells the house for $200,000 (i.e., its value has increased to $200,000), society is still better off even after taking into account the interest that was paid, regardless of the interest rate. The total quantity of currency in existence after the transaction is greater than before the rehab project, yet there are disproportionally more productive assets than there is currency. So, the increase in the value of the project plus the cost of financing is less than the increase in the money stock. In lending this way, the ratio of real

---

4 Banks are allowed to create deposits, which are obligations of a private entity, yet these deposits are gladly accepted as means for payment of taxes, which are obligations to the federal government, which obviously is a public entity. Even state governments, which can’t print money, accept privately issued obligations (deposits) as payment for various taxes and fees. The FDIC exists exactly to mitigate this private creditor risk, at least for certain private parties and to a certain extent. Banks are not allowed to print currency, but they are allowed to purchase currency from the treasury in a one-to-one ratio, thereby swapping private obligations for public, taxpayer backed obligations, gratis. It is as if one can take a corporate bond and swap it for a treasury bond at an equal yield yet at zero cost. There is an argument the FDIC insurance limit is counterproductive because sophisticated deposit holders’ balances are often large, and they therefore keep a sharp eye on the condition of the bank that holds their deposit. They are also more likely to move their deposit in case trouble is ahead, thereby increasing volatility in a bank’s deposit base, which is the exact opposite of one of the intended effects that FDIC insurance is supposed to have.
assets to the quantity of money will increase. In this context, an increasing money stock over time is not a bad thing *per se*.

Quantity of money $\uparrow$, quantity of assets $\uparrow\uparrow$.

In a purely endogenous system, interest rates have a lower bound, which is set by the processing costs of the liquification process and the operational costs of keeping the banking system going. The quantity of money that can be borrowed, or created, in an endogenous system has an upper limit equal to the quantity of eligible collateral that exists in society minus a margin of safety. The margin of safety largely acts as an incentive for the borrower to repay the loan because the borrower is likely to lose part of the whole margin of safety in case of a default.

There are two big caveats. First, collateral can only be pledged once. Second, the proceeds of that loan cannot be used for another liquification. This is why the placement of liens or similar notices – for example UCCs - are important. They don’t explicitly show up on anyone’s balance sheet or income statement, so they are not obvious, yet they are the key to preventing potential multiple liquification of assets.

### 3.2.4 Aggregate assets in an endogenous system

On the consolidated balance sheet of a purely endogenous system, the deposits which have been created through the liquification of assets are in place of those assets, not in addition to them, as is the case in an exogenous system. As economic participants liquify assets, the quantity of money in the banking system increases and the quantity of unencumbered collateral decreases, and vice versa. The ability to create and destroy deposits at will in an endogenous system—within the confines of available collateral—creates an elastic deposit supply curve. Assuming “loans (deposits) are created responsibly against collateral that is correctly valued, in such a system one could never have more deposits than there are assets. And when liquidity is no longer needed, negative carry in the form of required interest payments creates an incentive to make the deposits disappear by repayment of the loan that had created the deposit in the first place.

### 3.2.5 Banking in the endogenous world

Banks in endogenous systems are gatekeepers of liquidity by rendering judgements on the value of assets. They don’t have to have equity because they don’t take any risk; they don’t take deposits from one party and lend it to another party. There is risk of non-payment, but that risk is borne by society at large, not the deposit-issuing bank. As long as the bank has enough income to finance daily operations, there won’t be any issues. Money is created through asset swaps and destroyed by repayment. Banks also facilitate the transfer of deposits between owners. The bank earns money by charging a rate of interest on deposits that it has created and in the form of loan fees. It also charges fees for various services it performs. If a loan is not repaid, the bank simply tears up the note that the borrower gave as collateral. What happens to the deposit that was created will be discussed later.

### 3.2.6 Central bank money creation

Central banks can also create a type of money called “reserves,” but the process is a bit different from how commercial banks create deposits. A quick review of bank reserves will help set the stage.
Bank reserves consist of a commercial bank’s deposits at the central bank plus cash in the commercial bank’s vault. When the central bank purchases an asset it increases the commercial bank’s reserve balance—its deposit at the central bank—not the bank’s checking account at its own (or another) bank. As such, reserves are not directly spendable assets.

Central banks create money through the following:

- The central bank communicates to commercial banks that it intends to purchase $X of certain government bonds.
- The commercial banks then purchase these bonds from their customers and create deposits in the customers’ accounts.
- Commercial banks then sell the bonds to the central bank and receive a credit to their reserve accounts in return.

Note that in both transactions, bonds are sold and not pledged. There is no loan creation and hence no interest involved in the transaction, just an exchange of a bond for a deposit and then for a reserve credit.

The cost of money, the negative carry, through deposit creation in an endogenous system, then, is lowest when the central bank purchases debt to create reserves and thereby indirectly creates bank deposits. Interest earned on the bond held by the Fed is repaid to the issuer, the U.S. Treasury, making for cheap reserve creation and, therefore by extension, cheap deposit creation. Regardless of the interest rate on the bonds that the Fed purchases, it will remit excess interest earnings minus operating expenses at the end of the year back to the Treasury, thus reducing the cost of interest to the taxpayer to the operational costs of the Fed.

As a consequence, if someone who currently has no dollars and wants to transact in dollars pledges an asset to create those dollars, the aggregate debt will increase.

Aggregate debt, in other words, does not necessarily imply financial stress. Rather, it can indicate liquidity needs of the financial system. Liquidity in the form of deposits in an endogenous monetary system show up as assets (mainly bonds) on the central bank’s balance sheet and as total debt on the national balance sheet. Liquidity (deposits) and debt then are two sides of the same coin, and the notion that the national debt level is indicative of federal financial mismanagement is not necessarily correct.

It stands to reason that governments should issue separate bonds to provide financing for operations to create reserves, and, therefore, support the ability of the commercial banking system to create money. By commingling the bonds—as is currently the practice—government debt figures don’t necessarily reflect financing needs but rather are a mix of the two. Mixing the two uses also overstates the interest cost to society. Bonds owned by the central bank have essentially a net interest cost of close to zero.

### 3.2.7 Multiple liquifications of Assets

In an endogenous system there are two ways by which assets can be legally liquified more than once. Despite its legality, liquifying real assets more than once causes an imbalance between aggregate deposits and real assets net of encumbrances.

#### 3.2.7.1 Commercial Banks and a Chain of Multiple Liquifications

As the initial transaction, a commercial bank can take an asset—for example, a house—valued at $100,000 as collateral and then create a deposit at 80% of the asset’s value. Let’s assume that the bank then takes that loan, that IOU of $80,000, and through securitization turns it into a bond with a value of $80,000.
In a second transaction, a bank—and this could be the same bank that created the first deposit of $80,000 or a different bank—accepts the bond as collateral for a second loan and creates a second deposit also at 80%. There is still only one real asset, the $100,000 house, but a total of $144,000 ($80,000+$64,000) worth of deposits have been created. The owner of the new IOU can securitize it into yet another bond, which can be pledged as collateral for yet another loan and deposit. This results in a mini fractional-reserve lending system that has the potential to create deposits that are multiples of the value of the actual real-world pledged asset. From a balance sheet point of view, everything is in balance. But the end result is that the sum of the value of the deposits can be vastly in excess of the value of the pledged real asset, the house in our example.

The equity margins required by the various steps in the securitization chain set the maximum multiple of freshly created deposits compared to the original pledged collateral. The equity margins, or leverage ratios, that are used are determined by the commercial banks executing the transactions and creating the deposits, not the Federal Reserve. There is a backstop, however. Banks are subject to overall equity-to-deposit requirements. These requirements, however, are in the low single digits, meaning that banks can easily create 20 times or more deposits than they have in real, as opposed to financial, collateral.

This chain works because although the IOU is directly connected to the initial deposit that was created, it appears that the bank owns the bond free and clear and that there is no offsetting obligation against it. The relationship between the deposit that was created and the IOU is one-way. The IOU can be wiped out by returning the deposit, but the IOU can also be freely transferred anytime and swapped for a bond, which acts as a canister, of the same value. A bank has a choice between keeping the loan it made on the books or selling it. For example, a bank makes a mortgage loan. It now has a choice whether to sell that loan, for example to FNMA, and then purchase a mortgage-backed security (MBS), or just keep the loan.

In theory, the money that is created in these transactions will be destroyed as loans are paid off. But in reality, these loans become a permanent part of the financial landscape. That is because loans are being repaid with new loans, so the total balance does not decrease, making something that is in theory temporary, effectively permanent.

The consequences if there is a problem with the collateral at the base of the chain – say the house burns down – can be significant because both the parties that made the various loans (and created money in the process) get hurt, but, as we will see, also society at large.

3.2.7.2 Central Bank Creation of Money through Multiple Liquifications

Central banks can create money, in the form of reserves, in a fashion similar to how a commercial bank creates deposits, by locking up financial assets and creating offsetting bank reserves. The mechanism of the process is materially different though. When a commercial bank makes a loan and creates a deposit, it places a lien on an asset but does not legally own the asset. In contrast, a central bank actually purchases the asset and uses newly created reserves to pay for it. The central bank now owns the security.

Several differences exist between commercial bank money creation and central bank money creation. One difference is that the central bank can purchase assets at 100% of value on the dollar and does not require a margin of safety. Another difference is that central banks buy only limited types of financial assets but not real assets.
Central banks also don’t re-lend or securitize the assets they have purchased, which stops the potential chain of securitizing already securitized assets. However, given that central banks purchase only financial assets, they by definition purchase assets based on underlying collateral that has already been liquefied once.

Some other details are different too. When a central bank purchases an asset—say, a bond—from a bank, it pays for it by adding to the balance in the bank’s reserve account at the central bank. It does not create money that the bank can use in a direct manner. However, there is a second step in the process. The bonds that the bank sells to the central bank were purchased from account holders at the bank. Those account holders now have a balance in their deposit account which they can spend. So, although technically the statement that QE does not create money is true in the first order, in the necessary second step—customers selling bonds to the bank—deposits are undoubtedly created.

Where the commercial bank made loans with a margin of safety, this double liquification by the central bank upends the initial safety margin, and now you have more money than there are assets.

The common thread between these multiple liquefications is that the assets which are pledged are not real assets but financial assets. As long as lending against, or creating deposits against, financial assets is allowed, it is inevitable that deposits will grow at a faster rate than the stock of real assets.

4. **Viewing the economic landscape through the lenses of exogenous and endogenous monetary systems**

In general, economic theories which make sense in an exogenous system are unlikely to offer explanatory value in an endogenous system. That is not to say that there are no limitations to endogenous money creation or liquification. But the limitations that exist are not related to the quantity of money that is available. Rather they are related to the quantity of assets that are available to be pledged.

For example, in an endogenous monetary system, there is no crowding out effect because money is created on demand and there is no particular (fixed) quantity of money that borrowers compete for. Demand for money in one part of the economy does not impact availability of collateral to create money in a different part of the economy.

The quantity theory of money is also irrelevant because money is created and destroyed on demand. There is no fixed quantity of money that changes hands at some velocity. The rate at which money is created and destroyed, for example overnight repos, has no bearing on price levels or anything else in the real economy.

As endogenous money is the flipside of debt, money created must have duration equal to the duration of the debt that was created when an asset was liquified. This has significant implications. Generally, there is an assumption that all money is created equal and has an infinite life, in other words, that money is exogenous. When the Fed buys 2-year Treasuries instead of 10s, it has implications beyond what part of the yield curve is impacted. For example, if the recipient of the proceeds of the sale to the Fed of a 2-year treasury takes the deposit and uses that money to make a 30-year loan, it will cause issues once the 2-year treasury bill reaches maturity and needs to be repaid.
4.1 Deposit creation and optionality
Endogenous monetary systems have the potential to have too many assets liquified, that is, have more deposits in the system than is optimal. If the cost of liquefying, and the ongoing negative carry, is low enough, liquification, and thereby creating money, is akin to purchasing an option. Deposits give the owner the option to purchase other assets almost immediately rather than have to first find a buyer for the asset that the borrower currently owns. In a way, creating a deposit by pledging an asset creates two choices for the borrower. The first choice is to keep the deposit and, essentially, sell the pledged asset by not repaying the loan or returning the deposit—defaulting. The second choice is to use the deposit to purchase another asset. The option premium consists of two parts: first, the interest and fees that the borrower pays for the loan; second, the effective negative carry on the deposit. Even if the deposit earns interest, the risk that it is exposed to is by definition greater than the return it is receiving.
If this optionality is deemed undesirable, a way to modulate this process is by increasing the negative carry on deposits.

5. Exogenous Monetary systems Past and Present
Exogenous monetary systems have been around as long as the written word, although probably never in their pure form. It is impossible for the exogenous money stock to perfectly match demand, so various forms of credit are used to bridge the gap. That credit could have been provided by businesses rather than the government, but the gap between demand for money and supply somehow had to be bridged.

5.1 The End of the Gold Standard
Discoveries of gold in faraway places that were added to the monetary stock were inflationary and unrelated to the needs of the local economy. For example, by receiving payment in gold for goods sold during World War II, the U.S. ended up with a disproportionately large share of the global gold stock, which caused several issues. Even with that gold hoard, however, the quantity of gold at the prevailing price was simply insufficient to collateralize an increasing number of dollars at a constant exchange rate (to gold). The dollar price of gold had to be increased again and again until the world lost confidence and the connection between gold and money was formally suspended, although not cut.5

5.2 Cryptocurrencies: Bitcoin
Cryptocurrency appears to be a new development in the monetary landscape and a step beyond traditional banking as we know it. However, when you take a closer look, it is clear that this is not the case. In fact, cryptocurrency is, in a way, full circle back to the days of gold and silver.

---

5 As an aside, note that historically gold has been extracted at a rate that is greater than the rate of economic growth so the argument that a gold based monetary system would not be inflationary sounds hollow.
Between 1940 and 2018 the global gold stock has increased by 3.66%/yr while the global population has grown by 1.75%/yr. The quantity of gold per capita went from 12.74g/pp to 18.24g/pp between 1940 and 2018. These numbers do not exactly give an impression of stability.
It just makes no sense to tie the quantity of a currency in existence to random physical discoveries and technological progress in the extraction of a metal which plays a minor role in industrial, productive applications. Gold as the basis of a monetary system simply doesn’t work which is why the world has moved away from it. As Ben Bernanke said when he was asked about the role of gold in the monetary system: “well, its - you know, it’s an asset.”5
When Bitcoin was designed, its developer decided that only 21 million Bitcoins would ever exist. Regardless of the actual number of currency units, a static quantity of currency cannot meet the needs of a dynamic economy. One argument against this criticism is that the limit doesn’t matter because Bitcoins can be divided into smaller units. Gold and currency like dollars can be divided too. That does not change the total quantity of currency units available. A solution to the problem of Bitcoin’s hard upper limit that has been floated is that other, similar currencies can be created, and competition will allow the best one to win. We can see this process in action by looking at the large collection of “shitcoins” that have been created for which there is no demand. In the current environment, anyone can determine which money is issued and how much of it can exist.

Bitcoin has all of the downsides of exogenous currencies and some specific attributes that create additional problems. Yet, it has one positive attribute from an extremely limited, exogenous money point of view: its one-sidedness.

Crypto is trying to be so secure that in effect it has become insecure. It implements security through a 2-key system—a public key and a private key. This security is complex and vulnerable to mistakes which can cause loss of access to the coin. If you lose your bank card and can’t withdraw money or make a payment, all you need to do is call or go to your bank, prove your identity, and get a new card after a few days. This process, depending on the bank, can be easier or harder. But unless you don’t have some basic bits of information, it is highly unlikely that you won’t be able to access your account. This verification method uses the power of many independent datapoints. Crypto works differently. Like gold, and even fiat currencies, crypto has two categories of storage. The first one is that the individual owner holds the asset in a digital wallet. In the case of Bitcoin, if you lose your private key or make too many mistakes entering it, you can permanently lose access to your wallet. The money is not just lost to you but also to the system. Even if you store your private key safely in a cold wallet, there is a greater-than-zero chance that the wallet could be lost or stolen.

Crypto storage and safekeeping are different from money in a bank account. You can keep crypto on a crypto exchange, but exchanges have no security standards, have minimal self-regulation, and get hacked regularly. There is no insurance, either private or federal, for those losses. There is no regulatory framework and no audit requirements. In the case of theft, the total amount of Bitcoin in accessible existence does not change but will change hands from victim to perpetrator.

Given the inevitable losses of cryptocurrency coins that occur due to security mechanisms, the total quantity of coins in circulation will decrease over time, reducing their attractiveness as a form of payment. These losses cause liquidity to decrease, which can easily set in motion a self-reenforcing cycle of price decreases, which then reduces the desirability of owning the cryptocurrency. In the long run, Bitcoin could end up having no value.

Because Bitcoin is on a blockchain, it is possible to trace exactly which wallet a coin that you received in a transaction came from. In other words, transactions are traceable back to inception of the coin. Bitcoin is anonymous only to the degree that the holder of the wallet is anonymous. In contrast to a transaction with a $20 bill, all Bitcoin transactions are registered on one publicly available location. The $20 bill you have can change hands many times, and as long as it doesn’t get converted to a deposit at a bank, it is effectively untraceable.

As we saw in the descriptions of both exogenous and endogenous moneys, moneys are valuable at least because of their value in exchange. You can use a dollar bill or a deposit to make
payments to virtually anybody, and they will accept it. The value in exchange for Bitcoin is extremely limited because the number of participants in the Bitcoin ecosystem are limited. Try to pay for your next morning coffee or next tank of gas with Bitcoin. As the number of possible exchanges of goods and services with Bitcoin are extremely limited, the value of Bitcoin itself is uncertain. Bitcoin suffers from a lack of usability because of a lack of willing economic participants. Widespread adoption is limited due to low acceptance value.\(^6\)

Paul Krugman’s question “What problem is Bitcoin trying to solve?” is relevant and remains unanswered (Krugman, 2018). A clear, convincing case for widespread adoption of Bitcoin has not yet been articulated. It is difficult to see a future where Bitcoin will be widely adopted as a means of payment unless some very large sponsors like banks come to the rescue or if, for example, the government were to start accepting Bitcoin as a means of payment. Absent such sponsorship it is likely that Bitcoin is a curiosity that will fade into nothingness.

6. Losses
Losses tend to get little attention because they appear to be self-explanatory. You used to have something, and now you no longer do without selling it or being compensated for it. On closer examination, however, there are different types of losses, economic losses and monetary losses.

6.1 Economic Losses
Economic loss refers to the loss of a real asset, for example when a house burns down. Real, productive value that took inputs to create has been removed from the economy, causing a permanent loss of productivity. Value was lost, but money may or may not have been lost. Economic losses are often physical in nature but not always. For example, society may choose to no longer use a particular type of asset, such as a nuclear power plant that is taken offline before the end of its useful life. The physical asset will no longer generate an output that is valued by society. This type of loss can occur in systems that use either exogenous or endogenous money. Losses of financial assets are different. Financial assets are either claims on real assets—for example stocks and bonds—or are risk transfer instruments—for example, options, futures, and various swaps. These risk transfer instruments all net to zero and in aggregate generate no economic losses.

Losses in financial assets that are claims on real assets can occur in two ways. First, the underlying asset can become impaired, and that impairment is reflected in the financial asset, which effectively functions as a token. If the financial asset loses value because of an issue with the underlying asset, the loss is economic. However, if the underlying asset is not impaired yet the value of the financial asset is impaired, it is more like a monetary loss; value was transferred but not destroyed. This is what happens when a stock trades below book value. Second, the counterparty that holds the other side of a risk transferring asset is no longer able or willing to perform. This is not an economic loss to the system as a whole. No productive asset was impaired although monetary value changed hands.

In current economic systems, it is very hard to lose a financial asset. The loss of most financial assets by themselves is not an economic loss on a macro level because there are two parties—the issuer in a primary transaction, or the seller in a secondary market transaction, and the purchaser.

---

\(^6\) David Graeber in his book *Debt: The First 5,000 Years* talks about how colonizing countries frequently forced new colonies to use the invader’s currency by demanding payment of taxes in said currency. This forced adoption of the new currency from the top. But as Bitcoin is not issued by a government that can force adoption, this approach can’t work.
Even if you lose a bank deposit, you have years and years to recover it. Even if you don’t recover it and experience a personal loss, on a macro level nothing changes. All that happens is that ownership of the deposit transfers from you to (eventually) the government. One exception are bearer instruments, like bearer bonds. If your bearer bond gets stolen you may have a loss, but the thief has a gain, so in aggregate nothing has changed. If the bearer bond is destroyed, the holder has a recourse-less loss. But the bond issuer has a gain because it becomes impossible to repay the bond at maturity.

6.2 Monetary Losses
The second type of losses are losses of money itself (not losses of assets expressed in monetary units). These losses can be separated into macro level loss and micro level loss. In a macro monetary loss, money actually disappears from the system. For example, a gold coin may get worn down and its value becomes diminished, or a private key to a Bitcoin may be lost. This type of loss impacts both the party that incurs the loss and society overall because the currency unit itself has permanently disappeared from the available currency stock and cannot be recovered. When Pablo Escobar buried a billion dollars in the jungle and it turned into worm food, a monetary loss occurred. The value represented in the creation of those bills was enormous, even if the cost to produce the dollar bills was quite limited. However, in the case of a pure monetary loss there is no loss of productive assets in society. The micro sub-type of monetary loss occurs when one party loses money, but it is found or received by another party. For example, one person drops a coin and somebody else picks it up. Money was lost by an individual but not to the system. Modern accounting tracks balance sheet and income statement items, but it is not designed to explicitly track money creation and destruction. When Luca Pacioli figured out double entry bookkeeping in 1494, money was exogenous, so it made sense that tracking money creation and destruction didn’t enter the picture. However, losses are not just important for tracking profit and loss. They also play an important role from a monetary viewpoint. Real economic losses and monetary losses are not necessarily related. Often, they are, but the connection is far from 100%. An added layer of complexity is that the losses in the two types of monetary systems have different consequences.

6.3 Losses in an exogenous system
Economic losses in exogenous systems are real. In an exogenous system, the quantity of money is constant, so economic losses don’t impact the monetary side of the economy. When a lender makes a loan in an exogenous system the lender is at risk of monetary loss, but the economic system as a whole is not. If a lender loses money, it is because the money was transferred to another party and not returned. However, in total no money was created or destroyed. In an extreme case, if there is wholesale destruction—for example a natural disaster—the quantity of money remains constant while the quantity of assets could shrink considerably. You now have an unchanged quantity of money chasing fewer goods. But no matter what happens, the quantity of money in an exogenous system remains the same.

\[
\text{Quantity of money} =, \text{Quantity of assets} \downarrow.
\]

6.4 Credit losses in exogenous monetary systems
Let’s assume somebody borrows existing, exogenous money and pledges a house as collateral. The money is transferred from lender to borrower. Unfortunately, the house burns down, and the borrower can’t repay the loan. The total quantity of money in the system remains constant, but the quantity of assets has decreased so the ratio of money to real assets has changed. Quantity of money = Quantity of assets ↓.

Now let’s look at a case where a bank is defrauded. No collateral was pledged, but a signature loan was made and not repaid. From a strict aggregate monetary view nothing happened. No assets were destroyed. Money was transferred between two parties, but no money was created or destroyed. Quantity of money = Quantity of assets =

In an exogenous system, credit losses never impact the quantity of money that exists, but they can impact the value of assets. When looking at the ratio of money to assets, in an exogenous system only the asset side of the equation has an impact. In an exogenous system, credit losses run solely through the income statement, ultimately hitting the shareholder’s equity account on the balance sheet. Quantity of money = quantity of assets ↓.

6.5 Credit losses in endogenous monetary systems
Under the credit creation theory of banking, in a perfect world the quantity of money that exists should follow the value of the collective assets very closely. Assets are liened, money is created, over time those loans are repaid, and the money disappears.

In reality, however, banks in their role as deposit creators, have loan losses. What is the impact of these loan losses?

When a bank that created money through making a loan has a credit loss there will be, in contrast to exogenous systems, an impact on both the monetary side and the asset side.

7. The Problem of Stranded money
In a 100% credit creation world, after all loans have been repaid no money should be left. But in reality, some loans fail and don’t get repaid in full. Absent a mechanism to mop up the excess money created by failures to repay loans, money will be left without pledged assets as an offset. This causes an imbalance between the assets and liabilities on a bank’s balance sheet. An entry to the P&L account fixes the accounting side, but that does not cause deposits to disappear from the financial system. The money is stranded. Let’s look at an example.

Suppose we have a bank that has no equity, just a banking license that allows it to create deposits by taking collateral in exchange.

The bank creates $80,000 of money against a $100,000 house pledged as collateral. The borrower uses all the money to pay bills. His ending bank balance is zero because the deposit was transferred to other, unrelated accounts.

Unfortunately, the house burns down with the homeowner in it, so any personal guarantee the borrower may have signed is useless. The asset has disappeared, the bank has no recourse, and now the bank will realize a loss. The asset, the note secured by the house, is lost. But the liability, the deposit it created, still exists. The bank can’t take back the deposit because it has been transferred to an unrelated party.

The bank takes ownership of the property (which is only land at this point) and sells it for $20,000. The loss to the bank is the $80,000 loan amount minus the $20,000 proceeds of the land sale for a net loss of $60,000.
The bank will take the loss on its income statement, but that doesn’t change the quantity of money in existence because taking a loss doesn’t reduce the amount of money in an unrelated party’s checking account. Loan losses reduce shareholders’ equity on the balance sheet but don’t “un-create” money that was created through endogenous lending.

Before the loan was made there was a pre-existing money stock of $X and collateral with a value of $100,000. After the loan was made, the money stock became $X+$80,000, and the free collateral pool shrank by $80,000 (the lien is the value of the note, not the value of the collateral). Once the fire occurred, the money stock was still $X+$80,000, but the collateral value became $20,000 (the land value). $60,000 of excess currency creation versus the asset base occurred at the same time the asset disappeared. So, the difference is $60 = 100 in excess money versus assets. Quantity of money↑, quantity of assets ↓.

I refer to this endogenous money, which exists in excess of loans, as stranded money. It used to be the flipside of an asset, but it has become one-sided. In effect, this stranded money is now exogenous.

The value of the land in our example is not relevant from a money creation point of view. If the bank sells the land, it will receive dollars that were either already in existence or that were created by the bank financing the purchase transaction. In either case, the number of dollars in existence will not be reduced. There is no way for the bank to extinguish the dollars that were created because the debt is gone. Although in the perfect endogenous world, money and debt are two sides of the same coin, losses of pledged collateral break the link between money and collateral. When that link is broken, money moves from the endogenous system to the exogenous system, and there is no way—at least without external intervention—to remove money from the exogenous system.

In the case where a commercial bank made a signature loan—that is, no collateral with intrinsic value (that can be sold) was pledged—and the loan defaulted, the total quantity of assets in society remains unchanged, but the quantity of money increased. Quantity of money↑, quantity of assets =.

Endogenous loan losses impact both the issuing bank and society at large. The bank is hurt by the loss of an asset—the IOU that was pledged and will show up as a loss on the income statement—but society now has more currency units in circulation relative to its assets. In the case where there are multiple iterations of securitizations, the excess deposits that are created can be multiples of the lost asset.

It would make sense then that society, which bears the costs of loans gone bad through an increasing quantity of money, plays a part in setting the parameters on lending rather than just the particular bank that is executing the loan.

Stranded money is problematic. Once cash balances exceed transactional demand, currency users will start to invest it in risky assets. Cash is a negative carry asset and that creates an incentive to invest in anything that at least does not cost money during the holding period. Deposits in a pure endogenous system have negative carry caused by two factors. The first factor is expenses incurred during the creation of the loan and deposit, including the costs incurred in analyzing the collateral, recording of liens, and other onetime expenses. Only endogenous money incurs this cost in its creation. The second factor is the ongoing maintenance costs of the deposit—the accounting, statements and general continuing administrative expenses. Both endogenous and exogenous money incur these expenses.
But stranded money has less negative carry in addition to associated holding costs because it is not connected to an interest-bearing loan. So, the minimum investment rate of return to potential contemplated investments is much closer to zero. Therefore, it makes sense that the excess money will be invested until returns are very close to zero because it is better to get something than nothing.

When stranded money builds up, interest rates are repressed by an ever-increasing quantity of zero cost money. As a result, risk is no longer priced correctly. The supply of money simply overwhelms demand, and projects with lower and lower returns are engaged in while the level of risk may remain the same, lowering the risk/reward ratio. This continuous buying pressure puts downwards pressure on yields and arguably puts ongoing upwards pressure on asset prices. If an asset purchased by a central bank disappears, the effect we saw in commercial bank lending is even more pronounced. To continue with the prior example, a bank created money of $80,000 against an asset worth $100,000. Suppose then that the central bank purchased the resulting bond for $80,000, creating in total $160,000 in new money versus $100,000 in collateral. If an asset has been re-securitized, the effect is directionally the same – more deposits than there were assets were created. In the case of asset destruction, the quantity of stranded money can be greater than the value of the asset. If the asset gets destroyed – say the house burns down - society loses a total of $80,000 in assets (the house minus the land value). At the same time, $160,000 in deposits were created. While the commercial bank will have a loss, and the loss that the central bank incurred will be borne by the taxpayer, society at large will have double consequences. The loans cannot be repaid, the newly created money remains in the system, and the asset has disappeared. That obviously changes the ratio of money to assets in the system.

\[
\text{Quantity of money} \uparrow \uparrow, \text{Quantity of assets} \downarrow.
\]

From a P&L perspective, the bank may have taken a loss, reduced its loss reserve, or even taken no loss because of an insurance payment, but that does not impact the quantity of the money stock. Any reductions in the loss incurred come out of the existing stock of money and are simply transferred between checking accounts.

There is an important difference between losses in an exogenous system and an endogenous system. In an exogenous system, the quantity of money in the system remains constant, but the quantity of assets is reduced. In an endogenous system, however, the quantity of money has increased through lending, and the quantity of assets have decreased. So, there is a gap equal to the sum of the lost value of the asset plus the money that was created. As such, pledged asset losses in an endogenous system have a greater impact on the asset / money ratio than an identical loss of pledged assets in an exogenous system.

Banks in the private sector don’t have a way to siphon off the stranded money that was created by bad loans. However, the public sector can remove stranded money from the system. For example, the government can impose a tax on deposits equal to the endogenous loan losses that were incurred by banks—the money that was created but for which collateral no longer exists. That tax revenue can be used to retire debt by paying off government bonds which are held at the Fed, thus reducing the loan side of the collective portfolio. Both money and bonds, which are just the different aspects of a loan, would disappear. Mechanically, the treasury would tax deposits and send the proceeds to the Federal reserve. The Federal reserve could then rip up some of the treasury bonds in its vaults, and both the treasury bonds—which were paid for by freshly created reserves—and the money would cease to exist.
8. The invisibility of stranded money
The reason stranded money is not in the forefront of economics is because it is so hard to spot in the system. Debits equal credits, so everything seems in balance. Our banking system is a hybrid of banks acting as financial intermediaries and banks in their role of credit (money) creators. Even if the functions were separate, or even if the institutions were separate, there is no functional separation between money in the exogenous system and money in the endogenous system because these two systems interact seamlessly. Even for banks themselves the difference between the two systems is not obvious. When a bank makes a loan, it may not be clear if the money came from a customer who made a deposit earlier in the day or if the bank had to obtain a deposit—which very well may be the exact deposit that the bank had created in making the loan—on the interbank market to balance the books after the loan proceeds were transferred to a different bank. Additionally, accounting and reporting focus on balance sheet items and the income statement, not money creation or destruction. On a day-to-day basis, it certainly looks like commercial bank lending is a Savings=Investment business when in reality a part of it is not. We know this because the quantity of deposits in the system increases over time, which would be impossible if our system were based on exogenous money. Reporting on money creation would mean that we first have to admit that money is created by liquefying assets and that the risk of liquification is to a degree socialized. This makes charging interest significantly above processing costs a bit more difficult to swallow. Stranded money is further hidden from sight because historically on average loan losses have been less than GDP growth. During normal economic times, loan losses are small, but the stranded money created by them accumulates. Loan losses are an income statement item and are offset there by the reduction of value of the loan asset, which balances the books. However, that is completely separate from deposits that were created and now cannot be un-created. These types of credit losses are a one-way street and cumulative because the commercial banking system can’t remove the excess money. Over time then more and more money, which started its life out as credit money with associated collateral, builds up in the system without offsetting assets as collateral.

9. The Evidence for Exogenous and Endogenous Monetary Systems
Unfortunately, no data is available to directly support the assertions about exogenous and endogenous monetary systems. Although the FDIC in their Quarterly Banking Profile reporting publishes balance sheets and income statements (FDIC, 2022) including lending losses, it does not break out whether the losses came from existing deposits or if the losses were incurred on deposits that were created in the process of lending. By adding the dollar amounts of annual losses starting with the available data from Q1 1984, we can tally total losses, but that, again, does not tell us which part of the losses came from the lending of existing deposits versus deposits that were created during lending. What the FDIC data does tell us, when combined with the Fed’s balance sheet, is that the quantity of deposits is not constant and overall increases over time. This data supports the assertion that we have – at least in part – an endogenous monetary system. If that weren’t the case, the quantity of deposits plus currency in circulation would be constant. Another implication of stranded money is that collateralized loans that default have a greater impact on the mismatch between assets and money than unsecured loans. This is because both
sides of the collateralized loan transaction are impacted. With an unsecured loan, only the monetary side is impacted, and no collateral is destroyed. From an individual lender’s point of view secured loans are less risky. But from society’s point of view, they are in fact riskier than unsecured loans because losses are socialized to a larger degree. It is no wonder then that more and more money is floating around in the exogenous part of the monetary system. The additional, stranded money looks like excess savings and lowers interest rates below where they otherwise would be.

Over time, more and more money will become stranded due to inevitable loan losses. Unless it is removed from the system, that part of the money stock becomes permanent and has all the characteristics of exogenous money. If stranded money is left unmanaged, even a system that started out as a pure endogenous system will develop the characteristics of an exogenous system. There is a second consequence to the low interest rates caused by stranded money. Low rates reduce the cost of liquefying assets to the point that it is rational to take a loan to have cash just in case a higher yielding investment opportunity comes along. This has the potential of increasing asset price volatility. Money can flow in and out of assets at a moment’s notice. One way to reduce this effect is to increase the negative carry incurred on deposits. Negative carry has a real expense component like transaction and administrative expenses, but it is possible for the government to increase that cost by increasing taxes on deposits over some reasonable level.

9.1 Stranded money as an explanatory factor
Stranded money potentially explains several real-world phenomena that economists observe. It creates a situation where the quantity of money in existence keeps increasing outside of deposits that are created during the lending process. Stranded money also puts part of the burden of loans gone bad on society at large rather than on the bank that made the lending decision. It seems extremely unlikely that loans that were made 30 or 40 years ago were fundamentally riskier than loans made today at the same lending parameters. Yet the compensation in real terms to lenders has decreased in virtually all developed financial markets (Gamber, 2020). Apparently, lenders world-wide are willing to decrease their risk-reward ratios. Explanations like savings gluts, lower growth rates, and a shortage of safe assets are all floated as explanations for the secular decline in real interest rates, but none of these explanations is satisfactory. Stranded money may be part of the explanation.

Stranded money also may help explain why asset prices are bid up over time. The marginal cost of money is small. So, assets that yield more than the cost to create the deposit will be purchased. Banks have almost no choice but to purchase more and more stocks/bonds to soak up the cash they have. Banks can use exogenous cash to make a loan, but absent loan demand, they will buy stocks/bonds. In financial markets, stranded money may contribute to collateral shortages, or more accurately put, deposit excesses. The quantity of safe collateral – usually treasuries – is shrinking relative to the quantity of money sitting in deposit accounts. Stranded money is not unique to the US monetary system. Any system where deposits are created through lending will have a certain amount of stranded money. For example, in Japan many interest rates have been nominally around zero for years on end. The Japanese banking system had huge losses between 1992 and 2001 (Bank of Japan, 2002), creating lots of stranded money in the process. The BOJ also had very large holdings of government bonds, private sector bonds, and equities, which meant that it had been creating deposits for assets (bonds and
mortgages) that already themselves were the product of deposit creations. By buying and creating deposits against equities, even the expected net present value of earnings was turned into deposits. The resulting glut of stranded money combined with large scale liquification of financial assets suppressed interest rates and kept them stuck around zero for a long time. Just like in the US, interest rates in Europe have been trending downwards for years, and stranded money is a possible partial explanation.

9.2 Interest rate determinants
Usually, interest rates are explained by demand and supply dynamics. But looking at the monetary world through either an exogenous or an endogenous lens leads to different answers. The question of what causes a particular level of interest rates is largely a function of the system in which a loan is originated. Because the quantity of money is fixed in an exogenous system, one can argue that rates reflect demand for money and, to a limited degree, the willingness to provide money.
Not so in a purely endogenous system. In theory, in an endogenous system, interest rates should reflect the cost of liquification of assets. So as long as there is enough eligible collateral, costs (interest rates) can be minimal. These costs include the cost of valuing the collateral and various closing costs.
Rates in economies with time to accumulate stranded money trend towards transaction costs. This makes sense because the cost of the deposit connected to the loan is basically zero. Another reason for interest rates to have a downward trend is the liquification of financial assets, which leads to deposit creation in excess of existing real assets. Because interest rates have a real floor set by transaction costs, at least from a bank’s point of view, the Fed’s ability to drive down rates by purchasing treasuries and mortgages starts to hit a wall. There is a minimum cost in the bond/deposit transformation, so in order to drive down rates more than that, The Fed must make more and more purchases.

10. Conclusion
In a way, we are coming full circle. What started as an exogenous system over time transformed into an endogenous system. But as more and more stranded money has built up, the monetary system is starting to work more like an exogenous system again.
Comparing losses and deposits provides some insight. As of the end of the third quarter of 2022, according to the FDIC there are $19.3 trillion in deposits. According to the same dataset, since 1984 average loan losses have been 16 basis points of the deposit base. The cumulative net charge-offs add up to $1.7 trillion, which is approximately 9% of the current deposit base. It is unclear how much in losses were incurred before the FDIC started tracking these variables, but it is safe to say that the number is greater than zero. However, there is an alternative way of looking at the quantity of stranded money. The uncertainty with respect to the quantity of stranded money revolves around how losses are recovered.
If a bank creates a deposit, the borrower defaults, and at some point partially repays what was owed, the difference between the deposit that was created and the repayment becomes stranded money. If, however, the bank creates a deposit, a borrower defaults, and then the bank takes possession of the property (the collateral) and resells it, if the purchaser of the property also borrows money to finance the purchase, i.e., a bank creates yet another deposit, the total quantity of deposits that were created can be in excess of the value of the asset that was pledged as
collateral. In this case it makes more sense to look at the FDIC “Total charge-off” data rather than the “Net charge-offs.” That data show that the cumulative total charge-offs are $2.1 trillion, amounting to approximately 11% of the 2022 Q3 deposit base.

These unanchored deposits have accumulated over time and are all looking for yield. So instead of deposits that are created because the asset owner requires liquidity these deposits are looking for assets which have a positive yield. In other words, the causal relationship has changed direction. That search for yield is slowly but surely driving down interest rates to the point where rates no longer are related to the risk that is taken. This is because in addition to deposits that are created when real assets are pledged, the quantity of stranded money keeps increasing. Banks can use stranded money to invest in the reverse way of how they create deposits: They now have (stranded) deposits and can invest them in financial assets. So, although from an accounting point of new the books will still be in balance and there is no outward sign that there is an issue, the order of operations is reversed. And because these deposits are in addition to the ones that are created in the normal course of lending, they start to put downward pressure on yields. At the same time, the quantity of deposits that are created through multiple liquifications of financial assets decreases the stability of the financial system. In practical terms, the taxpayer will need to shoulder actions that may be needed to bail out the system once it crashes, socializing risk.

Our current monetary system has two complementary parts, but current monetary theory does not appear to make the distinction. If traditional monetary theory is to offer insight into economic events, we need to be as clear as possible whether we are applying rules and remedies to the exogenous monetary system or the endogenous monetary system. But before we can do that, we need to be explicit about identifying which deposits are created through liquification and which are not. If we are to gain understanding of the endogenous monetary system, we must begin to measure the economic action that occurs in it. The twin sources of excess money creation can be remedied through two mechanisms. Excess deposits caused by loan losses can be mopped up by instituting a tax on deposits. Then those deposit-related tax funds could be used to extinguish treasury securities held by the Fed. There probably should be a number of parameters to limit the social impact of impacted deposits by such a tax. The second source of excess deposits created by multiple liquifications of financial assets will likely require regulatory measures. This would also entail regulation of the shadow banking sector which, although in principle operates in the Savings = Investment part of financial markets also, in certain circumstances, engages in pledging financial assets as collateral at commercial banks for more deposits. The continuously increasing quantity of excess deposits puts upwards pressure on asset prices because deposits have negative carry, and holders of the deposits want to swap them for investments with positive expected returns. This is potentially a material factor which puts upwards pressure on financial asset prices specifically but overall inflation as well.

In order to test any macroeconomic theory, we first need to know which monetary regime we are in or what the mix is between the two types. To get a better understanding of that we need to measure variables like which deposits are effectively stranded, or exogenous, and how much of the deposits in the banking system are due to multiple liquifications. Only then can we have a
better understanding of the various ebbs and flows in the financial system, enabling us to have better control over the banking system.

References

Electronic copy available at: https://ssrn.com/abstract=4637150
