Inflation and the growth rate of money in the long run and the short run

Javier Díaz-Giménez and Robert Kirkby

SEF WORKING PAPER 09/2016
The Working Paper series is published by the School of Economics and Finance to provide staff and research students the opportunity to expose their research to a wider audience. The opinions and views expressed in these papers are not necessarily reflective of views held by the school. Comments and feedback from readers would be welcomed by the author(s).

Further enquiries to:

The Administrator
School of Economics and Finance
Victoria University of Wellington
P O Box 600
Wellington 6140
New Zealand

Phone: +64 4 463 5353
Email: alice.fong@vuw.ac.nz
Abstract

Between 1960 and 2013, in the United States the inflation rate was essentially proportional to the growth rate of money in the long run, but that relationship did not hold in the short run. We ask whether three standard monetary model economies from the Cash-in-Advance, the New-Keynesian, and the Search-Money frameworks replicate these two facts. We find that all three deliver the first fact, but that they fail to deliver the second fact, since in all three of them the inflation rate is proportional to the growth rate of money both in the long run and in the short run. This is because in all three model economies the price level responds too quickly to changes in the growth rate of money.

Keywords: Monetary Economics; Quantity Theory of Money; Cash-in-Advance; New-Keynesian; Search-Money.
1 Introduction

According to Robert E. Lucas, “the central predictions of the quantity theory are that money growth should be neutral in its effects on the growth rate of production and should affect the inflation rate on a one-for-one basis”.\(^1\) In this article we focus almost exclusively on the second one of these two predictions and we try to find out whether money growth affects the inflation rate on a one-for-one basis in the United States and in three model economies that have been designed and calibrated for other purposes. To do this, we follow the method proposed in Lucas (1980) classical paper and we examine the relationship between the growth rate of money —in excess of the growth rate of output— and the growth rate of the price level, both in the long run and in the short run.

This makes ours a peculiar paper in several ways. First, we have not found some obscure fact that contradicts some established theory and we have not solved a fancy model that delivers that fact. Instead, we revisit the Quantity Theory of Money (QTM) relationship between the growth rates of money and prices —arguably, one of the best-known relationships in monetary theory—and we use that relationship to test three simple model economies that have been designed for other purposes. Our model economies are standard examples of the Cash-in-Advance, the New-Keynesian, and the Search theories of money —arguably, the three monetary theories that have been most-widely used in empirical studies of money.

Our paper is also peculiar in that our test gives us mixed results. We show that in the United States the inflation rate is essentially proportional to the growth rate of money in the long run, but that it is not so in the short run. In contrast, the three model economies that we study deliver the first fact, but they fail to deliver the second fact, since in all three of them the inflation rate is proportional to the growth rate of money both in the long run and in the short run. If we were to grade the three models based on this test alone, each of them would score a B, but none of them would score an A.

The Tests. To verify whether the QTM relationship between growth rates of money —in excess of output— and the price level holds in the short run, we do the obvious thing: we plot the growth rate of the price level plus the growth rate of output —in the United States we use the Consumer Price Index plus real GDP— against the growth rate of money —in the United States we use both M1 and M2— and we compute the average distance from the 45 degree line that goes through the grand mean of our samples. If the QTM relationship holds, the data points will trace that 45 degree line. If it does not hold, the scatter plot of the growth rates of money and the price level will form a vague cloud, and no clear pattern will emerge.\(^2\)

To verify whether the QTM relationship holds in the long run, is somewhat more complicated.

---

\(^1\)See Lucas (1996), page 665.

\(^2\)Subtracting the growth rate of output from the growth rate of money —or, equivalently, adding it to the growth rate of the price level— is a slight twist on Lucas (1980) which we justify in Section 2 below.
As Lucas (1996) states: “The modifier ‘long run’ is not free of ambiguity but, by any definition the use of data that are heavily averaged over time should isolate only long-run effects”. To make the meaning of “long run” operative, we follow the methods described in Lucas (1980). In that article, Lucas defines the long-run as the low frequency fluctuations of the series and he uses a two-sided moving average filter to extract these low frequency fluctuations. We choose this method to test the QTM relationship because it does not depend on any structural assumptions. Therefore, it can be used to make genuine comparisons across very different monetary frameworks. This differs from impulse response functions and from other alternative methods used in monetary economics which typically depend on model specific structural assumptions and, therefore, cannot be used to make clean comparisons across monetary frameworks that differ widely.

The Model Economies. The three monetary frameworks that we study here represent different approaches to modelling money. The Cash-in-Advance framework focuses on the role of money as an exogenous facilitator of transactions, the New-Keynesian framework focuses on the role of money as a nominal anchor around which prices are sticky, and the Search-Money framework focuses on the role of money as an endogenous mechanism of exchange that solves the problems created by the absence of double-coincidence of wants in barter. We have chosen a canonical model economy to represent each one of these frameworks. These model economies were not designed with the specific aim of delivering the QTM relationship between the growth rates of money and the price level observed in the data. This makes them ideal for our test, because they have not been rigged to ace it.

Three reasons have lead us to choose standard implementations of all three frameworks: (i) we believe that using more complicated models models would not solve the difficulties of the frameworks, (ii) many of the modelling complications, such as capital-adjustment costs or consumption habit formation, could be just as easily applied to any of the three frameworks and would not help us to decide which one of them is the best approach to modelling money, and (iii) the standard implementations of the models economies are easier to follow and understand, and allow us to provide a more intuitive account of our results. We return to discuss point (i) in Section 6 below.

The Cash-in-Advance framework. The Cash-in-Advance framework makes the use of money in exchange compulsory by forcing households to buy consumption goods using money carried over from the previous period. In general, this cash-in-advance constraint is inefficient, but it solves the informational problem that would arise when trying to coordinate all the simultaneous trades; a problem that non-monetary economies ignore. In the words of Lucas (1980) the cash-in-advance framework “is an attempt to study the transaction demand for money in as simple as possible a general equilibrium setting”. In this article, we use Cooley and Hansen (1989)’s cash-in-advance business cycle model as our canonical example of the cash-in-advance framework. We chose this model economy because it combines a cash-in-advance constraint and the standard neoclassical

---

3 See Lucas (1996), page 665.
model of business cycles.

The New-Keynesian framework. The New-Keynesian framework models the relationship between money and interest rates using a money demand equation. To break away from the short-run neutrality of money, New-Keynesian models assume that prices are sticky. In this framework, prices either cannot be changed every period by assumption, or doing so is costly. This price stickiness assumption is justified using empirical evidence that prices do not change often in the real world. In this article, we use the New-Keynesian monetary model economy described in Chapter 3 of Galí (2008) as our canonical example of the New-Keynesian framework. We chose this model economy because it includes money explicitly.\(^4\)

The Search-Money framework. The Search-Money framework is a successful attempt to satisfy Wallace (1998)'s dictum that “money should not be a primitive in monetary theory”; that is, that there must be an endogenous reason that justifies the existence of money. In the Search-Money framework this reason is to enable trade. Search-Money models assume that people meet in pairs and exchange goods using barter. But this means that trade only occurs when both trading partners have a good that the other one wants. This is the well-known problem of barter: trade is often limited by the absence of a double-coincidence of wants. Money solves this problem because everyone always wants money, at least for future use as a trade-enabler. In this article, we use a stochastic extension of the Search-Money model described in Aruoba, Waller, and Wright (2011) as our canonical example of the Search-Money framework.\(^5\) We chose this model economy because it includes capital. Capital accumulation plays an important role in these economies because it reduces the number of monetary trades and, consequently, it amplifies the effects of monetary innovations on the rest of the economy, where monetary exchanges take place.\(^6\)

Findings. First the long run. We plot the filtered data of the growth rates of M1 and of M2 against the growth rate of the Consumer Price Index plus the growth rate of real GDP in the United States in the 1960–2013 period and we show that the Quantity Theory of Money relationship between the growth rates of money and the price level held in the long-run during that period. Our measure of the average distance to the 45 degree line and we confirms this finding. Then we carry out the same exercise for observations on the growth rates of money and the price level in our three model economies. In all three model economies the filtered points lie along the 45 degree line that goes

\(^4\) Many New-Keynesian models often omit money entirely and they use a Taylor rule on interest rates instead. They rationalize this modelling choice on the grounds that modern central banks tend to focus on interest rates, and not so much on monetary aggregates. When they model money explicitly their standard approach is to use of a money-demand equation. See, for example, Christiano, Eichenbaum, and Evans (2005) and Sargent and Surico (2011).

\(^5\) Berentson, Menzio, and Wright (2011) also solve a stochastic extension of the original Search-Money model described in Lagos and Wright (2005). Their extension differs slightly from ours in the timing of the money shock. It also differs because they impose an AR(1) process on interest rates, which implies a Markov process on money, while we impose an AR(1) process on (log) money, which implies a Markov process on interest rates.

\(^6\) Specifically, in Lagos and Wright (2005) model economy which does not include capital, monetary trades account for 20.6% of real output on average. In contrast, when capital is added to a Search-Money model economy, as in Aruoba, Waller, and Wright (2011), monetary-trades account for 1.6% of real output on average.
through the grand mean of the sample. This means that the Cash-in-Advance, the New-Keynesian, and the Search-Money frameworks also display the Quantity Theory of Money relationship in the long run and, therefore, that they replicate the long-run behavior of the United States. Finally, we simulate 100 stochastic realizations of the equilibrium processes of the three model economies and we find that the Cartesian distances of the filtered points from the 45 degree lines are close to zero in all three cases. This confirms our qualitative results. We conclude that, in the long run, the differences between the three frameworks, if any, are tiny, and that all three of them pass the long-run Quantity Theory of Money test with flying colors.

Then the short run. When we repeat this exercise in the short run, we find that the Quantity Theory of Money relationship between the growth rates of money and prices does not hold in the United States and that the three model economies that we study fail to replicate this finding. While the graph of the United States unfiltered data contains no suggestion of the Quantity Theory of Money relationship, our simulations of the three model economies produce data points that lie very close to the 45 degree line. This suggests that the three model economies display a short-run Quantity Theory of Money relationship between the growth rates of money and the price level that is too tight, even though it is not exact. The Cartesian distances of the data points from the 45 degree lines confirm our qualitative results.

Conclusions. We conclude that in our three model economies the price levels respond too quickly to changes in the rate of growth of money. Therefore, the search for a model economy in which the short term response of prices to monetary innovations replicates the sluggishness found in the data still remains an important challenge for monetary economics.

2 Illustrating the Quantity Theory of Money

Multiply the supply of money by $\alpha$ and prices will become $\alpha$ times larger —this is a rough but useful characterization of the Quantity Theory of Money. More precisely, the Quantity Theory of Money claims that the rate of growth of nominal prices plus the rate of growth of output is approximately equal to the rate of growth of the money supply. The formal expression of the Quantity Theory of Money is the following

$$MV = PY$$

(1)

where $M$ is the nominal money supply, $V$ is the velocity of circulation of money, $P$ is the price level, and $Y$ is real output. Let $g_x$ be the growth rate of variable $x$. Then, if we assume that $V$ is relatively constant and, consequently, that $g_V = 0$, it follows that

$$g_M \simeq g_P + g_Y$$

(2)

According to Lucas (1996) interpretation of this equation, when the Quantity Theory of Money holds, $g_Y = 0$ and $g_P = g_M$. Therefore, when the Quantity Theory of Money holds, if we graph...
$g_P + g_Y$ against $g_M$, we will get a 45 degree line. And, when the Quantity Theory of Money does not hold, we will get a meaningless bird-shot scatter plot. This is the central idea behind the Quantity Theory of Money illustrations provided in Lucas (1980).

As we have mentioned above, by adding the growth rate of output to the growth rate of the price level, we are in fact using a slight twist on the original method of Lucas (1980). Output was absent, at least explicitly, from the first formulation of the Quantity Theory of Money made by David Hume. Hume presented the theory as the following thought experiment: “Were all the gold in England annihilated at once, and one and twenty shillings substituted in the place of every guinea, would money be more plentiful or interest lower? No surely: We should only use silver instead of gold”.\footnote{See Hume (1742), Of Interest.} It is not clear when output was first included explicitly, but it plays a prominent role in Fisher (1911) and in Friedman and Schwartz (1963). Writing about the Greenback period, 1867-1879, Friedman and Schwartz observe that prices decreased slightly, despite an increase in the money supply, and they attribute the difference as being substantially due to the large increase in real output that occurred during this period.

So our inclusion of output in our formulation of the Quantity Theory of Money is certainly not novel. It is only our inclusion of output in the method of Lucas (1980) which is novel. Other applications of Lucas (1980)’s methods by Whiteman (1984) and Sargent and Surico (2011) follow Lucas literally and they omit output also. However, when doing so, there is no reason to expect a forty-five degree line in the QTM plots, as the following example illustrates.

Start with the standard undergraduate textbook formulation of the QTM that we have mentioned before: $g_P \simeq g_M - g_Y$. Then, given some exogenous starting reference levels for $g_M$, $g_P$ and $g_Y$, and if nothing else changes, any change in $g_M$ will have no effect — that is, it will “be neutral” — on $g_Y$ and it will lead to one-for-one changes in $g_P$. This formalizes Lucas (1996)’s quote which we cited above. Consistently, with this argument Lucas (1980) omits the exogenous starting reference level, he simply plots $g_P$ against $g_M$, and he looks for one-for-one changes that will plot of a forty-five degree line.

But we find this misleading. Dropping the exogenous reference level only works when the reference level remains constant over the relevant time period, or when $g_M$ moves exactly one-for-one with $g_Y$. This idea is best illustrated with a numerical example. Suppose that $g_Y$ is exogenous, and that in three consecutive periods, $g_Y = (0, 2, 4)$, and that $g_M$ is also exogenous and that it is $g_M = (2, 3, 4)$.\footnote{This example is consistent with the the money supply is being set according to the rule $g_M = 2 + 0.5g_Y$.} If we use the QTM to compute $g_P$, we obtain that $g_P = g_M - g_Y = (2, 1, 0)$. If we plotted the $(g_M, g_P)$ points we would plot $(2, 2)$, $(3, 1)$ and $(4, 0)$, which clearly do not lie on a forty-five degree line. But if we plotted the $(g_M, g_P + g_Y)$ points, we would plot $(2, 2)$, $(3, 3)$ and $(4, 4)$, which clearly do lie on a forty-five degree line. Therefore, since the period between 1960 and 2009, which is the period that we consider, is a long period and the growth rates of U.S. GDP...
changed substantially during that period, we plot \( g_P + g_Y \) against \( g_M \) to test whether the QTM relationship between the growth rates of money, output and the price level holds in the data.\(^9\)

**The Quantity Theory of Money in the Long Run**

To illustrate whether the Quantity Theory of Money holds in the long run, Lucas (1980) associates the short-run with the high-frequency fluctuations of the growth rates of the quantity theory time series expressed, and the long-run with the low-frequency fluctuations of the growth rates of those series. To remove the high-frequency fluctuations and to obtain the low-frequency signal, Lucas transforms the original series using the following two-sided, exponentially-weighted, moving-average filter

\[
x_t(\beta) = \alpha \sum_{k=1}^{T} \beta^{|t-k|} x_k
\]

where

\[
\alpha = \frac{(1 - \beta)^2}{1 - \beta^2 - 2\beta(T+1)/2(1 - \beta)} \quad 0 \leq \beta < 1
\]

and where \( T \) is the number of observations in the time series.\(^10\)

A value of \( \beta = 0.0 \) returns the original time series. Increasingly higher values of \( \beta \) filter out the higher frequency fluctuations from the original time series and leave only the increasingly lower frequency fluctuations in the transformed series. Figure 1 illustrates how our version of Lucas’ filter transforms the original U.S. time series as we change the value of parameter \( \beta \).\(^11\) The filter is two-sided because the behavior of households is likely to be affected both by what happened to them in the past and by their expectations of what might happen to them in the future.\(^12\)

One important advantage of using Lucas’ methods to find out whether the Quantity Theory of Money holds in the long run in model economies is that his filter is atheoretical. This means that its results do not depend on any modelling assumptions. In contrast, other methods that are more sophisticated econometrically, such as structural VARs, require identifying assumptions that are model-dependent. Those methods are less useful to compare model economies that are fundamentally different, like those that we consider in this article.\(^13\)

\(^9\)In the model economies including or not \( g_Y \) makes little difference because we assume that their exogenous growth rates of output are zero.

\(^10\)Parameter \( \alpha \) guarantees that the means of the original and the filtered time series coincide. In fact, Lucas (1980) uses a slightly different definition of this parameter. He makes \( \alpha = (1 - \beta)/(1 + \beta) \). His definition guarantees that the means coincide assuming that the lengths of the unfiltered series are infinite. Instead, we use Sargent and Surico (2011)’s small-sample correction to the value of \( \alpha \). This correction preserves the means of the series, but assuming that the lengths of the unfiltered series are finite.

\(^11\)To prevent clutter, in all our figures we follow Lucas (1980) exactly and plot only the fourth quarter of every year. To prevent end-of-sample distortions, we drop the first five and last five years from each graph, even though we use them in the filter.

\(^12\)The choice of filter is not important. For example, Benati (2005, 2009) reports similar conclusions using a band-pass filter.

\(^13\)See Lucas (1980) for a discussion of this filter and of its frequency interpretation, and see Whiteman (1984) for further details on this discussion.
Higher values of $\beta$ extract the higher frequency fluctuations from the original series. Therefore, if the Quantity Theory of Money relationship holds in the long-run, as we increase the value of $\beta$, the plots of the filtered time series should look increasingly like the 45 degree line that runs through the grand mean of the unfiltered series. And, if it does not hold, we have no theory to account for the relationship between those variables and we expect the filtered data to become a blob around the grand mean of the unfiltered data. In fact, Lucas (1980) shows that this is precisely what happens when he plots the unemployment rate against the rate of growth of money, for the 1955–1975 period.

Figure 1: Lucas’ Illustrations in the United States

Panel A: M1 ($\beta = 0.00$)  
Panel B: M1 ($\beta = 0.90$)  
Panel C: M1 ($\beta = 0.95$)  
Panel D: M2 ($\beta = 0.00$)  
Panel E: M2 ($\beta = 0.90$)  
Panel F: M2 ($\beta = 0.95$)

*The coordinates of the center of the white circle in each panel are the grand mean of the unfiltered sample.

Quantifying Lucas’ Illustrations

To quantify Lucas (1980)’s illustrations, we use a relatively straight-forward method. Specifically, we compute the average Cartesian distance of the points in the plots from the 45 degree line that runs through the grand mean of the unfiltered observations. The Cartesian distance of a point, $(x_i, y_i)$, from a line, $ax + by + c = 0$, is $d = |ax_i + by_i + c|/\sqrt{a^2 + b^2}$.  

Another method that has been used in the literature is to compute the slope of an ordinary least squares (OLS)
The formal definition of the our distance statistic is the following

$$D_{45} = \frac{1}{\sqrt{2T}} \sum_{i} |x_i - y_i + (\bar{y} - \bar{x})|$$  \hspace{1cm} (5)$$

where $y_i$ is the value of the $i$-th observation of the growth rate of prices plus the growth rate of output, either of the original or of the filtered time series; $x_i$ is the corresponding observation of the growth rate of money and $\bar{x}$ and $\bar{y}$ are the average values of the unfiltered $x_i$ and $y_i$. Obviously, if the Quantity Theory of Money relationship holds, the value of the $D_{45}$ statistic will be small and, if it does not hold, it will be large.

*The Quantity Theory of Money in the Model Economies*

As we have already mentioned, the main purpose of this paper is to explore the extent to which the Quantity Theory of Money relationship between the growth rates of money and the price level holds in three of the modelling frameworks most frequently used by economists to think about monetary policy: the Cash-in-Advance framework, the New-Keynesian framework, and the Search-Money framework. For each one of these three frameworks we choose a representative model economy: for the Cash-in-Advance framework, we use the model economy described in Cooley and Hansen (1989); for the New-Keynesian framework, the model economy described in Chapter 3 of Galí (2008); and, for the Search-Money framework, a stochastic extension of the model economy described in Aruoba, Waller, and Wright (2011). Since these three model economies are standard, in the literature we relegate their detailed description to Appendix A.

We also describe in detail our calibration procedure in that appendix. To make our comparisons meaningful, we use the same functional forms and parameter values for the utility functions and for the processes on the technology and the monetary shocks, whenever possible.\footnote{We repeated our calculations with the functional forms and the calibration targets used in the original articles, and we found that this does not change our results qualitatively. This is partly due to the fact that the original articles target similar data moments and study similar time periods.} We also use the same methods to characterize the equilibrium processes of our three model economies and to find their solutions.

In all three cases, we describe the equilibrium processes as systems of stochastic difference equations and we solve these systems using the default perturbation methods of Dynare that allow us to obtain quadratic laws-of-motion. Then we simulate the three model economies and we obtain samples of 204 quarterly observations to replicate the number of observations in our United States sample. To obtain these samples, we use the same seeds for the random number generators. Consequently, the sequences of realizations of the random processes are identical in the three model economies.
3 Illustrating the Quantity Theory of Money in the Long Run

The Quantity Theory of Money in the Long Run in the United States

In Panel C of Figure 1 we plot the Quantity Theory of Money relationship between the growth rates of M1 and the price level in the United States in the long-run or, more precisely, when $\beta = 0.95$, and in Panel F of Figure 1 we repeat that plot for M2. In both panels of that figure we see that, when we filter out the high-frequency fluctuations, the original bird-shot scatters displayed in Panels A and D disappear and the observations approach the 45 degree line that runs through the grand mean of the unfiltered sample. Therefore, the scatter plots displayed in Figure 1 illustrate that the Quantity Theory of Money held in the long-run in the United States both for M1 and for M2 during the 1960–2009 period, and they confirm Lucas (1980)’s findings.

The long-run scatter plot for M1 is interesting from the perspective of the monetary history of the United States. In Panel C of Figure 1 the observations march roughly up the 45 degree line during the late 1960s and the 1970s. When they reach the top-right-hand corner of the graph, they suddenly drop down almost vertically. This period of sharply falling average growth rates of prices represents the beginning of the 1980s when the Federal Reserve, under Paul Volcker, started tightening monetary policy to fight inflation —and eventually defeat it. Then, in the 1990s and 2000s the points return to the 45 degree line as the U.S. economy transitions to a new monetary regime with a lower inflation rate and lower money growth rate, ending with the lower-leftmost points in the graph.

Table 1: The Quantity Theory of Money Statistics in the Long Run

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D45 (std dev)</td>
<td>0.50003</td>
<td>0.3953</td>
<td>0.0555</td>
<td>0.0149</td>
<td>0.0188</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0147)</td>
<td>(0.0031)</td>
<td>(0.0039)</td>
</tr>
</tbody>
</table>

In the second and third columns Table 1 we report the values of D45 statistics for M1 and M2 the United States in the long run, that is when the smoothing parameter of the filter is $\beta = 0.95$. Our numerical results confirm what we learnt from Lucas’ Illustrations. The Quantity Theory of Money relationship between the growth rates of money and the price level held in the long run in the United States between 1960 and 2009 for both monetary aggregates. Moreover, according to our D45 statistic, the Quantity Theory of Money relationship was tighter for M2 than for M1, since the distance from the 45 degree line was smaller for M2.

The Quantity Theory of Money in the Long Run in the Model Economies

Panels C, F, and I of Figure 2 represent Lucas’ Illustrations in the long run, that is, for $\beta = 0.95$, for the monetary aggregates in our Cash-in-Advance, New-Keynesian, and Search-Money model
Figure 2: Lucas’ Illustrations in the Model Economies

A: Cash-in-Advance ($\beta = 0.00$)  
B: Cash-in-Advance ($\beta = 0.50$)  
C: Cash-in-Advance ($\beta = 0.95$)

D: New-Keynesian ($\beta = 0.00$)  
E: New-Keynesian ($\beta = 0.50$)  
F: New-Keynesian ($\beta = 0.95$)

G: Search-Money ($\beta = 0.00$)  
H: Search-Money ($\beta = 0.50$)  
I: Search-Money ($\beta = 0.95$)

*The coordinates of the center of the white circle in each panel are the grand mean of the unfiltered sample.
economies. We find that the Quantity Theory of Money relationship between the growth rates of money and the price level clearly holds in the long run in all three of them. In fact, the Quantity Theory of Money relationship is so tight in every model economy that we are hard put to say in which one of them it is tightest. For that purpose, we must turn to the values of the D45 statistics that we report in the last three columns of Table 1.

While we have only one United States time series from which to compute the D45 statistics, we can simulate many stochastic realizations of the equilibrium processes of our model economies. To reduce the size of the sampling error, we compute the D45 statistics for our model economies using 100 independent random samples. The average values of the D45 statistics show that the Quantity Theory of Money relationship between the growth rates of money and the price level is tightest in the New-Keynesian model economy, followed by the Search-Money model economy and by the Cash-in-Advance model economy. But the differences between them are small. And in all three cases the Quantity Theory of Money relationship is much tighter than in the United States.

Conclusion. We conclude that in the long run the Quantity Theory of Money relationship is present both in the United States and in our three model economies. But, once again, it is sizably tighter in the model economies.

4 Illustrating the Quantity Theory of Money in the Short Run

The Quantity Theory of Money in the Short Run in the United States

To find out whether the Quantity Theory of Money relationship between the growth rates of money and prices held in the short run in the United States in Panels A and D of Figure 1 we plot the rate of growth of the Consumer Price Index against the rates of growth of M1 and M2 using quarterly data from the United States economy for the period 1960:Q1–2013:Q4. We obtain two bird-shot scatter plots. This shows that the Quantity Theory of Money did not hold in the short run in the United States during that period and it confirms and updates Lucas (1980)'s findings.\footnote{We have taken all the data from FRED2 (http://research.stlouisfed.org/fred2/). The time series that we have used are GNPC96, M1SL, M2SL, and CPIAUCNS. Our results are robust to using alternative measures for inflation: CPIAUCSL, CPIFENS, CPIFESL, GDPDEF, and real output GDPC1. We measure the growth rates as the percentage changes on the same quarter of the previous year.}

In the second and third columns Table 2 we report the values of D45 statistics for M1 and M2 the United States in the short run, that is when the smoothing parameter of the filter is $\beta = 0.00$. Our numerical results confirm what we learnt from Lucas’ Illustrations. The Quantity Theory of Money relationship between the growth rates of money and the price level did not hold in the short run in the United States between 1960 and 2009 for both monetary aggregates. Moreover, according to our D45 statistic, the scatter plot was more disperse for M1 than for M2, since the distance from the 45 degree line was bigger for M1.

11
### Table 2: The Quantity Theory of Money Statistics in the Short Run

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D45</td>
<td>2.9850</td>
<td>2.5420</td>
<td>0.9338</td>
<td>1.5611</td>
<td>1.9278</td>
</tr>
<tr>
<td>(std dev)</td>
<td>–</td>
<td>–</td>
<td>(0.0890)</td>
<td>(0.1185)</td>
<td>(0.1408)</td>
</tr>
</tbody>
</table>

The Quantity Theory of Money in the Short Run in the Model Economies

Panels A, D and G of Figure 2 represent Lucas' Illustrations in the short run—that is for \( \beta = 0 \)—for the monetary aggregates of our three model economies. We observe that the Quantity Theory of Money relationship is much stronger in our three model economies than in the United States (see Panel A of Figure 1). While in our three model economies the points lie close to the 45 degree line, in the United States data it is hard discern any pattern.

In the last three columns of Table 2 we report the sample means and the sample standard deviations of the D45 statistics in our three model economies. These statistics confirm what we found using Lucas' Illustrations, and they establish that our graphs are not the result of a sampling oddity. While the D45 statistics for M1 and M2 in the United States were 2.99 and 2.54, in all three model economies it is below 2.0. This result arises from the fact that the points in the graph for the United States form a shapeless cloud, while those in the graphs for the three model economies form clouds are substantially closer to the 45 degree line. Therefore, the D45 statistics confirm that our three model economies display too much of the Quantity Theory of Money relationship in the short run, when compared with the United States economy. When we compare the values of the D45 statistics of the three model economies we find that the money and price level growth rate pairs are furthest away from the 45 degree line in the Search-Money model economy, next in the New-Keynesian model economy and, finally, in the Cash-in-Advance model economy. Specifically their values are 1.93, 1.56 and 0.93.

Conclusion. We interpret these results to mean that in our three model economies the rate of growth of prices responds too quickly to changes in the rate of growth of money, relative to the United States, or that our three model economies do not display enough short-run sluggishness in the response of prices.

The Departures from the Quantity Theory of Money in the Short Run

In this subsection we describe how the three model economies depart from the Quantity Theory of Money in the short run using only one equation for each one of them. Specifically, we provide an expression for the equilibrium values of the term \( PY/M \) for each model economy.\(^{17}\) We provide

---

\(^{17}\)These expressions are also related to the issue of money demand as discussed in Lucas (2000). Lucas defines money demand as the relationship between nominal interest rates and the ratio of real money holdings to real output,
the derivation of these equations in Appendix A, together with the full descriptions of the model economies. If the Quantity Theory of Money held exactly $PY/M$ would be constant. Therefore, these single equation expressions thus help to understand how each model economy departs from the Quantity Theory of Money in the short-run.

The Cash-in-Advance Model Economy

In the Cash-in-Advance model economy we obtain that

$$\frac{PY}{M} = \frac{P(C + X)}{M} = 1 + \frac{PX}{M}$$  \hspace{1cm} (6)

where $C$ is consumption, and $X$ is investment. So the Cash-in-Advance framework succeeds in departing from the Quantity Theory of Money in as far as monetary policy distorts investment decisions. These distortions take place on the cash-good (consumption) and credit-good (investment) margin.

The New-Keynesian Model Economy

In the New-Keynesian model economy when we rewrite $PY/M$ in logs we obtain that

$$pt + yt - mt = \eta i_t = \eta r^n_t + \eta E_t\{f(\pi_t, \pi_{t+1}, \pi_{t+2})\}$$  \hspace{1cm} (7)

where $i_t$ is the nominal interest rate, $\eta$ is the elasticity of money demand, $r^n_t$ is the natural real interest rate, and $f(\pi_t, \pi_{t+1}, \pi_{t+2})$ is a linear function of current and future inflation. It is evident from expression (7) that the elasticity of money demand and the changing values of the nominal interest rates play an important role in allowing the New-Keynesian model to get away from the Quantity Theory of Money in the short run.

What role do sticky prices play in this? The natural real interest rate, $r^n_t$, is independent of both monetary variables and the parameters that determine the degree of price stickiness. So, for sticky prices to be part of the story, they must operate through the inflation rate which evolves according to

$$\pi_t = (1 - \theta)(p^*_t - p_{t-1})$$  \hspace{1cm} (8)

where $1 - \theta$ is the fraction of firms that get to adjust their prices at time $t$, and $p^*_t$ is the price level that they choose. So sticky prices affect the rate of inflation and, therefore, the nominal interest rates and they contribute to the short-run departure from the Quantity Theory of Money relationship. In practice, however, this effect is small.\(^{18}\) This is because the nominal interest rate or $(M/P)/Y$. This ratio is the inverse of the $PY/M$ term which we consider here.

\(^{18}\)We experimented with various parameter values, and we found that the Quantity Theory of Money relationship was largely unaffected by the degree of price-stickiness, $\theta$, but that it was very sensitive to the value of the elasticity of money demand, $\eta$. 

13
does not change immediately as predicted by Fisher’s equation, \( i = r + \pi \), after a monetary shock because these shocks generate a liquidity effect. But this liquidity effect is not very long-lasting and it all but disappears, when we filter out the high frequency fluctuations of the nominal time series. Part of the problem for the New-Keynesian model in breaking free from the Quantity Theory of Money relationship is that any large change in money growth, which might allow the model to get away from the Quantity Theory of Money, is persistent and so is picked up in expectations, meaning that inflation moves as well, resulting in shifts along the forty-five degree line rather than movements away from it. We conclude that the short-run behavior of the New-Keynesian model arises directly from the money demand equation, and that the role played by the degree of price stickiness is small.

**The Search-Money Model Economy**

In the Search-Money model economy we obtain that,

\[
\frac{PY}{M} = \frac{1}{z(q, K)} F_N(K, N)
\]

(9)

where \( \gamma \) is a parameter that quantifies the disutility of labor. In the simulations of this model economy total output, \( Y \), the stock of capital, \( K \), and the marginal product of labour, \( F_N(K, N) \), are almost constant. Consequently, they do not account for the departure from the Quantity Theory of Money relationship in the short run. They are almost constant because most of the trades are non-monetary, the centralized night market is much bigger than the decentralized day market, and this market is almost unaffected by changes in the money supply. Almost all the variability in expression (9) comes from changes in \( z(q, K) \), which represents the terms of trade in monetary exchanges and, more specifically, from changes in \( q \) —the amount produced and traded in the monetary exchanges that take place in the decentralized market. These changes in \( q \) are caused by the unexpected changes in the amount of money and by changes in the inflation rate, which is the cost of holding money.

In summary, the Search-Money framework succeeds in departing from the Quantity Theory of Money relationship in the short run to a certain extent because of the effects of changes in the money supply on the value of money. People want to hold money because it is useful for monetary trades. The value of money is jointly determined by this demand for money and by the money supply. Since the nominal price of consumption goods is the inverse of the cost of acquiring money, changes in the value of money result in changes in the price level.\(^{19}\) Therefore, changes in the money supply affect the value of money, and thereby the price level. Moreover, the resulting inflation also affects output because of a holdup problem.\(^{20}\) This effect is magnified because monetary trades account only for a small fraction of total exchanges and, therefore, changes in the supply of money

---

\(^{19}\)The price of consumption goods is the number of units of money that agents exchange for one unit of the consumption good. The cost of acquiring money is the number of units of the consumption good that agents give up to obtain one unit of money.

\(^{20}\)In every Search-Money model inflation decreases output. Sellers in single-coincidence meetings know that they
are large relative to the size of the total amount of monetary trades. Consequently, the effect of a
given change in money supply on the value of money and, hence, on prices in the Search-Money
framework, is larger than in the other two frameworks.

5 Which Framework Wins?

We have shown that all three of our model frameworks display the Quantity Theory of Money
relationship in the long run and, therefore, succeed in replicating the long-run behavior of the
United States economy. But we have also shown that the Quantity Theory of Money relationship
is substantially tighter in all three model economies than in the United States in the short run.
Given the importance of departures from the Quantity Theory of Money for the behavior of money,
prices, and hence monetary policy, we think that this is an important shortcoming for all three
model economies.

The difficulties in capturing the short-run departures from the Quantity Theory of Money have
been known to afflict the Cash-in-Advance framework since the work of Hodrick, Kocherlakota,
and Lucas (1991). While the New-Keynesian and Search-Money frameworks have cast light on a
number of other issues in monetary economics, they have not resolved these difficulties. Perhaps
further research within these frameworks will succeed in enabling them to depart from the Quantity
Theory of Money relationship in the short run.

Progress within the Cash-in-Advance framework in the attempt to slow down the response of
prices to monetary shocks—a problem closely related to departing from the Quantity Theory of
Money relationship in the short-run—appears to have stalled. Early attempts to solve this
problem use constructs such as portfolio-adjustment costs (see, for instance, Christiano, 1991,
and Christiano and Eichenbaum, 1995). But this line of research has trailed off since then, after having
been only moderately successful. Perhaps a partial exception to this rule can be found in the recent
work of Alvarez, Atkeson, and Edmond (2009) who allow for cash-in-advance constraints that last
for multiple periods.

On the surface progress within the New-Keynesian framework seems to be more promising. For
example, Christiano, Eichenbaum, and Evans (2005) develop a New-Keynesian model capable of
reproducing the slower reaction of the economy to monetary shocks observed in empirical studies
that use impulse-response functions. This suggests that these more advanced New-Keynesian mod-
els might offer better hopes for departing from the Quantity Theory of Money relationship in the
short run. However this progress may turn out to be superficial. The key shock driving inflation in
the more advanced New-Keynesian models is the cost-push shock. The problem with this shock is
can increase the price of the consumption good because the outside option of the buyer—to hold onto the money
until next period—is less attractive when the inflation rate is higher. These increased prices—known as the holdup
problem—decrease economic efficiency and output. See, eg., Lagos and Wright (2005) for a discussion of the holdup
problem.
that it seems much more like a reduced-form shock than a structural shock, simply filling the gap between model inflation and observed inflation without actually explaining it. This is related to the criticisms of Chari, Kehoe, and McGrattan (2009). We return to this issue in the conclusion.

The Search-Money framework is a more recent construct, and bringing productive capital into that framework is a very recent achievement. Therefore, future refinements within this framework might enable it to depart from the Quantity Theory of Money relationship in the short run. Time will tell.

6 Concluding Comments

In this article we show that the Quantity Theory of Money held in the long-run in the United States between 1960 and 2009. And we also show that it failed to hold in the short-run during that period. Given the prominence of the Quantity Theory of Money in monetary theory, we argue that monetary model economies should replicate both the long-run success and short-run failure of the Quantity Theory of Money observed in the United States, if we are to trust their prescriptions for monetary policy.

Our analysis, based on the Lucas (1980) Illustrations, shows that every one of the three main frameworks that are currently used to study monetary policy — the Cash-in-Advance framework, the New-Keynesian framework, and the Search-Money framework — display the Quantity Theory of Money relationship both in the long-run and in the short-run. This failure of all three frameworks to depart from the Quantity Theory of Money in the short-run casts some doubts on their usefulness for the analysis of monetary policy — which most monetary theorists consider to be an inherently short-run phenomenon.

Perhaps our result may be attributed to the simplicity of the method that we have used to show that it holds. But we contend that this is not the case at least for our New-Keynesian model economy. In an earlier version of this paper we carried out at Bayesian estimation of that economy. To do, we included an additional shock into the model and we followed the standard modelling approach for Bayesian estimations of Monetary Dynamic Stochastic General Equilibrium (DGSE) model economies. We included the shock between the inflation rate observed in the data and the “fundamental” inflation rate predicted by the model. That is, \( \pi_o^t = \pi_f^t + \epsilon_p^t \), where \( \pi_o^t \) is the observed inflation rate, \( \pi_f^t \) is the fundamental inflation rate, and shock \( \epsilon_p^t \) is an AR(1) process. This shock is commonly referred to as a “price mark-up disturbance”. The estimation of our model economy allocated most of the variance in the inflation rate to this price mark-up disturbance. In light of our work in this paper, we interpret this result as an attempt of the the estimation process to use the price mark-up disturbance to allow the data to escape from the strong QTM present in the model economy in the short run. Given that this substantial role of the price mark-up disturbance

---

21 See, for example, Smets and Wouters (2007).
in accounting for most of the variance in inflation is present in many mid-scale Bayesian-estimated Monetary DSGE models, such as Smets and Wouters (2007), our interpretation suggests that these New-Keynesian model economies suffer also from the same tight short-run relationship between the growth rates of money and the price level that we have found here.

To break away from the Quantity Theory of Money in the short-run, the three monetary frameworks that we study here need a more sluggish response of the growth rate of prices to changes in the growth rate of money. We are not sure about what causes this sluggish response of prices to changes in money in the real world. But the generally accepted conjecture is that the way money is introduced into the economy most probably makes a difference.

When David Hume first formulated the Quantity Theory of Money he appealed to the thought experiment which he summarized in the following quote, “Were all the gold in England annihilated at once, and one and twenty shillings substituted in the place of every guinea, would money be more plentiful or interest lower? No surely: We should only use silver instead of gold.” (Hume, 1742, Of Interest). A modern version of Hume’s experiment, and one which requires no imagination, occurred in the eleven European countries who decided to “annihilate” their local currencies on January 1st, 2002 and “substitute” them for the Euro. Paraphrasing Hume, was money less plentiful or interest higher in those countries after they changed their currencies? No surely: Their residents only used Euros instead of their former currencies, and went about their business as if nothing much had happened. In this and in other similar real world experiments, when currencies are redenominated by knocking off one or two zeros, for example, the Quantity Theory of Money relationship holds both in the short and in the long run, almost exactly.

But the words “at once” are key in Hume’s quotation. When central banks conduct monetary policy, changes in the quantity of money are not introduced evenly and “at once”. Instead, money is injected into one part of the economy, typically the banking system, and it spreads out gradually from there. A consequence of this relatively slow spreading out of money is that the Quantity Theory of Money fails to hold in the short-run, while the spreading out is taking place.

When money changes are universal and simultaneous —that is, when they affect every agent at the same time— the rate of growth of prices responds immediately to changes in the rate of growth on money. But when money enters the economy at a specific point, it has to spread around from there. The time it takes in this spreading around probably creates the sluggishness. Like the Quantity Theory of Money itself, this idea can also be traced back to David Hume; “[T]he money in its progress through the whole commonwealth...first quicken[s] the diligence of every individual before it encrease the price of labour.” (Hume, 1742, Of Money).

In representative agent model economies there is only one point at which money can enter the economy. Once it reaches this agent, it has nowhere to spread around, and the sluggish response of prices is very hard to achieve. In this type of model economies every change in the growth rate of
money is both universal and simultaneous by construction. This reasoning allows us to conjecture that agent heterogeneity may very well turn out to be a necessary condition for model economies to display the needed sluggishness.

Díaz-Giménez, Prescott, Alvarez, and Fitzgerald (1992) model the role of money as an asset in a heterogeneous household setup, and they give an early quantitative step in what could turn out to be the correct direction. The findings of Alvarez, Atkeson, and Edmond (2009), Telyukova and Visschers (2013), and Williamson (2008), each of which includes a degree of agent heterogeneity, suggest that agent heterogeneity may indeed be key in replicating the sluggishness observed in the data. As far as the Quantity Theory of Money relationship is concerned, the explicit modeling of agent heterogeneity is probably one of the best bets for future research.

References


A The Monetary Model Economies

In this appendix we describe in detail each of the three model economies used as canonical examples for the three monetary frameworks: Cooley and Hansen (1989) for the Cash-in-Advance framework; Galí (2008) Chapter 3 for the New-Keynesian framework; and Aruoba, Waller, and Wright (2011) for the Search-Money framework. We then describe the details of the calibration and computation of all three models.

A.1 The Cash-in-Advance Model Economy

The cash-in-advance abstraction is an explicit way to model the transactions function of money by requiring that at least some goods have to be purchased with cash. This abstraction was first developed and analyzed in Lucas (1980), and more generally by Stokey and Lucas (1983, 1987), although the idea to model frictions in this way dates back to Clower (1967). Quantitative explorations of the business cycle implications of this abstraction can be found in Cooley and Hansen (1989, 1995). In this article, to represent the cash-in-advance abstraction, we use a minor variation of the model economy described in Cooley and Hansen (1989), but in its actual description we follow Nason and Cogley (1994).

In this model economy there are three goods: a consumption good, an investment good, and leisure. We assume that only the consumption good must be bought with cash carried over from the previous period, while the investment good and leisure can be purchased on credit. Model economies with this type of cash-in-advance constraint attempt to account for the distortionary effects of inflation on real activity. These distortions create an incentive for people to substitute away from activities that require cash—from consumption, in our case—towards activities that are exempt from this requirement—towards investment and leisure, in our case.

As shown by Hodrick, Kocherlakota, and Lucas (1991), one of the shortcomings of the cash-in-advance abstraction is that the model economies react too quickly to monetary shocks. Numerous extensions have attempted to deal with this shortcoming by adding liquidity effects via portfolio adjustment costs (see Lucas (1990); Fuerst (1992); Christiano and Eichenbaum (1992); and Christiano and Eichenbaum (1995), amongst others). But these extensions, while they have succeeded in addressing the issue of the liquidity effects, have had very limited success in generating a sluggish response of prices. See Christiano (1991) for an interesting discussion of the motivations, strengths, and weaknesses of the cash-in-advance approach to modelling money.

A.1.1 Households

The economy is inhabited by a continuum of identical households of measure one who order their preferences over stochastic processes of consumption and labor according to the following utility
function:

$$\max E \sum_{t=0}^{\infty} \beta^t \left( \frac{C_t^{1-\sigma}}{1-\sigma} - \gamma N_t^{1+\varphi} \right)$$ \quad (10)$$

where $0 < \beta < 1$ is the discount factor, $C_t$ is consumption, and $N_t$ is labor\textsuperscript{22}. Households in this model economy are endowed with one unit of time which they can allocate to the supply of labour services to the firm or to the enjoyment of leisure, that is $N_t \in [0,1]$ for all $t$. The households face a budget constraint given by

$$P_tC_t + P_tX_t + M_t \leq P_tW_tN_t + P_tR_tK_t + M_{t-1} + T_t$$ \quad (11)$$

where $P_t$ is the price level, $X_t$ is investment in capital, $M_t$ are money holdings, $W_t$ is the real wage, $R_t$ is the real interest rate, $K_t$ is capital holdings, and $T_t$ is the lump-sum transfer of the cash injections made by monetary authorities.

The stock of capital evolves according to

$$K_{t+1} = (1-\delta)K_t + X_t$$ \quad (12)$$

where $0 < \delta < 1$ is the depreciation rate.

The innovation of the cash-in-advance abstraction that makes money necessary is to add a cash-in-advance constraint. This constraint requires that the consumption good must be purchased with money, in particular with money that must be ‘held in advance’. That is, with money holdings that are chosen one period ahead plus the money injected into the economy in the current period. This cash-in-advance constraint is

$$P_tC_t \leq M_{t-1} + T_t$$ \quad (13)$$

The process on money defined later, following Cooley and Hansen (1989), will make the cash-in-advance constraint always binding\textsuperscript{23}.

Therefore, the problem of the representative household is to choose $C_t$, $N_t$, $M_t$, $X_t$ and $K_t$ in order to maximize (10) subject to (11), (12), (13), and $N_t \in [0,1]$.

\textbf{A.1.2 Firms}

Firms in the economy operate in competitive factor and product markets and produce output according to a constant returns-to-scale production function. These assumptions allow us to use a

\textsuperscript{22}The utility function in Cooley and Hansen (1989) is $\log(C_t) - \gamma N_t$, which is a subcase of ours.

\textsuperscript{23}This assumption, that the cash-in-advance constraint always binds, was shown to be unconsequential by Hodrick, Kocherlakota, and Lucas (1991), since when it is allowed to be occasionally binding it remains the case that for quantitatively plausible calibrations it will bind almost all of the time anyway.
representative firm with a production function that takes the following form

\[ Y_t = A_t K_{f,t}^\alpha N_{f,t}^{1-\alpha} \]  

(14)

where \( Y_t \) is output, \( K_{f,t} \) and \( N_{f,t} \) are the capital and labour inputs, and \( A_t \) is a technology shock. Each period \( t \) the firm's decision problem written in real terms is

\[
\max_{Y_t,K_{f,t},N_{f,t}} Y_t - W_t N_{f,t} - R_t K_{f,t}
\]

(15)

The technology shock follows an exogenous AR(1) process in logs, given by

\[ a_t = \rho a_{t-1} + \varsigma_t \]

(16)

where \( a_t \equiv \log(A_t) \), \( \varsigma_t \) is an identically and independently distributed process that follows a normal distribution with zero mean and variance \( \sigma^2 \).

A.1.3 Money

The monetary authority of this economy issues non-interest bearing currency, \( M^s \), according to the following rule

\[ M_{t+1}^s = e^{\nu_t} M_t^s \]

(17)

where the stochastic money growth rate, \( \nu_t \), is revealed at the beginning of period \( t \) and evolves according to

\[ \nu_t = (1 - \rho_m) \bar{\nu} + \rho_m \nu_{t-1} + \xi_t \]

(18)

where \( 0 < \rho_m < 1 \) and where \( \xi_t \) is an identically and independently distributed process that follows a normal distribution with zero mean and variance \( \sigma^2 \).

Given the money supply rule, the government makes the required money injections to implement it each period. These injections take the following form

\[ T_t = M_{t+1}^s - M_t^s \]

(19)

and are given as lump-sum payments to the households, adding directly to their money holdings.

A.1.4 Prices and Market Clearance

Prices in this model economy are completely flexible and they adjust instantaneously so that labor, capital and money markets always clear. That is,

\[
N_t = N_{f,t} \\
K_t = K_{f,t} \\
M_t = M_t^s
\]

(20)
A.1.5 Equilibrium

To solve the model it must first be made stationary. The first step to achieve this is to divide equations (11) and (13) by the price level, $P_t$. The second step is to replace $M_t$ and $P_t$ in those two equations with $\hat{M}_t = M_t/M_t^s$ and $\hat{P}_t = P_t/M_t^s$, this allows us to remove the trending variables $M_t$, $P_t$.

Once the problem is stationary, the equilibrium of the cash-in-advance model economy can be characterized by the following system of equations that combines optimality conditions, budget and technology constraints, and market clearing conditions.

\begin{align*}
    K_{t+1} + \hat{M}_t/\hat{P}_t &= W_t N_t + (R_t + 1 - \delta)K_t \quad (21) \\
    C_t &= \frac{\hat{M}_{t-1} + e^{\nu_t} - 1}{e^{\nu_t} P_t} \quad (22) \\
    W_t &= (1 - \alpha) e^{a_t} K_{t-1}^{\alpha} N_t^{-\alpha} \quad (23) \\
    R_t &= \alpha e^{a_t} K_{t-1}^{\alpha-1} N_t^{1-\alpha} \quad (24) \\
    \frac{N_t^p}{W_t} &= \beta E \left\{ \frac{N_{t+1}^p}{W_{t+1}} (R_{t+1} + 1 - \delta) \right\} \quad (25) \\
    \frac{W_t}{N_t^p} &= \frac{\gamma}{\beta} E \left\{ C_{t+1} e^{\nu_{t+1}} \frac{\hat{P}_{t+1}}{P_t} \right\} \quad (26) \\
    \hat{M}_t &= 1 \quad (27) \\
    a_t &= \rho_a a_{t-1} + \varsigma_t \quad (28) \\
    \nu_t &= (1 - \rho_m) \bar{\nu} + \rho_m \nu_{t-1} + \xi_t \quad (29)
\end{align*}

A.1.6 The Quantity Theory of Money in a Single Equation

We now describe with a single equation the quantity theory of money in way that makes it easier to see how the Cash-in-Advance framework temporarily escapes from the quantity theory of money. Specifically, we give an expression for the term $PY/M$. Were the Quantity Theory of Money to hold exactly this term would be equal to a constant.

Using the cash-in-advance constraint, $P_t C_t = M_{t-1} + T_t$, with the aggregate resource constraint, $Y_t = C_t + X_t$, we have

\[
\frac{P_t Y_t}{M_t} = \frac{P_t(C_t + X_t)}{M_t} = \frac{M_{t-1} + T_t}{M_t} + \frac{P_t X_t}{M_t} = 1 + \frac{P_t X_t}{M_t} \quad (30)
\]

So the Cash-in-Advance framework succeeds in breaking away from the Quantity Theory of Money in-so-far as monetary policy distorts investment decisions (distorts the cash goods vs. credit goods margin).
A.2 The New-Keynesian Model Economy

To represent New-Keynesian abstraction we use the model economy described in Chapter 3 of Galí (2008). If money were absent, both the cash-in-advance model economy described above and the New-Keynesian model economy described below would simplify to similar versions of the standard real business cycle model economy.

The main purpose of New-Keynesian model economies is to analyze monetary policy. These model economies use sticky prices, which they justify with a mixture of theoretical justifications like rational inattention with empirical evidence that prices change infrequently. Sticky prices allow money to have short-run effects, while remaining long-run neutral. The New-Keynesian approach is perhaps more interested in modelling the effects of monetary policy on the economy, than in the modelling of money itself.

In the subsections below we discuss a version of the text-book description of the basic New Keynesian model economy which we have taken from Chapter 3 of Galí (2008). Even though this model economy can be characterized fully by a system of equations obtained by log-linearization about the steady-state of an explicit model economy, we provide the details of the full model economy to highlight its similarities with the cash-in-advance economy that we have just described.

This model economy has a representative household and it assumes that prices are sticky and that they change according to a Calvo rule, (Calvo, 1983). From these micro-foundations we derive the New Keynesian Phillips curve and the dynamic Investment-Savings equation. To close the model we add a process on nominal interest rates and a money demand function that define the monetary policy rule and the relationship between the money supply and the interest rate.

A.2.1 Households

The model has a representative household who chooses consumption, labor, and savings so as to maximize an expected discounted utility function

$$\max E \sum_{t=0}^{\infty} \beta^t \left( \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} \right)$$

(31)

where $0 < \beta < 1$ is the discount factor, $C_t$ is consumption, and $N_t$ is labor. Note that this is identical to the utility function in expression (10) for the cash-in-advance model.

However, in this model economy we assume that the household consumes a continuum of goods indexed by $i \in [0, 1]$. These goods are transformed into a composite good according to the following equation

$$C_t = \left[ \int_0^1 C_t(i)^{\frac{1}{1+\varphi}} \, di \right]^{\frac{1}{1+\varphi}}$$

(32)
In this model economy the maximization of expected discounted utility is subject to the following series of budget constraints,

\[
\int_0^1 P_t(i) C_t(i)di + I_tB_t \leq B_{t-1} + P_tW_tN_t + T_t
\]  

(33)

where \(B_t\) are purchases of nominal one-period bonds which have gross rate of return \(I_t\), \(W_t\) is the wage, \(T_t\) is a lump-sum component of income, which may include dividends from firm ownership, and \(P_t\) is the aggregate price level which is given by

\[
P_t = \left[ \int_0^1 P_t(i)^{1-\epsilon} di \right]^{1/\epsilon}
\]  

(34)

The representative household demands money according to a money demand function that depends on the nominal interest rates. However it is more convenient to write this demand function in logs and we provide it in expression (38) below.

### A.2.2 Firms

Each differentiated consumption good is produced by a different firm. All firms have the same production technology given by \(Y_t(i) = A_tN_t(i)^{1-\alpha}\), where \(Y_t(i)\) is the production of firm \(i\), \(A_t\) is a common technology level, and \(N_t(i)\) is the labour used by firm \(i\). The firms set prices a la Calvo, that is, each period firms are allowed to change prices only with probability \(1 - \theta\). Firms set prices to maximize their expected discounted future profits for the period in which that price is in place. Thus, problem for firm setting price in period \(t\) is

\[
\max_{P^*_t} \sum_{k=0}^{\infty} \theta^k E_t\{I_{t,t+k}(P^*_tY_{t+k|t} - \Psi_{t+k}(Y_{t+k|t}))\}
\]  

(35)

subject to a demand function

\[
Y_{t+k|t} = \left( \frac{P^*_t}{P_{t+k}} \right)^{-\epsilon} C_{t+k}
\]  

(36)

which comes from the first-order conditions of the representative agents problem. Where \(Y_t = \left( \int_0^1 Y_t(i)^{1-\epsilon} di \right)^{1/1-\epsilon}\) is production of the final (composite) good, \(P^*_t\) is the price being set, \(\Psi_{t+k}(\cdot)\) is the cost function, \(Y_{t+k|t}\) is the production at time \(t+k\) of a firm that last changed price in period \(t\), \(I_{t,t+k}\) is the stochastic discount factor for nominal payoffs, and \(P_t = \left[ \int_0^1 P_t(i)^{1-\epsilon} di \right]^{1/\epsilon}\) is the aggregate price level.

The technology process, \(A_t\), follows an AR(1) process in logs, \(a_t\),

\[
a_t = \rho_a a_{t-1} + \varsigma_t
\]  

(37)

where \(\rho_a \in [0,1)\), \(\varsigma_t\) is iid \(\mathcal{N}(0,\sigma^2_\varsigma)\).
A.2.3 Money

The money demand function in logs is

\[ m_t - p_t = y_t - \eta_i t \] (38)

where \( m_t \) is (log) money, and \( p_t \) are (log) prices.

Monetary policy, in keeping with all of the models covered in this paper is given by an exogenous AR(1) process,

\[ \nu_t = (1 - \rho_m)\bar{\nu} + \rho_m \nu_{t-1} + \xi_t \] (39)

where \( \nu_t \equiv \Delta m_t \), \( \rho_m \in [0, 1) \), \( \xi_t \) is white noise. Note that the process on money in the New Keynesian model (equation (39)) in exactly the same one as was used the Cash-in-Advance model (equations (17) & (18)), just that here we write the process with \( m_t \) in logs.

A.2.4 Prices and Market Clearance

The evolution of the aggregate consumer price level is given by

\[ P_t = \left[ \theta P_{t-1}^{1-\epsilon} + (1 - \theta)(P_{t}^*)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \] (40)

Thus consumer price inflation is

\[ \Pi_t^{1-\epsilon} = \theta + (1 - \theta) \left( \frac{P_t}{P_{t-1}} \right)^{1-\epsilon} \] (41)

where \( \Pi_t = P_t/P_{t-1} \) is the consumer price inflation rate.

The remaining component of the model is the requirement for market clearing. The market clearing conditions are given by, \( \forall t \): that the markets for each consumption good clear, \( C_t(i) = Y_t(i) \), \( \forall i \in [0, 1] \), and that the labour market clears, \( N_t = \int_0^1 N_t(i)di \).

A.2.5 Equilibrium

The system of equations that constitute the reduced form of the basic New Keynesian model are now given\(^\text{24}\). They are derived from the microfoundations listed previously. The difference in notation, with lowercase letters replacing the uppercase letters, is that all of the variables listed here are now in log-linear form rather than the levels represented by the uppercase letters, eg. \( y_t \) is log-deviation of output while \( Y_t \) is output. The New Keynesian Phillips curve is given by

\[ \pi_t = \beta E_t\{\pi_{t+1}\} + \kappa y_t \] (42)

\(^\text{24}\)This system of equations already incorporates the parametrization of \( \gamma = 1 \), to which the models are later calibrated.
where \( \kappa \equiv (\sigma + \frac{\varphi + \alpha}{1 - \sigma}) \frac{(1 - \beta)(1 - \beta \theta)}{\beta} \Theta \), and \( \Theta \equiv \frac{1 - \alpha}{1 - \alpha + \alpha \sigma} \leq 1 \); \( \pi_t \) is the inflation rate, and \( \bar{y}_t = y_t - y^n_t \) is the output gap, that is the difference between current output, \( y_t \), and the natural level of output \( y^n_t \) which would occur if prices were flexible. The dynamic IS equation is

\[
\bar{y}_t = -\frac{1}{\sigma} (i_t - E_t \{ \pi_{t+1} \} - r^n_t) + E_t \{ \bar{y}_{t+1} \}
\]

(43)

where \( i_t \) is the nominal interest rate, \( r^n_t \) is the natural interest rate (again, that which would result if prices were flexible). Both of these two equations are derived from the models micro-foundations.\(^{25}\)

Letting \( l_t = m_t - p_t \) be real money holdings and rewriting the money market equilibrium condition as \( \bar{y}_t - \eta i_t = l_t - y^n_t \), we can substitute out for \( i_t \) and get the following system of equations from the three above,

\[
\begin{align*}
\pi_t & = \beta E_t \{ \pi_{t+1} \} + \kappa \bar{y}_t \\
(1 + \sigma \eta) \bar{y}_t & = \sigma \eta E_t \{ \bar{y}_{t+1} \} + l_t + \eta E_t \{ \pi_{t+1} \} + \eta r^n_t - y^n_t \\
l_{t-1} & = l_t + \pi_t - \Delta m_t
\end{align*}
\]

(44)-(45)-(46)

where \( \hat{r}^n_t \) is the deviation from steady-state of the natural rate of interest.

The two other formulae necessary to complete the model are those for the natural level of output and the natural rate of interest expressed in terms of deviation from steady-state, both of which depend on the technology level.

\[
\begin{align*}
y^n_t & = \varphi^n_n a_t + \vartheta^n_y \\
\hat{r}^n_t & = -\sigma \vartheta^n_y a_t (1 - \rho_a) a_t
\end{align*}
\]

(47)-(48)

where \( \vartheta^n_y = -\frac{(1 - \alpha)(\mu - \log(1 - \alpha))}{\sigma (1 - \alpha) + \varphi + \alpha} > 0 \) and \( \psi^n_y = \frac{1 + \varphi}{\sigma (1 - \alpha) + \varphi + \alpha} \). The model is thus the system of equations given by (44)-(48) together with the processes on the changes in the money supply (39) and technology shocks (37).

A.2.6 The Quantity Theory of Money in a Single Equation

We now describe with a single equation the Quantity Theory of Money in way way that makes it easier to see how the New-Keynesian framework temporarily escape from the Quantity Theory of Money. Specifically, we give an expression for the term \( \frac{P}{M} \). Since the New-Keynesian model is log-linearized we will look at the log of this term, namely \( p + y - m \). Were the Quantity Theory of Money to hold exactly this term would be equal to a constant.

First observe that simply rewriting the money demand equation, (38), we get

\[
p_t + y_t - m_t = \eta i_t
\]

(49)

where $i_t$ is the nominal interest rate and $\eta$ is the elasticity of money demand. Combining the New-Keynesian Phillips Curve, equation (44), with the dynamic IS, equation (43), we get that

$$i_t = r^m_t - \frac{\sigma}{\kappa} \pi_t + (1 + \beta \kappa - \frac{\sigma}{\kappa}) E_t\{\pi_{t+1}\} + \frac{\beta}{\kappa} E_t\{\pi_{t+2}\}$$

(50)

So the nominal interest rate depends on the natural real rate of interest $r^m_t$ (which depends on the current technology shock) and current and future expected inflation. Thus we have that

$$p_t + y_t - m_t = \eta r^m_t + \eta E_t\{f(\pi_t, \pi_{t+1}, \pi_{t+2})\}$$

(51)

Now, $r^m_t$ is independent of monetary factors and the parameters relating to sticky prices. So for sticky prices to be part of the story they must be operating through inflation. From equation, (41), we have inflation evolves as

$$\pi_t = (1 - \theta)(p^*_t - p_{t-1})$$

(52)

where $p^*_t$ is the price level being chosen by those firms that get to reset their prices. So in principle, sticky prices may affect the rate of inflation, and thus the nominal interest rates — helping to break away from the Quantity Theory of Money. In practice however the effect quantitatively negligible.

A.3 The Search-Money Model Economy

The aim of Search-Money models is to provide structural reasons that justify the existence of money. This abstraction focuses on money as a facilitator of exchange based on the idea that money exists mainly to solve problems related to the presence of single-coincidence of wants. Search-Money models go a step deeper than the other two abstractions that we consider here, in which money exists simply because the modeler assumes that it does, rather than to solve an explicit problem; like the absence of a double coincidence of wants in exchange. For this reason the model is the only one of the three we consider that satisfies Wallace’s Dictum for monetary economics, that “Money should not be a primitive in monetary theory — in the same way that firm should not be a primitive in industrial organization theory or bond a primitive in finance theory” (Wallace, 1998).

Search-Money models have become more popular in recent years as they have begun to overcome some teething problems that plagued them in their earlier days: for instance in Kiyotaki and Wright (1989) money holdings were restricted to being 0 or 1 units per agent. Lagos and Wright (2005) overcame these issues by introducing the concept of a centralized (Arrow-Debreu) night-market alongside the decentralized (Kiyotaki-Wright) day-market. The use of the night-market remains integral to the latest generation of Search-Money models such as Head, Liu, Menzio, and Wright (2012) and Berentson, Menzio, and Wright (2011). To represent the Search-Money abstraction we use a stochastic extension of the model economy described in Aruoba, Waller, and Wright (2011) — the stochastic extension is necessary to allow us to use the same process on money
growth as in the other models\textsuperscript{26} The model of Aruoba, Waller, and Wright (2011) uses the same combination of decentralized day-market and competitive night-market as Lagos and Wright (2005) and incorporates physical capital.

In this model economy there are continuum of agents, a decentralized day-market, and a centralized night-market. Money is essential in the day-market because meetings are anonymous, and credit is precluded in a fraction of these meetings because there is no possibility of credibly promising to repay at a later date. As a result exchange must be quid pro quo and so without money some trades would never take place — namely, those in which there was no double-coincidence of wants. Capital investments are made during the competitive night-market, and capital is used in production during both markets\textsuperscript{27}. The model of Aruoba, Waller, and Wright (2011) includes a government sector, we eliminate this, which requires some recalibration of the model\textsuperscript{28,29}.

\subsection*{A.3.1 Households}

There is a continuum of households indexed by $i$ who live forever and whose measure we normalize to 1. Time is discrete and households discount the future at rate $\beta \in (0, 1)$. Each period is divided into two subperiods which are commonly referred to as “day” and “night”. Households consume and supply labour in both subperiods, and their preferences over sequences of consumption and labor are ordered according to the following period utility function

\begin{equation}
U(c, n, C, N) = u(c) - h(n) + U(C) - N
\end{equation}

where $c$ and $C$ denote consumption and $n$ and $N$ denote labour in the day and night subperiods. Assume that $u, h,$ and $U$ are twice continuously differentiable with $u' > 0, h' > 0, U' > 0, u'' < 0, h'' \geq 0$ and $U'' \leq 0$. Also, $u(0) = c(0) = 0$, and suppose that there exists $q^* \in (0, \infty)$ such that $u'(q^*) = h'(q^*)$ and $C^* \in (0, \infty)$ such that $U'(C^*) = 1$ with $U(C^*) > C^*$.

Aruoba et al. (2011) propose to use the following functional form to take the model to the data

\begin{equation}
U(c, n, C, N) = \left\{ \left[ (c + \chi(1-\sigma) - \chi(1-\sigma)) / (1 - \sigma) - \gamma n \right] + \Xi \log(C) - N \right\}
\end{equation}

With the exception of the inclusion of parameter $\chi$, $u(c) = [(c + \chi(1-\sigma) - \chi(1-\sigma)) / (1 - \sigma)$ is the same constant elasticity of substitution utility of consumption as the ones we have used in the other

\textsuperscript{26} An earlier version of this paper used the model of Lagos and Wright (2005). This model failed to break away from the Quantity Theory of Money in the short-run, performing much worse than the other models presented here.

\textsuperscript{27} The appearance of capital in both markets is important. Earlier work by Aruoba and Wright (2003) to introduce capital, with capital appearing in only one market, led to the results that the day and night markets could be solved for separately, and thus money had no effect on consumption, investment, or anything else in the competitive night-market.

\textsuperscript{28} Our results are robust to leaving the government sector in the model of Aruoba, Waller, and Wright (2011).

\textsuperscript{29} Aruoba, Waller, and Wright (2011) actually present three models. Here we follow their model 2. Their model 1 is the same model, but with a slightly different calibration. Their model 3 uses ‘competitive search’, setting prices in the decentralized day market by price taking, rather than Nash bargaining. For robustness we tried out using their model 3 and it makes no real difference to the results we found using their model 2.
two models economies; the utility of consumption is $U(C) = \Xi \log(C)$ and the disutility of labor is $h(n) = \gamma n$ in the day market. The assumption that utility is quasi-linear in labour is used by Aruoba et al. (2011) and is necessary to keep the model analytically tractable\footnote{Quasi linearity means that there are no wealth effects in the demand for money, so all agents in the centralized night markets choose the same money holdings. As a robustness test we simulated both the New-Keynesian and Cash-in-Advance models setting the parameters so that utility was quasi-linear in labour (ie. $\varphi = 0, \gamma = 1$; note that $\gamma = 1$ is the value to which this parameter is calibrated in those models anyway.). The effect on the results was negligible.}. 

A.3.2 Production and Trade

The day-good, $c$, comes in many differentiated varieties indexed by $i$. Each household consumes only a subset of these goods. Each household can transform its own labour into one of these goods that the household itself does not consume by the production function $F(K_i, N_i)$, namely household $i$ produces good $i$ which it does not consume. The production function is given by a standard Cobb-Douglas formulation $F(K, N) = K^\alpha N^{1-\alpha}$. Trade during the day is decentralized and anonymous and households are matched randomly in a typical search setup.

For two households $i$ and $j$ drawn randomly, there are three possible trading situations. The probability that one consumes what the other produces, but not vice-versa —and, therefore, there is a single coincidence of wants is $\omega$, and we assume that it is symmetric. Then, the probability that neither one of them consumes what the other one produces is $1 - 2\omega$. In a single-coincidence meeting, if $i$ wants the good that $j$ produces we call $i$ the buyer and $j$ the seller. In a fraction $\varpi$ of single-coincidence meetings the buyer can only pay with money, in the remaining fraction, $1 - \varpi$, the buyer has access to credit, $l$. By assumption capital can not be used for transaction purposes.

The night good, $C$, comes in a single and homogeneous variety, which is consumed by every household. Each household can transform its own labour into income at the market wage. Trade, during, the night occurs in a centralized Walrasian market. Consequently, the night-good can be purchased on credit. Since money is a good, it can be traded in the night market just like any other good. Investments in capital are also made during the night market.

All the differentiated day-goods and the night-good are perfectly divisible and non-storable, with the exceptions of money and capital which are storable.

A.3.3 Money

In this model economy there is an object called money that is perfectly divisible and storable in any non-negative quantity. The total money stock at time $t$ is $M_t$, and it evolves according to

$$M_{t+1} = e^{\nu_{\varpi+1}} M_t$$

(55)
The monetary injections, \((e^{\nu_{t+1}} - 1)M_t\), are made after the night market closes and they are distributed lump-sum and equally to every household. The rate of growth of money, \(\nu\), follows an AR(1) process given by

\[
\nu_t = (1 - \rho_m)\bar{\nu} + \rho_m\nu_{t-1} + \xi_t
\]

where \(0 < \rho_m < 1\) and where \(\xi\) is an identical and independently distributed process with zero mean and variance \(\sigma^2_\xi\). Although for the equilibrium proofs below we only need to assume that the rate of growth of money follows a first-order Markov process. So the process on money, as given by equations (55) and (56), is identical to that used in the New-Keynesian and Cash-in-Advance models.

**A.3.4 Prices and Market Clearance**

Let \(1/p_t\) be the price of money in the centralized night-market, that is, \(p_t\) is the nominal price of night good \(C\).

In the deterministic version of the model economy described in Aruoba, Waller, and Wright (2011), the only uncertainty comes from the random matching. In the stochastic extension that we use here, the rate of growth of the money supply, \(\nu_t\), is also uncertain. Consequently, in our model economy the decisions of each household at each point in time depend on its current money holdings, \(m\), on it’s capital holdings \(k\), during the centralized night market on it’s earlier borrowing during the day \(l\), and on the aggregate state which is the rate of growth of money, \(\nu\). Therefore, the households’ choices at time \(t\) can be characterized with a value function that has \(m, k, \nu\) as its arguments; as well as \(l\) in the centralized night-market.

Let \(V(m, k, \nu)\) be the value function for a household when it enters the the decentralized day-market, and \(W(m, k, l, \nu)\) its value function when it enters the centralized night-market. Since trade is bilateral in the day-market and the day-good is non-storable, the seller’s production, \(n\), must be equal to the buyer’s consumption, \(c\).

Let \(m\) be money holdings. The the value of trading at the day-market is

\[
V_t(m, k, \nu) = \omega V_t^b(m, kp) + \omega V_t^s(m, k, \nu) + (1 - 2\omega)W_t(m, k, 0, \nu)
\]

where \(V_t^b(m, k, \nu)\) and \(V_t^s(m, k, \nu)\) denote the values to being a buyer and being a seller, as given by

\[
V_t^b(m, k, \nu) = \omega[u(q_b) + W_t(m - d_b, k, 0, \nu)] + (1 - \omega)[u(\tilde{q}_b) + W_t(m, k, l_b, \nu)]
\]

\[
V_t^s(m, k, \nu) = \omega[-c(q_s, k) + W_t(m + d_s, k, 0, \nu)] + (1 - \omega)[-c(\tilde{q}_s, k) + W_t(m, k, -l_s, \nu)]
\]

In these expressions \(q_b\) and \(d_b\) (\(q_s\) and \(d_s\)) denote the quantity of goods and money exchanged when
buying (selling) for money, while \( \tilde{q}_b \) and \( l_b \) \((\tilde{q}_s \text{ and } -l_s)\) denote the quantity and the value of the loan for the buyer (seller) when trading on credit.

At the centralized night-market agents solve the following problem

\[
W_t(m, k, l, \nu) = \max_{C, N, m', k'} \{U(C) - N + \beta E_t\{V_{t+1}(m' + (\epsilon' - 1)M, \nu')|\nu\} \}
\]

subject to \( C = wN + (1 + r - \delta)k - k' + \frac{m - m'}{p}, C \geq 0, 0 \leq N \leq \tilde{N}, \) and \( m' \geq 0, \) where \( \tilde{N} \) is the endowment of night-hours, \( w \) is the wage, and \( r \) is the interest rate on capital.\(^{31}\)

It is assumed that the markets for captial and labour in the night-market are competitive, thus \( w = F_N(K, N) \) and \( r = F_K(K, N). \)

Now that we have defined the value functions, we consider the terms of trade in the decentralized day-market. In single-coincidence meetings, we use the generalized Nash solution in which the buyer has bargaining power \( \zeta > 0 \) and threat points which are given by the continuation values. In the fraction \( \pi \) of meetings where money is used \((q, d)\) is the consumption for money exchange pair that maximizes the following problem

\[
(q, d) = \arg\max \{ [u(q) + W_t(m_b - d, k_b, 0, \nu) - W_t(m_b, k_b, 0, \nu)]^\zeta \\
- [c(q, k_s) + W_t(m_s + d, k_s, 0, \nu) - W_t(m_s, k_s, 0, \nu)]^{1-\zeta} \}
\]

subject to \( d \leq m_b \) and \( q \geq 0. \) In the remaining fraction, \( 1 - \pi, \) of meetings where credit is available, \((\tilde{q}, l)\) is determined just like \((q, d)\), except that the Nash bargaining problem is no longer any constraint on \( l, \) the way \( d \leq m_b \) had to hold in monetary trades.

As Aruoba, Waller, and Wright (2011) observe, the solution to the bargaining problem in 61 will involve \( d = m_b. \) Substituting this into the bargaining problem and taking the first order condition with respect to \( q \) we have

\[
\frac{m_b}{p} = \frac{z(q, k_s)w}{\gamma}
\]

where

\[
z(q, k) \equiv \frac{\zeta c(q, k)u'(q) + (1 - \zeta)u(q)c_q(q, k)}{\zeta u'(q) + (1 - \zeta)c_q(q, k)}
\]

reflects the terms of trade in the bargaining meetings.

Real output, \( Y = Y_D + Y_N, \) is the combination of real output in the decentralized day market, \( Y_D = \omega\pi M/p + \omega\pi\omega l/p, \) and real output in the centralized night market \( F(K, N). \)

Following Aruoba, Waller, and Wright (2011) we measure inflation in terms of the price level in the centralized market \( p_t. \)^{32}

\^[31] Notice that \((\epsilon' - 1)M\) is the transfer of money that is added lump-sum to the households’ holdings after they exit the night-market.

\^[32] We also tried using a Laspeyres measure of inflation that included prices in the decentralized markets. But
A.3.5 Equilibrium

The system of equations that defines an equilibrium is now given. To make the model stationary we define $\hat{m}_t = m_t/M_t$ and $\hat{p}_t = p_t/M_t$; observe that in equilibrium it follows that $\hat{m}_t = 1$ for all $t$. The derivation of this system of equations follows almost exactly as described in Aruoba, Waller, and Wright (2011). The first three equations are related to the first-order conditions of the household.

$$z(q_t, K_t) = \beta E \left\{ \frac{z(q_{t+1}, K_{t+1})}{\exp(\nu_{t+1})} \left( 1 - \omega \zeta + \omega \zeta \frac{u'(q_{t+1})}{z(q_{t+1}, K_{t+1})} \right) \right\}$$  \hspace{1cm} (64)

$$U'(C_t) = \beta E \{U'(C_{t+1})[1 + F_K(K_{t+1}, N_{t+1}) - \delta] - \omega [\varpi \Gamma(q_{t+1}, K_{t+1}) + (1 - \varpi)(1 - \zeta)c_k(\hat{q}_{t+1}, K_{t+1})] \}$$  \hspace{1cm} (65)

$$U'(C_t) = \frac{1}{F_N(K_t, N_t)}$$  \hspace{1cm} (66)

The fourth equation is aggregate resource constraint

$$C_t = F(K_t, N_t) + (1 - \delta)K_t - K_{t+1}$$  \hspace{1cm} (67)

The next two equations determine the price level in the competitive night market, and the real value of the credit loans made in the decentralized day market (in the fraction $\varpi$ of meetings where credit is available):\footnote{While neither $l_t$ nor $p_t$ are stationary, by treating $l_t/p_t$ as a single variable the equation is stationary.}

$$\hat{p}_t = \gamma \frac{z(q_t, K_t)}{F_N(K_t, N_t)}$$  \hspace{1cm} (68)

$$\frac{l_t}{p_t} = F_N(K_t, N_t)[(1 - \zeta)u(\hat{q} + \zeta c(\hat{q}, K)]$$  \hspace{1cm} (69)

The next four equations are related to the terms of trade in the decentralized day market ($z(q, K)$, as defined in (63)), and some related derivatives and quantities.

$$z(q_t, K_t) = \frac{\zeta c(q_t, K_t)u'(q_t) + (1 - \zeta)u(q_t)c_q(q_t, K_t)}{\zeta u'(q_t) + (1 - \zeta)c_q(q_t, K_t)}$$  \hspace{1cm} (70)

$$z_q(q_t, K_t) = \frac{u'(q) c_q[\zeta u'(q) + (1 - \zeta)c_q] + \zeta(1 - \zeta)(u(q_t) - c)(u'(q_t)c_{qq} - c_qu''(q_t))}{[\zeta u'(q_t) + (1 - \zeta)c_q]^2}$$  \hspace{1cm} (71)

$$z_K(q_t, K_t) = \frac{\zeta u'(q_t)c_k[\zeta u'(q_t) + (1 - \zeta)c_q] + \zeta(1 - \zeta)(u(q_t) - c)u'(q_t)c_{Kq}}{[\zeta u'(q_t) + (1 - \zeta)c_q]^2}$$  \hspace{1cm} (72)

$$\Gamma(q_t, K_t) = c_K(q_t, K_t) - c_q(q_t, K_t)\frac{z_K(q_t, K_t)}{z_q(q_t, K_t)}$$  \hspace{1cm} (73)

since in the calibrated model the decentralized market accounts for only about 3% of total real output this made no noticeable difference.
where $c$ is shorthand for $c(q_t, K_t)$, $c_q$ for $c_q(q_t, K_t)$, $c_K$ for $c_K(q_t, K_t)$, $c_{qq}$ for $c_{qq}(q_t, K_t)$, and $c_{qK}$ for $c_{qK}(q_t, K_t)$. The next equation is simply the definition of real output,

$$Y = F(K, N) + \omega \omega M/p + \omega \omega \omega l/p \quad (74)$$

The final equation is that defining the money growth rate,

$$\nu_t = (1 - \rho_m)\bar{\nu} + \rho_m \nu_{t-1} + \xi_t \quad (75)$$

The Search-Money model with capital is thus given by the system of stochastic difference equations, (64)-(75).

### A.3.6 The Quantity Theory of Money in a Single Equation

We now describe with a single equation the Quantity Theory of Money in way way that makes it easier to see how the Search-Money framework temporarily escapes from the Quantity Theory of Money. Specifically, we give an expression for the term $PY/M$. Were the Quantity Theory of Money to hold exactly this term would be equal to a constant.

In the Search-Money model, by equation (68), we have that

$$\frac{PY}{M} = \frac{1}{z(q, K)} \frac{\gamma Y}{F_N(K, N)} \quad (76)$$

In the simulation results total output $(Y)$, capital stock $(K)$, and the marginal product of labour $(F_N(K, N))$ are almost constant, and thus not related to the ability of the Search-Money model to get away from the Quantity Theory of Money. They are almost constant because most of the economy is based on non-monetary trades, the centralized night market is much bigger than the decentralized day market, and so unaffected by changes in the money supply. All of the movement occurs in the $z(q, K)$ term, specifically from changes in $q$ — the amount produced/traded in the exchanges involving money in the decentralized market. The amount produced in monetary exchanges varies with the amount of money and inflation (the cost of holding money).

### A.4 Calibration and Computation

For our comparisons of the three model economies to be meaningful, we choose their functional forms and parameters so that they are as similar as possible. This use of identical parameter values wherever the models coincide, of identical exogenous processes, and of identical functional forms for the utility of consumption, as well as the fact that we have solved the three model economies using identical solution methods allows us to make a genuine comparison between them. Since we have removed all other possible sources of variation, we can safely attribute any differences in their outputs with respect to the Quantity Theory of Money relationship to the different ways in which these three frameworks model money.
Table 3: Parameter Values

<table>
<thead>
<tr>
<th></th>
<th>Cash-in-Advance</th>
<th>New Keynesian&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Search-Money</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preferences</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time Discount factor&lt;sup&gt;a&lt;/sup&gt;</td>
<td>$\beta$</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Curvature of Consumption&lt;sup&gt;b&lt;/sup&gt;</td>
<td>$\sigma$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Weight on Labour&lt;sup&gt;c&lt;/sup&gt;</td>
<td>$\gamma$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Curvature of Labour&lt;sup&gt;d&lt;/sup&gt;</td>
<td>$\varphi$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returns to Capital&lt;sup&gt;d,e&lt;/sup&gt;</td>
<td>$\alpha$</td>
<td>0.33</td>
<td>n.a.</td>
</tr>
<tr>
<td>Depreciation Rate&lt;sup&gt;e&lt;/sup&gt;</td>
<td>$\delta$</td>
<td>0.025</td>
<td>n.a.</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>$\rho_a$</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Variance of Shock</td>
<td>$\sigma_z$</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td><strong>Money</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elasticity of Money Demand&lt;sup&gt;e&lt;/sup&gt;</td>
<td>$\eta$</td>
<td>n.a.</td>
<td>4</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>$\rho_m$</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Variance of Shock</td>
<td>$\sigma_{\xi}$</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>Constant Term</td>
<td>$\nu$</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td>Dist. of Shock</td>
<td>$\xi$</td>
<td>log-normal</td>
<td>log-normal</td>
</tr>
<tr>
<td><strong>Price Setting</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market Power</td>
<td>$\epsilon$</td>
<td>n.a.</td>
<td>6/5</td>
</tr>
<tr>
<td>Calvo Stickiness</td>
<td>$\theta$</td>
<td>n.a.</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Search</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob. of Single Coincidence&lt;sup&gt;g&lt;/sup&gt;</td>
<td>$\omega$</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Bargaining Power</td>
<td>$\zeta$</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Night weight on consumption</td>
<td>$\Xi$</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Make $u(0) = 0$</td>
<td>$\chi$</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Prob. of credit availability</td>
<td>$\varpi$</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

<sup>a</sup>Every other parameter that appears in the equations that characterize the equilibrium of the New-Keynesian model economy can be derived from the parameters that we have identified in this table using the following system of equations: $M = \epsilon/(1 - \epsilon); \mu = logM; \rho = -log\beta; \Theta = (1 - \alpha)/(1 - \alpha + \alpha\epsilon); \lambda = \Theta(1 - \theta)(1 - \beta\theta)/\theta; \kappa = \lambda[\sigma + (\varphi + \alpha)/(1 - \alpha)]; \theta_{\mu} = (1 - \alpha)(\mu - log(1 - \alpha))/|\sigma(1 - \alpha) + \varphi + \alpha|; \psi_{m} = (1 + \varphi)/|\sigma(1 - \alpha) + \varphi + \alpha|.$

<sup>b</sup>In the Search-Money model this parameter is calibrated 2.1, as this is needed as part of the the calibration procedure of Aruoba, Waller, and Wright (2011) (setting this parameter to one in the Search-Money with capital model, while messing up the calibration, does not affect the results).

<sup>c</sup>In the Search-Money model the disutility of labor is linear in both the day-market and in the night-market.

<sup>d</sup>Abbreviation “n.a.” means “not applicable”.

<sup>e</sup>There is no ‘returns to capital’ or ‘depreciation’ in the New-Keynesian economy as there is no capital.
Parameter Choices

We have decided to use Galí (2008) as our main reference for our parameter choices, with the obvious exceptions of the parameters and functions of the Cash-in-Advance and Search-Money model economies that do not exist in the New-Keynesian framework, such as the parameters related to the search for trading partners in the Search-Money model economy.\(^{34}\)

Since Galí (2008) exploits the certainty equivalence principle in his solution method, he does not define the shocks to either the technology or the money supply. Instead, we take the processes for those shocks from Cooley and Hansen (1989). We report our chosen parameter values in Table 3.\(^{35}\) Our results are robust to using the original parameter calibrations of each model, that is those parameter values given in the papers from which the models are taken. Importantly, the original parameterizations of the models (in the papers from which they are taken) are all calibrated to similar postwar periods. Since the frameworks are quite different using exactly the same calibration targets for the different frameworks is not possible, although some common calibration targets, such as interest rates and capital-output ratios were used by a number of the the original papers.

Simulation

To simulate our model economies we have used identical seeds for the random number generator so that the sequences of the realizations of the random shocks are identical in all three model economies. To obtain the model economy time series we discard the first 200 periods of each equilibrium realization to purge away the initial conditions, and then we draw a sample of 204 quarterly observations to replicate the number of observations in our United States time series. Whenever we need to obtain multiple samples, we repeat this process as necessary.

Computation

The equilibria of the three models economies that we have described above can be reduced to systems of stochastic difference equations. We have solved these systems using the default perturbation methods of Dynare to calculate quadratic approximations to the decision rules.\(^{36}\)

\(^{34}\)The calibrations reported in Aruoba, Waller, and Wright (2011) are annual and so had to be adjusted. This was done using the same methodology and targets they report — some targets, such as the capital-output ratio, have to be adjusted to quarterly values.

\(^{35}\)Galí (2008) pg. 52 says that \(\epsilon_p = 6\), however this appears to be a typo. When we use this value, we fail to replicate his results. Therefore, we use \(\epsilon_p = 6/5\) instead following http://www.dynare.org/phpBB3/tviewtopic.php?f=1&andt=2978. In this case we replicated Gali’s results successfully.

\(^{36}\)We have run every code with Dynare Version 4.4.3 using Matlab 2013a.