b
UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

The Welfare Costs of Inflation Reconsidered

Luca Benati, Juan-Pablo Nicolini

25-08

October, 2025

DISCUSSION PAPERS

Schanzeneckstrasse 1
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

The Welfare Costs of Inflation Reconsidered*

Luca Benati
University of Bern†

Juan-Pablo Nicolini
Federal Reserve Bank of Minneapolis
and Universidad Di Tella‡

Abstract

Modern analysis of the welfare effects of monetary policy is based on moneyless models and therefore ignores the effect of inflation on the efficiency of transactions. A justification for this strategy is that these welfare effects are quantitatively very small, as argued by Ireland (2009). We revisit Ireland's result using recent data for the United States and several other developed countries. Our computations are influenced by the experience of very low short-term rates observed since Ireland's work in the countries we study. We estimate the welfare cost of a steady state nominal interest rate of 5% to be at least one order of magnitude higher than in Ireland (2009), which questions the validity of performing monetary policy evaluation in cashless models.

*We wish to thank Fernando Alvarez and Peter Ireland for comments on a previous draft and for very helpful suggestions. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Minneapolis, or of the Federal Reserve System.

†Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001, Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

‡Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55401, United States. Email: juanpa@minneapolisfed.org

1 Introduction

We provide new estimates of the welfare cost of inflation. We follow the tradition of Bailey (1956), Friedman (1969), Lucas (2000), and Ireland (2009) in that we estimate the welfare cost using the area under the real money demand curve. For a steady-state interest rate of 5%, Lucas (2000) calculates the cost to be 1.1% of lifetime consumption, which is a significant amount. However, Ireland (2009) challenges Lucas's interpretation of the data and obtains an estimate of a mere 0.037%.

Our main contribution is to bring more data to the debate. We do so in two ways. First, we use the additional data available since Ireland's work. This is a particularly abnormal and, at the same time, very interesting period, since it was characterized by several observations with very low interest rates. Thus, it helps identify the behavior of money demand at very low rates, which, as we will discuss, is highly relevant. Second, we also study evidence from developed countries whose inflation histories are similar to those of the United States. This additional evidence is reassuring, since the United States went through regulatory changes during the 1980s and 1990s that blurred the distinction between types of deposits. This started a debate regarding the proper way to measure monetary aggregates, an issue that Ireland emphasized. These issues are absent in the other countries we study.

There are two key aspects of the money demand relationship that affect the computation, as both Lucas and Ireland note. The first is the functional form of the money demand at very low rates. For example, a function with a satiation value at zero, like the semi-log, will tend to deliver lower estimates than one in which desired money balances increase without bound as the nominal interest rate approaches zero, as in the log-log. The second is the values assigned to its parameters.

Our tests using the entire sample tend to prefer a functional form with a finite satiation point, as argued by Ireland. However, out-of-sample tests for very low values of the interest rate prefer the log-log specification, the one preferred by Lucas, in all but one case. To be conservative, we choose the specification with a finite satiation point as our benchmark, but we also report the results for the log-log specification. Regarding the estimated parameters, our results strongly support the numbers preferred by Lucas, both for the United States and for all the other countries.

For our benchmark case in the United States, we obtain a cost of 0.35%, almost ten times that of Ireland. Alternative scenarios deliver higher values, but overall, we find that 0.8% is a likely upper bound.

Modern monetary policy analysis is based on moneyless models, and it therefore ignores the effect of inflation on the efficiency of transactions. For example, the well-known "divine coincidence" case, in which stabilizing prices also stabilize the output gap, is only optimal at the cashless limit.[1] The accuracy of this strategy in computing welfare effects of policy is a quantitative issue. If the welfare cost of distortions on transactions is very small relative to the one arising from price rigidities, then the cashless limit is a sensible approximation. Our calculations suggest that this is not the case.

Nakamura et al. (2018) showed that the welfare cost of inflation is quite small in moneyless New Keynesian models, around 0.02% of consumption for an inflation rate of 3% percent. Given a real rate of 2% this is consistent with a nominal interest rate of 5%, the value considered by Lucas and Ireland. More recently, Afrouzi et al. (2024) show that with network effects, the cost can be much higher, close to 0.4% of consumption. Coibion et al. (2012) studied a model with recurrent, though not very frequent, episodes with the nominal interest rate at the zero lower bound. They computed the welfare effect of an interest rate of 5% to be close to 0.6% of lifetime consumption.[2] Relative to these last two figures, the 0.037% estimated by Ireland may appear negligible. But 0.35%, the lowest number we estimate, is certainly not.

Our starting point is the evidence supporting the notion of a downward and stable long-

[1]Khan et al. (2003) study the trade-off between distortions resulting from price frictions and the ones resulting from lack of money satiation.

[2]Coibion et al. (2012) explicitly acknowledge that they do not take into account the costs derived from lack of money satiation. Taking into account the effect that we study would actually reinforce the argument of their paper.

run real money demand discussed in Lucas and Nicolini (2015) and Benati et al. (2021). In this paper, we go further in several ways. First, we provide formal tests to compare different functional forms. Second, we use recent data with very low nominal interest rates to discipline both the functional form and the parameter estimates. Finally, relative to Lucas (2000) and Ireland (2009), we bring evidence from countries other than the United States to shed light on the question. On the theory side, we innovate by constructing upper and lower bounds for the estimate of the welfare cost of inflation. Thus, we do not need to rely on linear approximations. The area under the money demand curve is an almost exact measure of the welfare cost for a very general class of monetary models in the neighborhood of zero, as Alvarez et al. (2019) show. For a quite general subclass of the models they analyze, we compute exact lower and upper bounds for the estimates of the costs, using the area under the money demand curve, for any value of the interest rate. As we show, the difference between the upper and lower bounds is extremely small for the range of interest rates observed in the United States.

The paper proceeds as follows. In Section 2, we discuss a family of monetary models for which we derive very tight lower and upper bounds for the welfare cost of inflation, using the area under the real-money demand curve. In Section 3, we provide a discussion of our main results, using simple plots, as Lucas (2000) and Ireland (2009) did. Section 4 presents the formal statistical analysis for three different empirical specifications used in the literature, including those Lucas and Ireland explored. Besides estimating the key parameters for each specification, we also develop and perform formal pairwise tests to evaluate the different specifications. Section 5 presents our computations for the welfare cost functions for the benchmark case, in which the lower bound of the interest rate is zero. The exploration of countries other than the United States highlights a feature that we bring to the analysis: the assumption regarding the true lower bound on the short-term nominal interest rate. This is relevant since it determines the lower limit of the integral under the real money demand curve. Both Lucas and Ireland assumed the lower bound to be zero, as did most of the monetary economics literature until 2010. And so did we in Sections 4 and 5. However, the negative interest rates observed in the euro area, Denmark, Sweden, and Switzerland motivate us to reconsider that assumption. We do so in Section 6. Section 7 concludes.

2 The Model

We study a labor-only economy with uncertainty, in which making transactions is costly. The economy is inhabited by a unit mass of identical agents with preferences given by

$$E_0 \sum_{t=0}^{\infty} \beta^t U(c_t),$$

where U is differentiable, increasing, and concave.

Every period, the representative agent chooses a number of portfolio transactions n_t that allow her to exchange interest-bearing bonds for money, which is needed to buy the consumption good. The total cost of those transactions, measured in units of time, is given by a differentiable function $\theta(n_t, \nu_t)$, where ν_t is an exogenous stochastic process. This formulation generalizes the linear function assumed by Baumol (1952) and Tobin (1956).

The production technology for the consumption good depends linearly on time devoted to production. There is a unit of time for each period that can be used to produce goods or make transactions. Thus, equilibrium in the labor market and feasibility imply

$$1 = l_t + \theta(n_t, \nu_t),$$

$$c_t = z_t(1 - \theta(n_t, \nu_t)),$$

where z_t is an exogenous stochastic process. The real wage is then equal to z_t.

Purchases are subject to a cash-in-advance constraint

$$P_t c_t \leq n_t M_t,$$

where M_t is average money balances.

We allow money to pay a nominal return r_t^m. At the beginning of each period, the agent starts with nominal wealth W_t, which can be allocated to money or interest-bearing bonds B_t. This decision, together with the time allocation and consumption decisions, faces the following constraint:

$$M_t + B_t \leq W_t,$$

$$W_{t+1} \leq M_t(1 + r_t^m) + B_t(1 + r_t^b) + T_t + [1 - \theta(n_t, \nu_t)]\, z_t P_t - P_t c_t,$$

where r_t^b is the return on bonds and T_t is a transfer made by the monetary authority.

The unconstrained efficient outcome is to allocate all the labor input to the production of the consumption good so as to set $c_t = z_{t.}$ Thus, the welfare cost of making transactions, as a fraction of consumption, is given by $\theta(n_t, \nu_t)$.

It is straightforward to show (see Online Appendix A for details) that an interior solution for n_t must satisfy

$$n_t^2 \frac{\theta_n(n_t, \nu_t)}{(1 - \theta(n_t, \nu_t))} = r_t^b - r_t^m. \tag{1}$$

As long as $r_t^b - r_t^m > 0$, the cash-in-advance constraint is binding, so real money demand, as a proportion of output, is equal to the inverse of n_t. In what follows, we let the interest rate differential between bonds and money be $r_t \equiv r_t^b - r_t^m$. Note that equation (1) is independent of z_t, so the theory implies a unit income elasticity of real money demand.

For the maximum problem of the agent to be well defined, it has to be the case that $r_t = r_t^b - r_t^m > 0$. The popular zero-bound restriction on policy rates is obtained using the standard assumption in the literature that $r_t^m = 0$. Both Lucas (2000) and Ireland (2009) made this assumption. Recent experiences with negative policy rates in European countries raise the issue of incorporating a negative lower bound. We will do so below, but in what follows, we maintain, for two reasons, the standard assumption that the lower bound on interest rates is zero. The first one is that we want to bring in new data both from the

US and from other countries, to shed light on the discrepancies between Lucas (2000) and Ireland (2009). The second is that negative policy rates do not necessarily imply negative rates on deposits for households and firms, which are the ones relevant for our computations. We briefly discuss these issues below.

The functional form of the real money demand function depends on the functional form of the transactions technology $\theta(n_t, \nu_t)$, and at this level of generality, the model is consistent with many different possibilities. In what follows, to clarify the main difference between Lucas (2000) and Ireland (2009), we consider three well-known functional forms that have been used in previous empirical work. All three exhibit a unit income elasticity, as implied by the model. The first specification is the log-log one,

$$\ln \frac{M_t}{P_t y_t} = a^1 - \eta \ln r_t + u_t^1, \tag{2}$$

which exhibits a constant interest rate elasticity equal to η. Notice that as $i_t \to 0$, real money demand goes to infinity. It is this asymptote at zero that Lucas used to argue that the welfare cost of inflation is sizable, even at low values for the interest rate. The other two formulations that we explore are the semi-log,

$$\ln \frac{M_t}{P_t y_t} = a^2 - \gamma r_t + u_t^2, \tag{3}$$

which exhibits a constant semi-elasticity γ, and the Selden-Latané,

$$\frac{M_t}{P_t y_t} = \frac{1}{a^3 + \phi r_t + u_t^3}. \tag{4}$$

Both formulations imply a finite level of the demand for real money balances when the interest rate differential becomes zero. This feature is emphasized by Ireland, who uses (3) in his revision of Lucas's estimate.

The welfare cost implications of the last two functional forms are similar. We choose to include the Selden-Latané specification because of its very good econometric performance.

The theoretical foundation to computing the welfare cost of inflation using the area under the money demand curve is based on linear approximations around zero nominal interest rates. In a recent contribution, Alvarez et al. (2019) show that to be the case for a very general class of models.[3] A minor contribution of our paper is to provide upper and lower bounds for the welfare cost of inflation for a class of models that is quite general, but less so than the class considered in Alvarez et al. (2019). Our bounds can be used for any value of the interest rate. As we show below, for the low inflation countries we consider, the distance between the upper and lower bounds is positive, but extremely small, so much so that in all the figures below the difference between the two is invisible to the eye.

2.1 Exact bounds for the welfare cost of inflation

In this section, we apply the techniques developed in Alvarez et al. (2019) to a class of models that are more restrictive than the ones they used. Specifically, we consider only representative agent models in which the cost of transforming liquid assets into illiquid ones is given by the differentiable function $\theta(n_t, \nu_t)$,described above. Ignoring time and state dependence, we can write (1) as

$$n^2 \frac{\theta_n(n)}{(1-\theta(n))} = r, \tag{5}$$

where $r \geq 0$. As previously discussed, the welfare cost of inflation, measured as a fraction of consumption, is given by

$$\omega^W(r) = \theta(n(r)), \text{ where } \omega^W(0) = \theta(n(0)) = 0.$$

It follows that

$$\frac{\partial \omega^W(r)}{\partial r} = \frac{\partial \theta(n)}{\partial n}\frac{\partial n}{\partial r}(r) > 0. \tag{6}$$

[3] Alvarez et al. (2019) also show in numerical examples that the approximation is remarkably accurate for a wide range of positive values of the opportunity cost.

Finally, note that the area under the demand curve is equal to

$$\omega^D(r) = \int_0^r m(z)dz - m(r)r. \tag{7}$$

The next proposition states that the function $\omega^W(r)$ can be bounded above and below using the integral under the money demand curve. The proof closely follows the analysis in Alvarez et al. (2019) and appears in Appendix B.

Proposition 1: The welfare cost of inflation $\omega^W(r)$ is bounded above and below by the following transformations of the area under demand curve $\omega^D(r)$ as follows:

$$\frac{\omega^D(r)}{(1+\omega^D(r))} \leq \omega^W(r) \leq \omega^D(r).$$

It is straightforward to see that the bounds are extremely tight. For example, for an opportunity cost equal to 3% of consumption, which is very large, the difference between the upper and the lower bound is equal to about one-tenth of a percentage point.

Explicit closed form solutions for the function $\omega^D(R)$ can be obtained for the three empirical specifications described in (2) to (4), as we show in Appendix B.

3 A Look at the Raw Data

For the empirical analysis, we work with quarterly post-WWII data. The series and their sources are described in detail in Appendix C. For all but one country, we consider M1 as the relevant monetary aggregate. The exception is the United States, for which we follow Lucas and Nicolini (2015) and use NewM1, which, as noted earlier, is obtained by adding Money Market Demand Accounts (MMDAs) to the standard M1 aggregate produced by the Federal Reserve.

Figure 1 shows scatterplots of the ratio between nominal M1 and nominal GDP against a short-term nominal interest rate. We present three groups of countries organized by region.

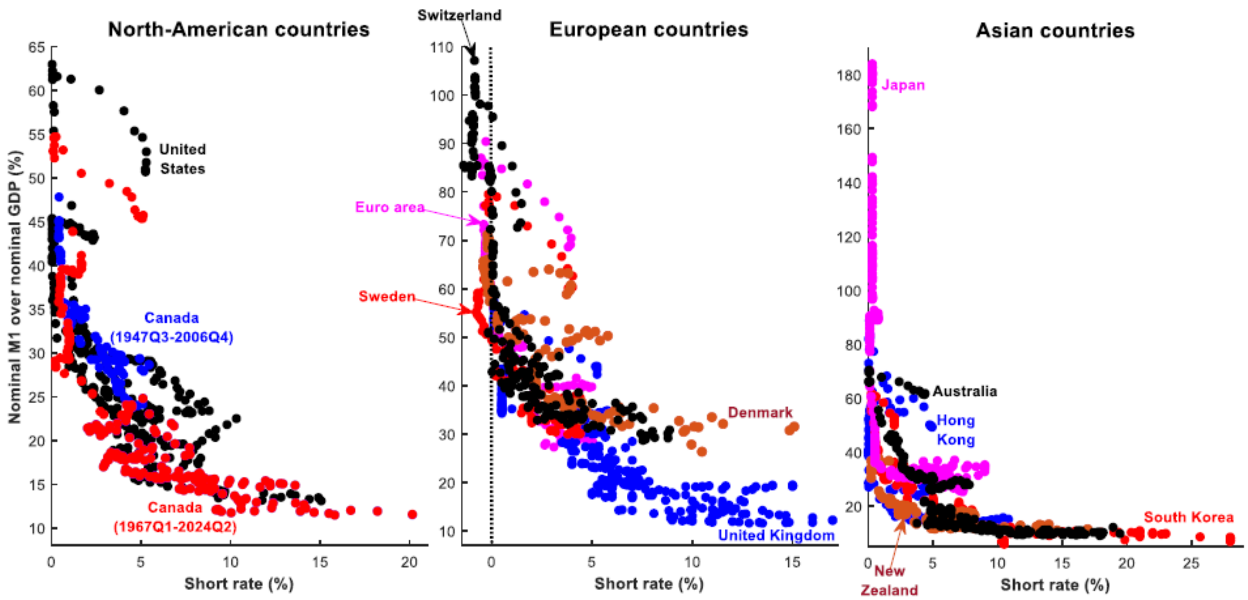

Figure 1: Scatterplots of nominal M1 over nominal GDP against the short rate

The three panels provide strong visual evidence of a negative relationship between the ratio of M1 to GDP and the short-term rate, a hallmark of the theory of real money demand. Comparing the three panels reveals several interesting features. The first is that there appear to be clear and sizable differences in the demand for real money balances across (groups of) countries. Australia and Asian countries exhibit starker heterogeneity: essentially, each country has its own demand curve. Finally, in several areas of Europe, short-term rates have consistently been negative.

3.1 Discussion of the results

In this section, we review money demand evidence using simple figures of the type used by both Lucas and Ireland. Formal econometric analysis will be presented in the section that follows.

Figure 2 shows a cross-plot of annual data for the money stock as a fraction of output and the short-term interest rates for the United States since 1915. It replicates Figure 2 in Lucas (2000), although we use NewM1 as the measure of the money stock. The colored

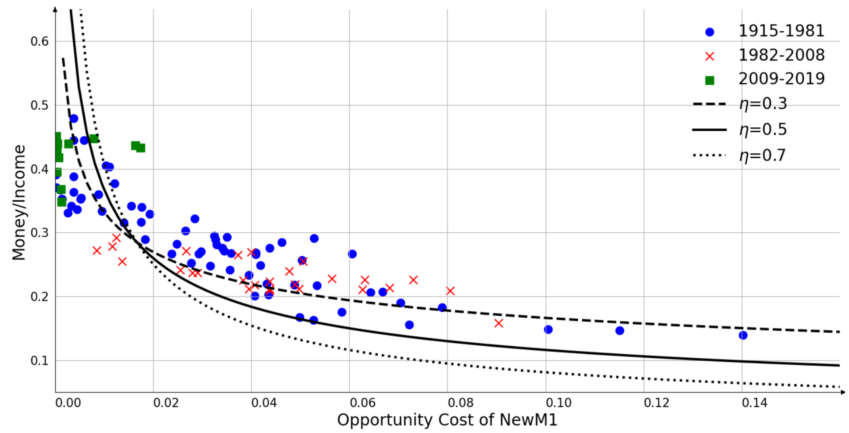

Figure 2: US log-log real money demand, 1915-2019

symbols represent specific sub-periods. The blue circles correspond to the 1915-1981 period. This is the period before the creation of the new deposits that create a divergence between the measures of M1 and NewM1. Using M1 as a measure of liquid assets until 1981 is not controversial. The red X's are obtained using NewM1 as the measure of the money stock. We highlight this period because the measure of money differs from the one used by both Lucas (2000) and Ireland (2009). The green squares correspond to the very low interest rate period that followed the Global Financial Crisis. These data were not available at the time of Ireland's work.

Following Lucas, we also plot three theoretical curves that correspond to the log-log formulation, as defined in equation (2). We show the curves with elasticities equal to 0.7, 0.5, and 0.3. In each case, the constant is chosen so that each curve crosses the point that contains the geometric means of the two series in the sample.

The figure shows that the 0.3 elasticity curve is the one that best matches the data. This conclusion differs from the one obtained by Lucas, who preferred the curve with elasticity equal to 0.5. There are two reasons for the difference. First, for the period 1982 to 1994, Lucas used the aggregate M1, which is substantially lower than NewM1. Aggregate M1 favors values with higher elasticity. Second, the evidence provided by the period with very low interest rates - that is, the period after 2008 (the green squares) - favors the specification with a lower elasticity.[4]

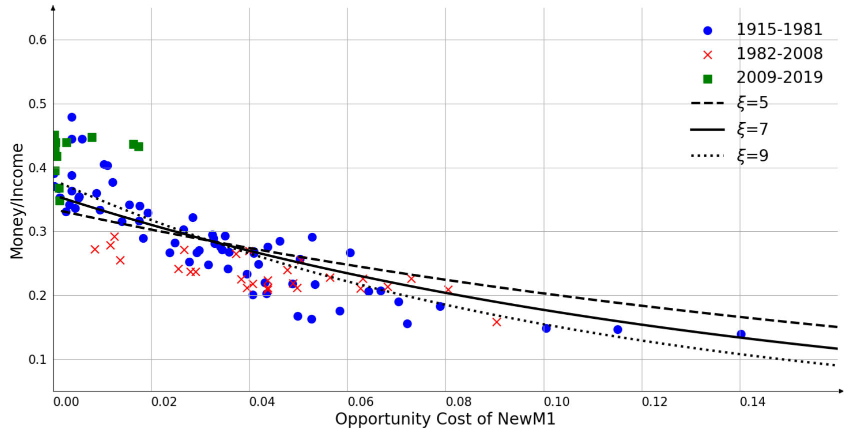

Figure 3: US semi-log real money demand, 1915-2019

Following Lucas again, we show in Figure 3 the same data points, but the theoretical curves correspond to the semi-log specification with semi-elasticities of 5, 7, and 9. From the figure, we conclude that the curve with semi-elasticity equal to 9 marginally provides the best fit. In Lucas's corresponding figure, a value of 7 seems to be the preferred choice.

Finally, in Figure 4 we show the data, together with the two preferred curves in each class. The figure suggests that the log-log performs better at capturing the data with low interest rates and high money-to-output ratios.[5] For intermediate values of the interest rate, the log-log seems to underestimate the money to output ratios, while the semi-log tends to overestimate them. Finally, for the points with higher interest rates, both specifications do reasonably well. This type of evidence led Lucas to pick the log-log over the semi-log.

In revisiting Lucas's results, Ireland first notes that the blue dots in Figure 4, which shows low interest rates and high money-to-output ratios, correspond to the years during WWII. One could certainly entertain the notion that other factors affected the behavior of real money demand in those years. Therefore, Ireland argues, using those years to evaluate the welfare cost of inflation under peacetime could be misleading.

This can be seen in Figure 5, where we plot the time series of NewM1 (solid black line)

[4]The econometric results deliver a value lower than 0.3, as we discuss below.
[5]This feature is consistent with formal out-of-sample tests we perform in Section 4.

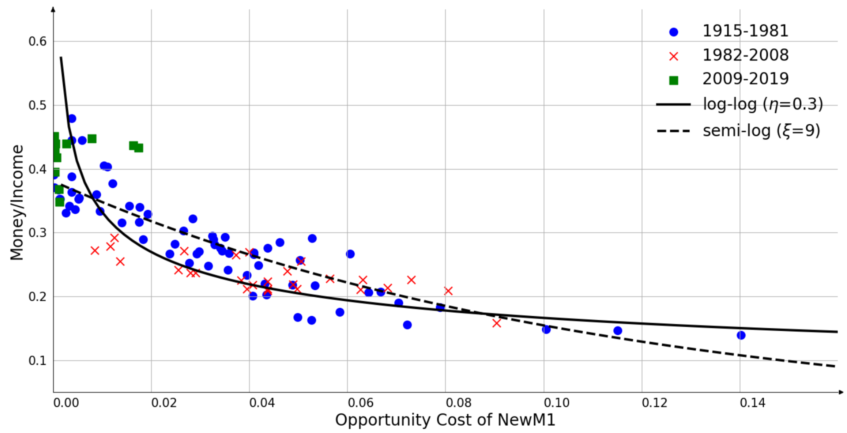

Figure 4: US real money demand, 1915-2019

and the monetary aggregate used by Ireland (solid black line until 1981 and dashed blue line until 2008, the last available data point), together with the short-term nominal interest rate (dotted red line). As Ireland rightly pointed out, the war years are the only ones with low interest rates in the sample used by Lucas. However, since Ireland's analysis, there have been several years with very low interest rates, very similar to those during WWII. And, as Figure 4 shows, the behavior of NewM1 after 2009 (the green squares) is in line with the data during the war used by Lucas. Thus, to the extent that one is willing to accept NewM1 as a correct measure of transactional assets, the years of WWII do not appear as outliers. The case for the log-log Lucas made survives the test of time.

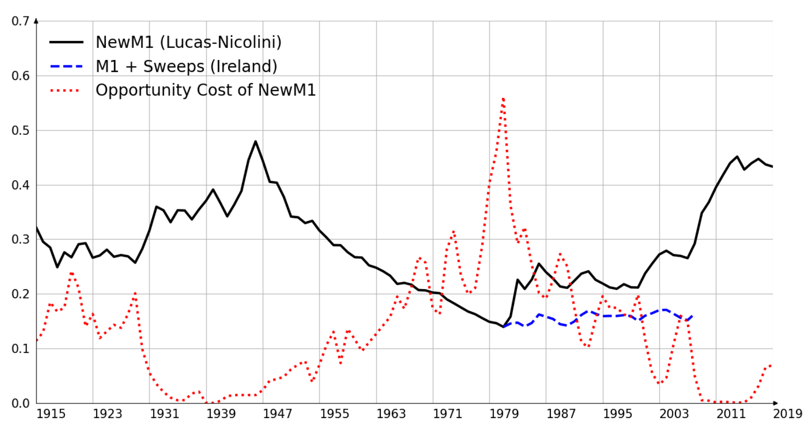

Figure 5: US money-income ratio and short-term interest rate, 1915-2019

Ireland also argued that if one extended the sample in Lucas (2000) until 2008 using M1, a break would become apparent around 1980, even if M1 was adjusted by the sweep programs. Therefore, the argument goes, the pre-1982 evidence was not very useful for estimating the money demand curve. Using post-1980 data alone, he then made two points. The first was that semi-log was the preferred specification. The second was that the semi-elasticity was closer to 1.8, much lower than Lucas's preferred value of 7.

Ireland's argument can clearly be seen in Figure 6, where we plot the same data as before, except that we adopted the definition of the money supply used by Ireland - M1 plus the sweep programs.

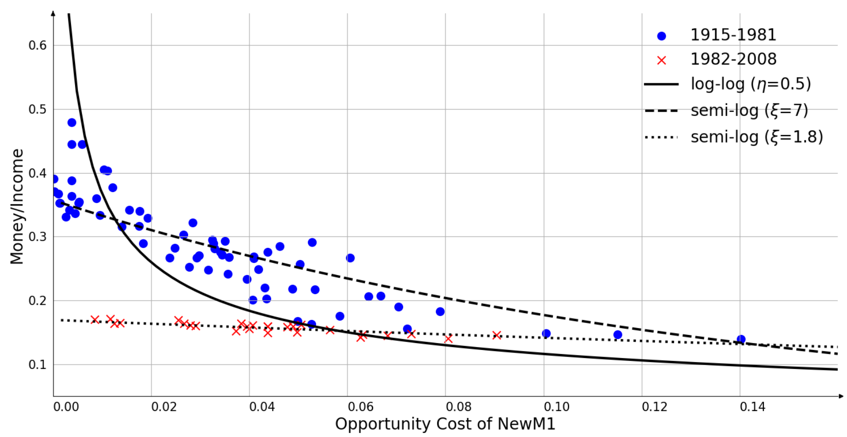

Figure 6: US real money demand, 1915-2008

The figure clearly shows the break in the behavior of the monetary aggregate chosen by Ireland. The data following 1982 - denoted by red X's - line up remarkably well along a semi-log money demand curve with a semi-elasticity of only 1.8.

The main difference between our analysis and that in Ireland is the measure of money, highlighted in Figure 7. The figure shows the theoretical curve corresponding to the log-log specification with an elasticity of 0.3, which matches the behavior of M1 from 1915 till 1981. We also show Ireland's measure in red X's and NewM1 in green squares. While a break in the slope is clear using Ireland's measure, this is not the case when we use NewM1.

As described earlier, NewM1 adds to the standard M1 measure the Money Market Demand Accounts (MMDA), which were created in 1982 and, in a couple of years, became

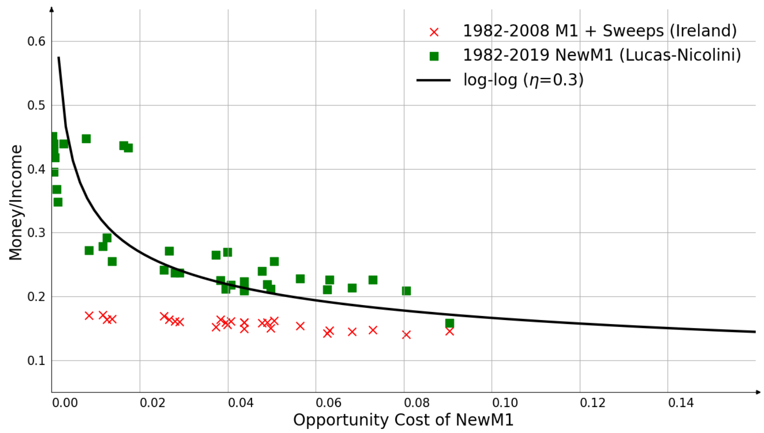

Figure 7: US money-income ratio and short-term interest rate, 1915-2019

around 10% of total output. The justification for doing so, as argued in Lucas and Nicolini, is that the MMDA provided transaction services that were very similar to the ones performed by checking accounts. In addition, they paid interest, which explains why they grew so much at the expense of standard checking accounts, which were banned from paying interest by Regulation Q.[6]

A new development occurred by the early 90s, when one bank adopted a software that automatically transferred funds from checking accounts to MMDAs of the same client in the same bank a few minutes before closing time, and would transfer them back to the checking account a few minutes after opening the following day. The profitability of these "sweeps", as they were called, is explained by the fact that the MMDA reserve requirement was only 1%, while it was 10% for the checking accounts. As reserve requirements are computed over end of day balances, the bank could make substantial savings on reserves. Thus, the sweeps were just a way to avoid the reserve requirement.

Concerned about the implications of this practice, other financial institutions raised the issued to the authorities, who did not react. Given the passive response, the practice extended to many other banks in a matter of a few years. In a nutshell, the sweeps were a way to bypass the reserve requirement without actually changing the regulation. For a detailed

[6]A reasonable interpretation is that the creation of these new accounts was a way to lessen the bite of Regulation Q without repealing it, which would have required congressional support.

account of this problem, see Cynamon et al.(2006).

The sweep programs completely blurred the distinction between demand deposits and a fraction of the stock of MMDAs as reported by the Federal Reserve, since that fraction of the MMDAs were, from the point of view of the holders, just demand deposits. Thus, the decision of Ireland (2009) is totally justified: from the point of view of the holder, the amounts in the sweep programs have the same transactional characteristics as the checking accounts. Ireland's choice is free from controversy: any attempt to measure purchasing power must include the sweeps.

The argument in Lucas and Nicolini (2015) is that Ireland's adjustment is not enough. They make the point that all the MMDAs, not only the ones artificially created by the sweep programs, are close enough substitutes to the demand deposits that the right practice is to include them in M1.

Ireland interpreted the regulatory changes of the 80s as permanent changes in the elasticity of real money demand. Thus, in order to evaluate the welfare cost of inflation in the 21st century, only data from 1980 onward should be used - which explains the title of his paper.[7] Lucas and Nicolini (2015), on the contrary, argue that the regulatory changes changed only the composition of transactional assets on the supply side, by creating a new asset that closely competed with demand deposits. But, according to Lucas and Nicolini (2015)'s theory, this change had minimal impact on the elasticity of money demand.[8]

To estimate the model, we adopt the proposal in Lucas and Nicolini (2015). That is certainly a debatable choice. Hoping that cross-country comparisons help shed light on the issue, we now briefly consider the experience of countries similar to the United States, for which we use the measure of M1 reported by their central banks for both before and after 1982. For space reasons, we restrict this informal discussion to three additional countries.

[7]The data for the total stock of MMDAs was not readily available when Ireland wrote his paper - though the data on sweeps had been collected by Cynamon et al. (2006). The measure estimated by Lucas and Nicolini (2015) was obtained from the call reports.

[8]The new deposits could - and did - pay interest that depends on the short-term interest rate on government debt. And the elasticity of real money demand does change when some deposits pay interest. However, the effect is quantitatively very small, so we ignore it in this discussion.

Our main interest in exploring these other countries is to see the extent to which they shed light on both the choice of functional form and the value of the parameters. Thus, in the cases that follow, we show the data together with the log-log theoretical curve with an elasticity of 0.5 - the one preferred by Lucas - and two theoretical curves for the semi-log, one with a semi-elasticity of 7 and one with a semi-elasticity of 1.8, as preferred by Lucas and Ireland, respectively.

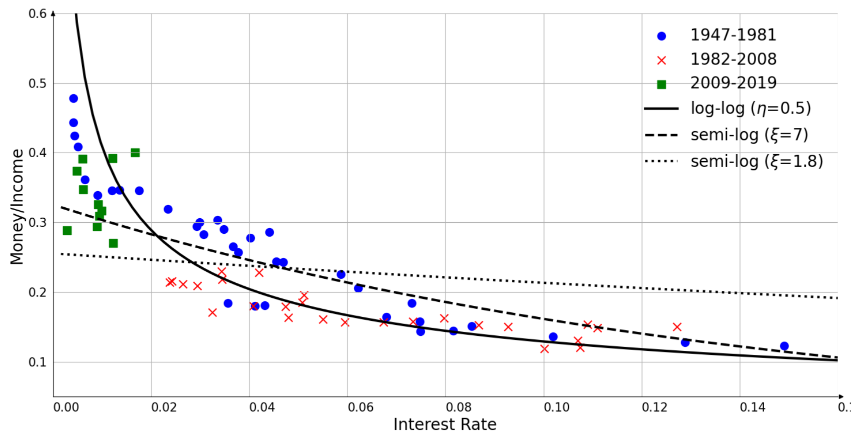

Figure 8: Canada real money demand, 1947-2019

In Figure 8, we show annual data for Canada from 1947 until 2019. Data until 1981 are shown as blue circles, data from 1982 to 2008 are shown as red X's, and data from 2009 to 2019 are represented with green squares. The first feature we would like to highlight is that the Canadian case differs from the US one in that there is no apparent break in the series in the 1980s, even though we use for the whole period M1 as defined by the Bank of Canada. The second feature is that the semi-log curve with the parameter equal to 1.8 does quite badly relative to the one with the parameter equal to 7. Finally, the log-log does much better than the semi-log at tracking the values with very low rates and high money to output ratios. Incidentally, note that as in the US, the data with low interest rates that followed the financial crisis of 2009 (the green squares) behave quite in the same fashion as the ones right after WWII (the blue circles with low interest rates). As in the US, the postwar years do not appear to be special in Canada.

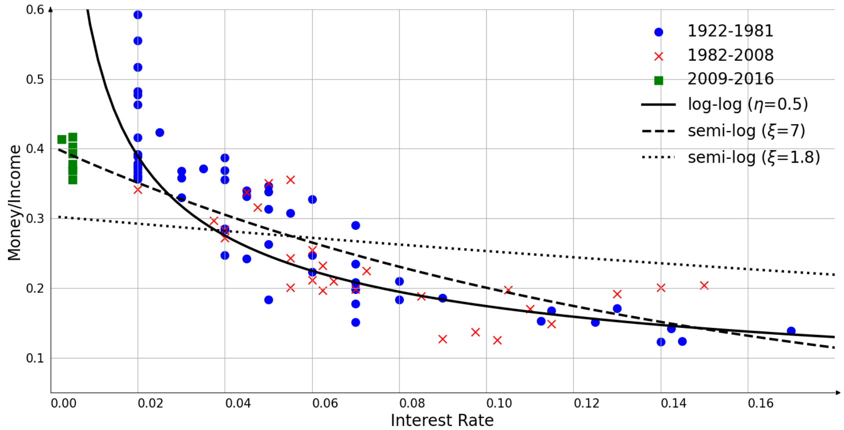

Figure 9: UK real money demand, 1922-2016

In Figure 9, we show similar evidence for the United Kingdom. The data from 1922 till 1981 are depicted with blue circles, the data from 1982 till 2008 with red X's, and the data from 2009 till 2016 with green squares. As in the case of Canada, no break is apparent in 1981. However, the UK case differs from the other ones in that the semi-log appears to do a better job than the log-log, as long as the semi-elasticity is chosen to be 7. The one with a semi-elasticity of 1.8 does very badly. The UK case is also different from that of both the US and Canada in that, as conjectured by Ireland, the war years (the blue circles with money to output ratios above 0.45) are very different from the years after 2008 (the green squares). It is precisely if one ignores those observations that the semi-log curve with elasticity of 7 performs particularly well.

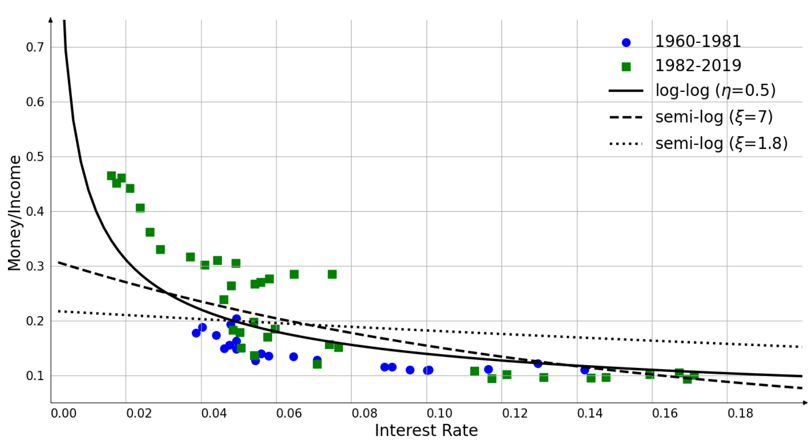

Figure 10: Australia real money demand, 1960-2019

Finally, in Figure 10, we plot the data for Australia starting in 1960. As was the case before, blue circles correspond to pre-1981 data, and green squares to the 1982-2019 period. Australia did not have interest rates very close to zero after 2009, so we do not differentiate that sub-period from the others. As in Canada and the UK, no break in behavior is apparent in 1981. The semi-log with an elasticity of 1.8 does not match the data well. Finally, in order to track the years of lower rates, the log-log specification performs better, although a slope higher than 0.5 would probably match the data better.

In summary, once the monetary aggregate for the US is adjusted to take into account new liquid deposits created in 1982, there is no apparent evidence of a change in behavior of money demand. In the other three countries, no evidence of a break is visible in the data using the standard M1 measure.

We find ambiguous evidence regarding the superiority of the log-log versus the semi-log specifications. The log-log clearly appears as the better specification for Australia and the semi-log for the UK. For Canada and the US, both specifications do a reasonable job at matching the data, except for observations with interest rates very close to zero, where the log-log performs better. The WWII years appear to be very special in the UK, but not in the US and Canada. When the log-log specification is preferred, the elasticity that best represents the data appears to be smaller than 0.5, the value preferred by Lucas. Finally, for the semi-log specification, the curve with a coefficient equal to 7 tracks the data very well, while the curve with a coefficient of 1.8 does not.

So far, we have focused on the two functional forms discussed by Lucas and Ireland. In what follows, we briefly discuss the performance of the functional form originally proposed and studied by Selden (1956) and Latané (1960), described in (4). As we show below, this functional form performs quite well in the econometric tests discussed below.

In Figure 11, we show three curves corresponding to the SL case: one with coefficient b equal to 25, one equal to 35, and one equal to 45. As was the case before, in each case, the constant is adjusted so that the curve crosses the grand mean of the data. It is apparent

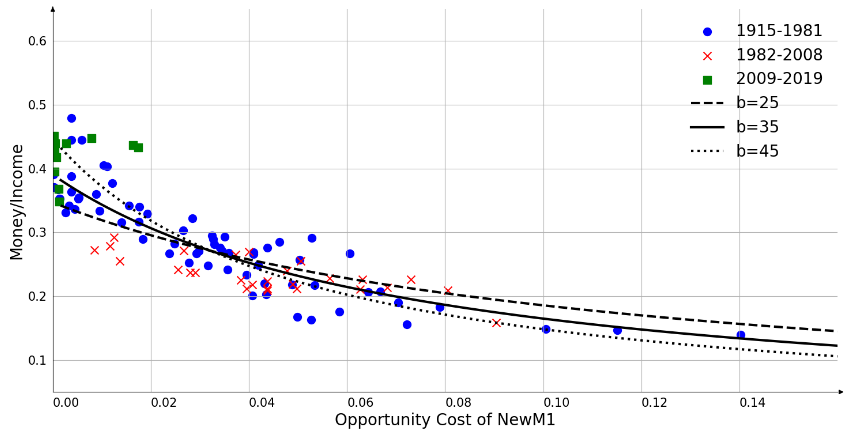

Figure 11: US Selden-Latané real money demand, 1915-2019

that the curve with coefficient b equal to 35 best matches the data.

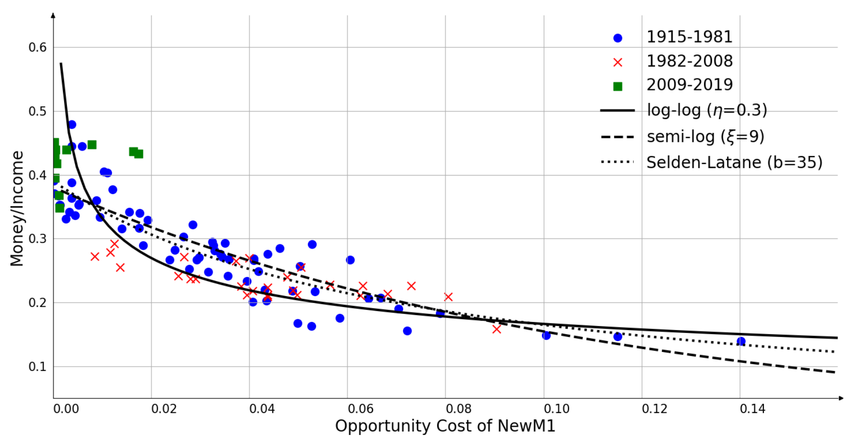

Figure 12: US real money demand, 1915-2019

In Figure 12, we plot the data, together with the three preferred specifications. Essentially, Figure 12 is the same as Figure 3, with the preferred SL specification added.

The SL specification does slightly better than the semi-log at very low values of interest rates, and it does better than both the semi-log and the log-log for values higher than, say, 2%. As we mentioned above, the log-log appears to underestimate the money-to-output ratio for intermediate values of the interest rate, and the semi-log appears to overestimate them. An attractive feature of the SL specification is that it is between the other two specifications in that range, so it provides a better approximation of the data. Below, we formally show that the SL has nice statistical properties.

4 Estimation and Testing

We study and compare the three functional forms, (2), (3), and (4) using formal econometrics. For the analysis of this section, we assume that the lower bound on interest rates is zero.

The methodology for estimating the three alternative specifications of the money demand curves closely follows the analysis by Benati et al. (2021). We first test for unit roots in the series. The evidence is overwhelming, as shown in Table A1 in Appendix D, which reports results from Elliot et al.'s (1996) unit root tests for either the levels or the logarithms of M1 velocity and the short rate.

In searching for a cointegration relationship between velocity and the short rate, we first take the unit root tests literally and use Johansen's tests. A plausible alternative interpretation of the results in Table A.1 is that the series are *local-to-unity*. So, we also search for cointegration based on Wright's (2000) test, which is valid for both exact unit roots and roots that are local-to-unity.[9] Results for both the Johansen and the Wright test are presented in Tables A.2 and A.3 in Appendix D.

4.1 Parameter estimates

Neither Johansen's nor Wright's tests directly provide point estimates for the parameters of the real money demand function.[10] We therefore estimate the money demand equations using Stock and Watson's dynamic OLS procedure, which delivers point estimates for the parameters.

Table 1 shows the point estimates, as well as 90% confidence intervals, for the coefficients ϕ for the Selden-Latané specification, γ for the semi-log, and η for the log-log.

The point estimate for the semi-elasticity parameter for the US, 9.1, is much closer to 7, the value adopted by Lucas (2000), than to 1.8, the one estimated by Ireland (2009). The

[9] All of the technical details about the implementation of the tests are identical to those of Benati (2020) and Benati et al. (2021), which the reader is referred to.

[10] For the Johansen test, the corresponding money demand equation is estimated in its VECM form, from which the money demand parameters can be indirectly obtained. Wright's test, on the other hand, does not produce point estimates, but rather confidence intervals at the $x\%$ level for the parameters.

Table 1 Point estimate and 90%-coverage bootstrapped[a] confidence interval for the coefficient on (the logarithm of) the short rate based on Stock and Watson's (1993) estimator

		Money demand specification:		
Country	*Period*	Selden Latané	Semi-log	Log-log
United States	1950Q1-2024Q2	-37.4 [-45.6 -25.8]	-9.1 [-11.6 -5.9]	-0.17 [-0.26 -0.11]
United Kingdom	1955Q1-2024Q2	-39.0 [-49.1 -23.7]	-8.5 [-11.6 -6.0]	-0.28 [-0.40 -0.18]
Canada	1947Q3-2006Q4	-40.4 [-50.6 -26.2]	-8.0 [-10.3 -5.2]	-0.38 [-0.47 -0.24]
	1967Q1-2024Q2	-39.2 [-51.9 -27.0]	-8.0 [-11.4 -5.1]	-0.31 [-0.41 -0.21]
Australia	1969Q3-2024Q2	-61.0 [-74.1 -37.7]	-11.4 [-14.1 -7.4]	-0.43 [-0.53 -0.29]
New Zealand	1988Q2-2024Q2	-36.7 [-53.3 -16.1]	-6.3 [-9.5 -3.0]	-0.27 [-0.36 -0.17]
South Korea	1964Q1-2024Q2	-44.0 [-47.7 -36.1]	-7.4 [-8.8 -4.9]	-0.48 [-0.56 -0.31]
Japan	1960Q1-2024Q2	-34.5 [-43.3 -20.4]	-18.0 [-23.3 -12.0]	-0.37 [-0.51 -0.19]
Hong Kong	1985Q1-2024Q2	-62.8 [-84.3 -41.7]	-15.8 [-22.1 -9.9]	-0.17 [-0.25 -0.11]
Switzerland	1972Q1-2024Q2	-26.7 [-32.5 -20.2]	-13.1 [-16.7 -9.4]	–[b]
Sweden	1998Q1-2024Q2	-26.8 [-38.4 -16.6]	-11.6 [-17.1 -6.9]	–[b]
Euro area	1999Q1-2024Q2	-32.1 [-41.6 -19.7]	-14.6 [-19.8 -9.5]	–[b]
Denmark	1991Q1-2024Q2	-15.3 [-20.4 -7.1]	-6.6 [-9.6 -2.8]	–[b]

[a] Based on 10,000 bootstrap replications.
[b] The last observations for the short rate are either zero or negative.

lower bound of the 90% confidence interval of our estimate is 5.9, substantially higher than the preferred value of Ireland.

An inspection of the results for the other countries shows that values close to the estimate for the US are quite common. In 6 cases, the point estimate is more than 10, and only for New Zealand and Denmark it is less than 7 - though barely. The value Ireland obtains for the US, 1.8, is substantially below the lowest bound of the 90 percent confidence interval for all the countries.

4.2 Which specification fits the data better?

It is not possible to nest the three specifications into a single encompassing one. However, we found a way to nest the semi-log with the log-log on one hand, and the semi-log with the Selden-Latané on the other.

We start from the comparison between the semi-log and the log-log. For each country, we regress $\ln(M_t/Y_t)$ on a constant, p lags of itself, and p lags of either the level of the short rate or its logarithm. A natural way of interpreting these regressions is the following. Under the assumption that cointegration is indeed there for all countries,[11] and based on either specification, both $Y_t^{SL} = [\ln(M_t/Y_t)\ R_t]'$ and $Y_t^{LL} = [\ln(M_t/Y_t)\ \ln(R_t)]'$ have a cointegrated VECM(p-1) representation, which maps into a restricted VAR(p) representation in levels (where the restrictions originate from the cointegration relationship). The equations we are estimating can therefore be thought of as the corresponding *unrestricted* form of the equations for $\ln(M_t/Y_t)$ in the VAR(p) representation in levels for either Y_t^{SL} or Y_t^{LL}. It is important to stress that the two specifications we are estimating are in fact nested. The easiest way of seeing this is to think of them as two polar cases—corresponding to either $\theta = 1$ or $\theta = 0$—in the following representation based on the Box-Cox transformation of R_t:

$$\ln\left(\frac{M_t}{Y_t}\right) = \alpha + \sum_{j=1}^{p} \beta_j \ln\left(\frac{M_{t-j}}{Y_{t-j}}\right) + \sum_{j=1}^{p} \delta_j \left(\frac{R_{t-j}^{\theta} - 1}{\theta}\right) + \varepsilon_t. \tag{8}$$

We estimate (8) via maximum likelihood, stochastically mapping the likelihood surface via Random-Walk Metropolis (RWM). The *only* difference between the "standard" RWM algorithm, which is routinely used for Bayesian estimation, and what we are doing here is that the jump to the new position in the Markov chain is accepted or rejected according to a rule that does not involve any Bayesian priors, as it uniquely involves the likelihood of the data.[12] So one way of thinking of this is as Bayesian estimation via RWM with completely uninformative priors, so that the log-posterior collapses to the log-likelihood of the data. All

[11] If this assumption did not hold, the entire model comparison exercise would obviously be meaningless.

[12] So, to be clear, the proposal draw for the parameter vector β, $\tilde{\beta}$, is accepted with probability min[1, $r(\beta_{s-1}, \tilde{\beta} \mid Y, X)$] and rejected otherwise, where β_{s-1} is the current position in the Markov chain and

$$r(\beta_{s-1}, \tilde{\beta} \mid Y, X) = \frac{L(\tilde{\beta} \mid Y, X)}{L(\beta_{s-1} \mid Y, X)},$$

which uniquely involves the likelihood. With Bayesian priors, it would be

$$r(\beta_{s-1}, \tilde{\beta} \mid Y, X) = \frac{L(\tilde{\beta} \mid Y, X) P(\tilde{\beta})}{L(\beta_{s-1} \mid Y, X) P(\beta_{s-1})},$$

where $P(\cdot)$ would encode the priors about β.

the other estimation details are identical to those of Benati (2008), to which the reader is referred.

Panel A of Table 2 reports, for either specification and for $p \in \{2, 4, 8\}$, the difference between the modes of the log-likelihood of the semi-log and the log-log specifications. The main message is that whereas the semi-log appears as the preferred functional form for the US, the UK, Canada, and Australia, the log-log produces a larger value of the likelihood for New Zealand, South Korea, Japan and Hong Kong, so neither specification is clearly superior to the other.[13]

Table 2 Model comparison exercise: Difference between the mode of the log-likelihoods

		A: Semi-log minus log-log			B: SL minus semi-log		
Country	*Period*	$p = 2$	$p = 4$	$p = 8$	$p = 2$	$p = 4$	$p = 8$
United States	1950Q1-2024Q2	20.13	21.24	35.00	1.73	5.04	3.50
United Kingdom	1955Q1-2024Q2	3.27	3.02	2.84	0.52	0.34	2.52
Canada	1947Q3-2006Q4	12.40	8.87	7.96	-0.27	1.78	2.06
	1967Q1-2024Q2	1.37	1.13	-1.11	1.26	1.22	8.74
Australia	1969Q3-2024Q2	12.88	13.35	12.53	2.66	1.87	3.35
New Zealand	1988Q2-2024Q2	-7.65	-2.12	-0.99	-0.40	2.37	1.48
Switzerland	1972Q1-2024Q2	-	-	-	0.65	1.40	3.97
Sweden	1998Q1-2024Q2	-	-	-	-0.13	-0.18	0.98
Euro area	1999Q1-2024Q2	-	-	-	0.17	0.36	3.00
Denmark	1991Q1-2024Q2	-	-	-	0.42	0.80	10.78
South Korea	1964Q1-2024Q2	-4.51	-8.81	-6.96	7.09	18.58	8.36
Japan	1960Q1-2024Q2	-7.89	-5.23	-4.39	0.14	0.73	0.22
Hong Kong	1985Q1-2024Q2	3.42	-0.36	-3.74	0.12	0.20	0.23

For Switzerland, Sweden, Euro area, and Denmark there is no comparison, because the last observations for the short rate are negative.

Turning to the comparison between the semi-log and the Selden-Latané, we adopt the same logic as before, but this time we "flip" the specifications for velocity on their head by regressing the interest rate on lags of itself and of either the level or the logarithm of velocity. Once again, these two regressions can be thought of as particular cases of the

[13] This crucially hinges on the fact that we are here focusing exclusively on low-inflation countries. As shown by Benati et al. (2021) and Benati (2021), for high-inflation countries, and especially hyperinflationary episodes, the data's preference for the log-log is overwhelming.

nested regression

$$R_t = \alpha + \sum_{j=1}^{p} \varphi_j R_{t-j} + \sum_{j=1}^{p} \xi_j \left[\frac{\left(\frac{Y_{t-j}}{M_{t-j}}\right)^{\theta} - 1}{\theta} \right] + \varepsilon_t, \tag{9}$$

with either $\theta = 1$ (corresponding to Selden-Latané) or $\theta = 0$ (corresponding to the semi-log).

At first sight, this approach might appear questionable: Since we are dealing with the demand for real M1 balances for a given level of the short-term nominal interest rate, why would it make sense to regress the short rate on M1 velocity? In fact, this approach is perfectly legitimate for the following reason. As shown by Benati (2020), M1 velocity is, to a first approximation (and up to a scale factor), the permanent component of the short-term rate,[14] so that if we focus, for example, on the Selden-Latané specification, $V_t = a^3 + \phi R_t^P$, where V_t is velocity, $a, \phi > 0$ are coefficients, and R_t^P is the unit-root component of the short rate (R_t), with $R_t = R_t^P + R_t^T$; R_t^T is the transitory component.

Regressing R_t on V_t therefore amounts to regressing the short rate on its (rescaled) stochastic trend—that is, the dominant driver of its long-horizon variation—and it is therefore conceptually akin to (e.g.) regressing GDP on consumption.[15]

The results are reported in Panel B of Table 2. The evidence is much sharper than that for the previous comparison: in particular, for p equal to 4 or 8, the Selden-Latané specification always preferred to the semi-log for all countries except for Sweden when $p = 4$.

Summing up, we note that whereas the Selden-Latané functional form appears to be quite clearly preferred to the semi-log, the log-log and the semi-log seem to be, from an empirical standpoint, on a roughly equal footing.

To be conservative, we choose the Selden-Latané to be our benchmark specification since it implies a finite satiation point. But given the evidence with low rates that we discuss below, we also discuss results using the log-log specification.

[14]This expresses in the language of time-series analysis Lucas's (1988) point that real M1 balances are very smooth compared with the short rate.

[15]See Cochrane (1994) on consumption being the permanent component of GDP.

4.3 The recent evidence with very low interest rates

Since the Global Financial Crisis, many of the countries in our sample kept nominal interest rates below the 2% threshold for most of the following decade and a half.

We exploit that feature of the data and make a comparison between the two functional forms precisely in the range that matters most. Specifically, we select countries that meet two requirements. First, there must be a long enough series of the interest rate that is always above a threshold. Second, that period must be followed by a long enough series of interest rates below that threshold.

Using data from countries that meet these requirements, we can estimate both the SL and the log-log functional forms, using data only above the threshold. Armed with those estimates, we compute confidence intervals for interest rates below that threshold. We then use the second part of the sample to compare the ability of the two econometric models to replicate the out-of-sample behavior of real money demand when interest rates are below 2%.

The results using the Selden-Latané functional form are reported in Figure 13. The wide red lines show the point estimates for real money demand. The thinner lines report the 66% and 90% confidence intervals. These estimates are based on data using interest rate observations greater than 2% for the eight countries that satisfy our criteria. Typically, this is the sample before 2008-2009. The figure also plots the observations for the subsequent period that exhibit interest rates below 2%. Except for a small number of observations in the UK, all the observations for the eight countries lie above the estimated curve - one would have expected roughly half the observations to be below and half above. But, more importantly, with the exception of Australia and the UK, many of the observations lie above the 90% confidence interval. This feature is particularly dramatic in Japan and Hong Kong. In general, the Selden-Latané specification tends to underestimate the elasticity of money demand for interest rates below 2%, except in the case of the UK.

A similar exercise can be performed for the log-log specification. The results are depicted

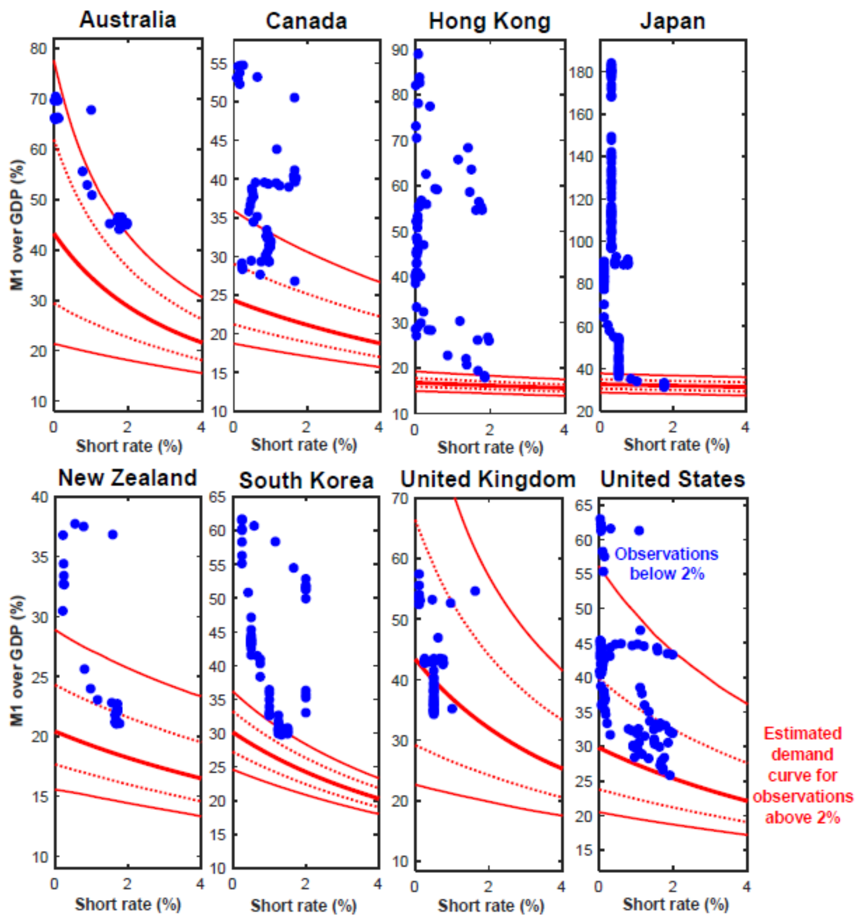

Figure 13: Observations below 2%, and estimated money demand curves based on observations above 2%. Based on the Selden-Latané functional form

in Figure 14. It is still the case that observations are much higher than the estimated value for Japan and Hong Kong. For the other countries, the most important difference is that, with the exception of the UK, the observations are now more evenly distributed around the point estimates, although in many cases, they are outside the confidence intervals. For the UK, only one observation lies above the estimated curve.[16] Interestingly, in the US all observations lie within the 90 percent confidence interval. These results are consistent with the eyeball inspection of simple plots in Section 3.

This exercise justifies presenting the results for the log-log case along our benchmark Selden-Latané computations.

[16]Incidentally,this is consistent with the visual evidence of the UK we discussed in Section 4, where the log-log seemed to perform worst.

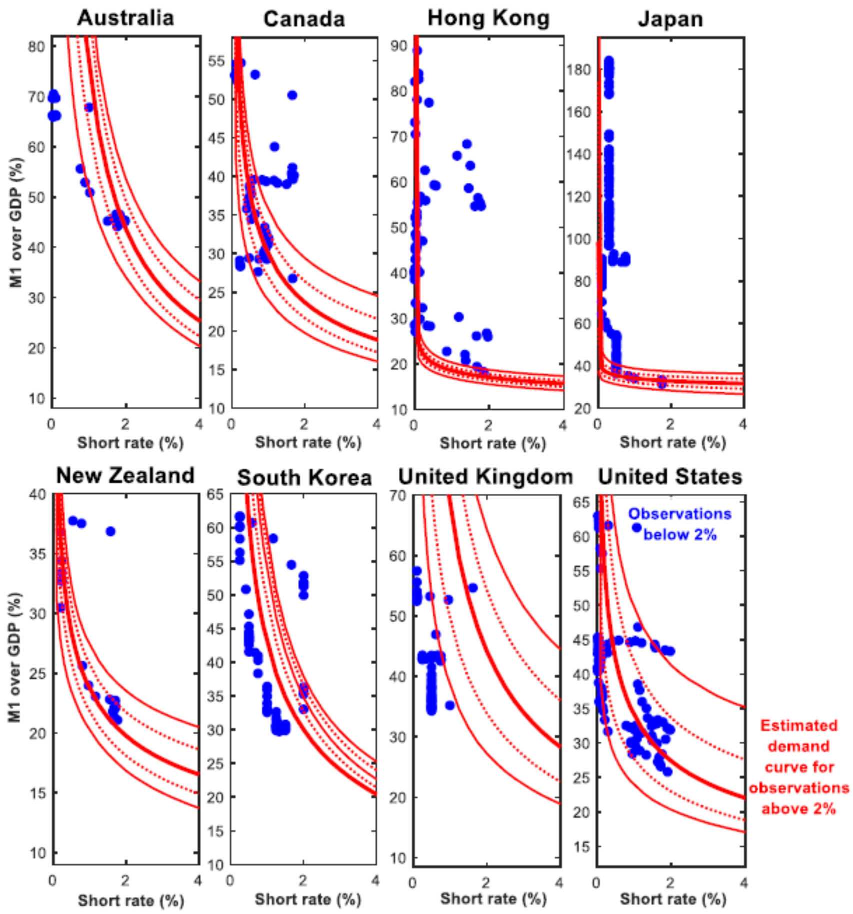

Figure 14: Observations below 2%, and estimated money demand curves based on observations above 2%. Based on the log-log functional form

5 The Estimated Welfare Cost Functions

The methodology we use in order to compute the welfare costs of inflation follows Luetkepohl (1993, pp. 370-371). We first estimate via OLS the cointegrating regression corresponding to any of the three specifications—that is, to (2), (3), or (4). This gives us the point estimates of the parameters - see Table 1 - we need in order to compute the point estimates of the welfare cost functions. We then estimate the relevant VECM via OLS by imposing in estimation the previously estimated cointegration vector, and we characterize uncertainty about the point estimates of the welfare cost function by bootstrapping the VECM as in Cavaliere et al. (2012).

As discussed in Section 4.1, this procedure is valid if the series contain *exact* unit roots.

Under the alternative possible interpretation of the results from unit root tests—that is, that the series are *local-to-unity*—we proceed as in Benati et al. (2021, Section 4.2.1). Specifically, we compute, from the just-mentioned VECM, the corresponding VAR in levels, which by construction features one, and only one, exact unit root. We turn it into its corresponding near unit root VAR by shrinking the unit root to λ=1-0.5($1/T$), where T is the sample length.[17] The bootstrapping procedure we implement for the second possible case, in which the processes feature near unit roots, is based on bootstrapping such a near unit root VAR. The two bootstrapping procedures deliver near-identical results, and in what follows, we will therefore report and discuss only those based on bootstrapping the VECM.

In this section, we focus on countries for which interest rates were always non-negative. The top panel of Figure 15 reports the welfare cost of inflation for the United States, Canada, New Zealand, and South Korea, using the Selden-Latané functional form.

The point estimates of the upper and lower bounds are depicted as continuous black lines: as we previously anticipated, in all cases the two lines are virtually indistinguishable, thus implying that the two bounds provide a very precise characterization of the welfare costs (the same holds for nearly all countries and all functional forms). The dotted and solid red lines respectively depict 67% and 90% confidence intervals obtained from the bootstrapped distributions.

For a steady-state interest rate of 5%, the value used by both Lucas and Ireland, we find the welfare cost to be between 0.18% (New Zealand) and 0.35% (US and South Korea) of permanent consumption.

The estimate for the US is about a third of the one obtained by Lucas (2000) when using the log-log specification and almost 10 times above Ireland's (2009) estimate of 0.037% when using the semi-log. As we show in Appendix E, the estimate of the welfare cost for the US using the semi-log is also close to 0.35.[18]

[17] For details see Benati et al.'s (2021) footnote 24.

[18] Results using the semi-log for all countries are reported in Appendix E. That specification is dominated by the Selden-Latané formula, and the computed welfare costs are similar between them.

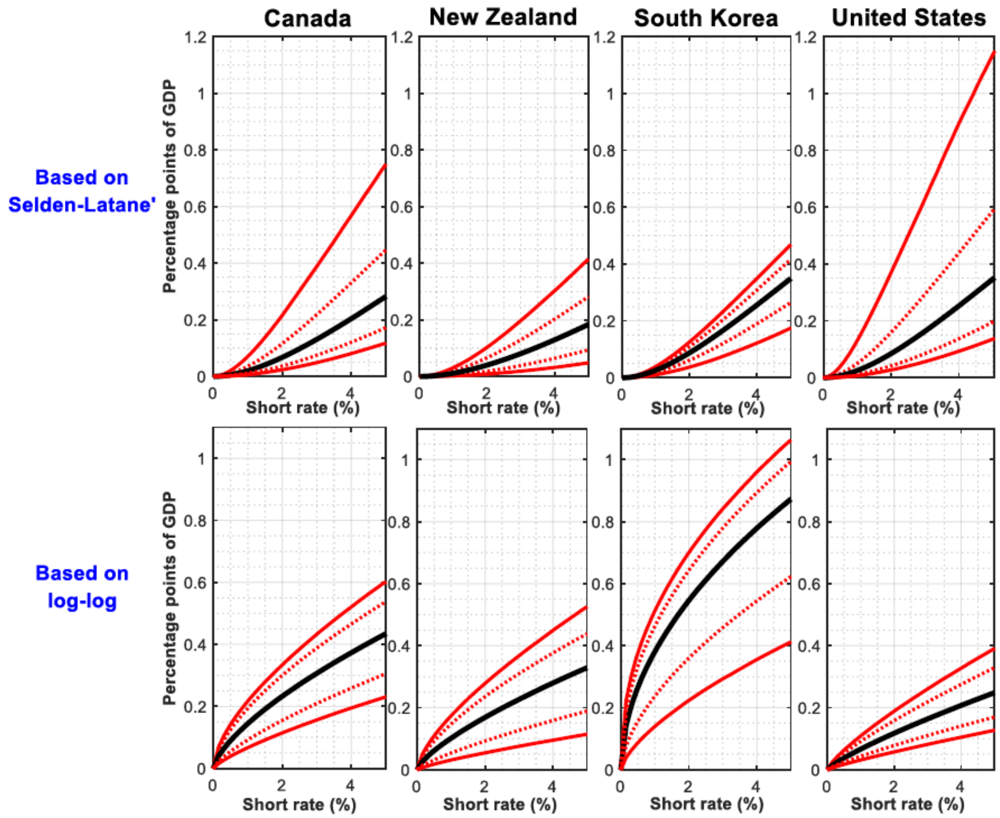

Figure 15: Estimated welfare cost functions based on the Selden-Latané and log-log specification: point estimates of the lower and upper bounds, 5th and 16th percentiles of the lower bounds, and 84th and 95th percentiles of the upper bounds of the bootstrapped distributions

The reason for the discrepancy is that our estimate of the semi-elasticity is substantially larger than the one used by Ireland (2009).

In the top panel of Figure 16, we show equivalent results for the UK, Japan, Hong Kong, and Australia. In this case, equivalent estimates range from 0.45% (UK and Hong Kong) to 0.80% of consumption (Japan), substantially larger than the ones we obtained for the first group.

The bottom panel of Figure 15 shows the results for the log-log case. The figure highlights the theoretical point made by Lucas (2000): as the log-log specification implies that the welfare cost is a convex function of the interest rate, it implies substantially higher costs at very low interest rates. This is clearly the case for Canada, New Zealand and South Korea, where the range of welfare cost of a 5% interest rate goes from 0.3% to 0.4%, 0.2% to 0.3%, and 0.35% to 0.85% percent, respectively.

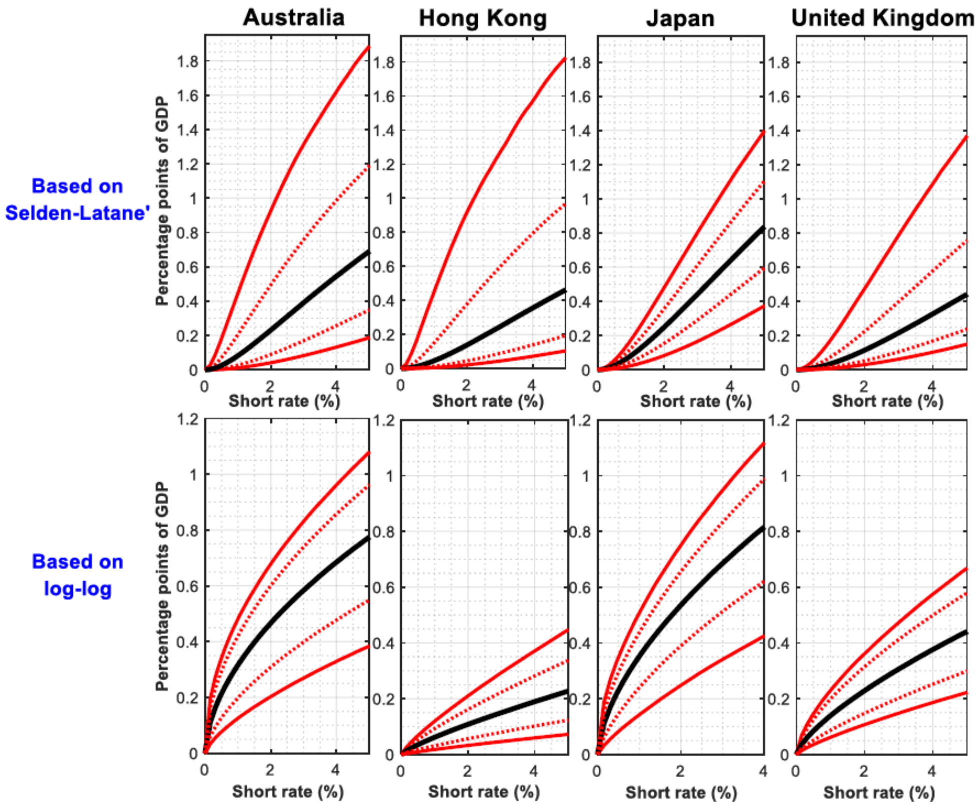

Figure 16: Estimated welfare cost functions based on the Selden-Latané and log-log specification: Point estimates of the lower and upper bounds, 5th and 16th percentiles of the lower bounds, and 84th and 95th percentiles of the upper bounds of the bootstrapped distributions

However, the point at which the curves cross each other - and therefore at which the log-log delivers lower welfare costs for higher interest rates - depends on the estimated slope and level parameters, and those differ across countries significantly. For the US, the estimated welfare cost for a 5% rate for the log-log is lower than for the Selden-Latané. What explains this fact is that our point estimate for the elasticity (0.17) in the United States is much smaller than the one used by Lucas (0.5). In fact, the US is the country for which the estimated elasticity is the lowest. In fact, the estimated elasticity is also substantially lower than 0.3, the value we chose when discussing the evidence in Figure 2.

The bottom panel of Figure 16 shows corresponding results for Australia, Japan, Hong Kong, and the UK. The log-log implies higher costs for Australia, Japan, and the UK, while the case of Hong Kong is like that of the US in Figure 15. As it happens, Hong Kong and the US are the two countries with the lowest point estimate of the elasticity in the log-log case (which in both cases is equal to 0.17).

6 Allowing for Negative Short-Term Interest Rates

So far, we have followed the literature in assuming that as the nominal return on bonds goes to zero, so does the nominal return on money. Under this condition, then, the lower bound on r_t^b is zero. The recent experience of prolonged negative short-term interest rates in several countries challenges this notion. As the opportunity cost of money r_t must be non-negative, the interest rate on bonds can be negative only if the own return on money is negative, at least when r_t^b becomes small.

The relevant opportunity cost for the representative agent is the difference $r_t^b - r_t^m$. Our model does not explicitly model banks, but its equilibrium can be decentralized with a competitive banking sector in which negative rates are passed to depositors.[19] An alternative model, in which banks have monopoly power, may have banks that do not pass the negative rate to their households, and collect income through higher fees.

Did the negative policy rates translate into negative rates for depositors in these experiences? There is evidence that small deposits did not pay negative rates, even in Switzerland, where interest rates were the lowest. But there is also evidence that for large deposits - affecting mostly firms - the nominal return was negative.[20] There is also evidence of heterogeneity among customers and banks. For instance, Michaelis (2022) shows that by early 2018, while 40% percent of German banks were paying negative rates on average on overnight deposits for non-financial corporations, only 10% were doing so for households. However, by 2022, close to the end of the negative policy rate period, approximately half of the banks were paying negative rates for both corporations and households. Michaelis also shows that fee income substantially increased during this period.

[19] See, for example, Prescott (1987).

[20] See https://www.reuters.com/business/finance/credit-suisse-group-ending-negative-interest-rates-private-clients-2022-06-29.

At one extreme, one could assume that the negative policy rates were just a tax on banks and irrelevant to depositors. If this were the case, the results of the previous section would be the valid ones. The purpose of this section is to illustrate the robustness of those results to alternative assumptions.

To account for negative policy rates, we proceed as follows. As we identify our measure of money with M1 in the data, it is natural to think of the return on money as an average of the return of the two components of M1, cash and demand deposits. For cash, a negative return can be rationalized by the risk of being lost or stolen, as Alvarez and Lippi (2009) measure using survey data.[21] For deposits, we use a linear relation between their nominal return and the interest rate on bonds. Kurlat (2019) provides very strong empirical support for such a relationship. These assumptions, taken together, are consistent with the return on money satisfying

$$r_t^m = -a + b r_t^b, \tag{10}$$

for $a \geq 0$ and $b < 1$.[22] This linear relationship implies that r_t^m will be negative for small enough values of r_t^b, and it implies that $r_t^b \geq -a/(1-b)$.

Thus, for $a > 0$, the lower bound on the short-term rate is negative. The standard assumption in the literature is obtained by imposing that $a = b = 0$. Kurlat (2019) estimates b to be close to 0.15, very precisely using micro-data from the US. We adopt that value. We then let $a = 1$, which corresponds to a lower bound on the short-term interest rate of roughly -1.2% percent.

This can account for the observations on short-term rates in Denmark, the euro area, and Sweden. It cannot account for Switzerland, for which the lowest value for the short-term interest rate was around -1.8%, so we do not discuss the log-log case for that country.[23]

[21] Alvarez and Lippi (2009) calibrate this return at -0.02, using survey data from Italy.

[22] Further details are provided in the Online Appendix F.

[23] Switzerland is special not only because its low short rate, but also in that it exports banking services. It is likely that its measure of M1 and its ability to implement negative rates in deposits may very well be country-specific.

In addition to being based on empirical evidence, the linear relationship has the advantage that the relevant opportunity cost r_t becomes

$$r_t = r_t^b - r_t^m = a + (1 - b)r_t^b,$$

which is a linear transformation of the observable short-term interest rate r_t^b. As the last two functional forms we adopted for the money demand, equations (3) and (4) , are either a linear function of r_t or the inverse of a linear function of r_t, one needs only to estimate those two specifications under the benchmark case of $a = b = 0$, then adjust the estimates by the corresponding linear transformation. Then, we use those estimates to compute the welfare cost.

However, for the log-log specification, this is not the case, and both the cointegration tests and the estimates will depend on the specific assumption regarding the lower bound. As it turns out, both are quite sensitive to the assumed lower bound, particularly so for the case of the United States.

We discuss the effects of the assumed lower bound on the cointegration tests and the comparison between the log-log and the semi-log to Appendix G. In a nutshell, all cointegration tests uniformly improve for the log-log specification when the lower bound is reduced. In testing between the log-log and the semi-log, the performance of the log-log also improves uniformly, but only for Canada does the result reverse so that the log-log outperforms the semi-log. Finally, for the three countries with negative rates, the semi-log outperforms the log-log.

6.1 Estimation results and welfare computations

Table 3 presents the estimation results for the log-log case, under the two assumptions regarding the zero bound. We first show the results for the three cases in which interest rates visited negative territory and then the rest of the countries. For all these countries,

Table 3 Point estimate and 90%-coverage bootstrapped[a] confidence interval for the coefficient on the logarithm of the short rate based on Stock and Watson's (1993) estimator

Country	*Period*	a=0, b=0	a=-1, b=0.15
Sweden	1998Q1-2019Q4	–[b]	0.250 [0.212 0.291]
Euro area	1999Q1-2019Q4	–[b]	0.398 [0.341 0.465]
Denmark	1991Q1-2019Q4	–[b]	0.298 [0.183 0.396]
United States	1959Q1-2019Q4	0.165 [0.087 0.235]	0.406 [0.255 0.531]
United Kingdom	1955Q1-2019Q4	0.284 [0.155 0.404]	0.468 [0.259 0.630]
Canada	1947Q3-2006Q4	0.373 [0.236 0.468]	0.544 [0.357 0.676]
	1967Q1-2019Q4	0.305 [0.200 0.382]	0.467 [0.295 0.561]
Australia	1969Q3-2019Q4	0.749 [0.518 0.892]	0.916 [0.640 1.083]
South Korea	1964Q1-2019Q4	0.477 [0.401 0.539]	0.655 [0.565 0.722]
Japan	1960Q1-2019Q4	0.328 [0.172 0.440]	0.646 [0.281 0.917]
Hong Kong	1985Q1-2019Q4	0.171 [0.096 0.241]	0.587 [0.363 0.824]

[a] Based on 10,000 bootstrap replications. [b] The last observations for the interest rate are either zero or negative.

the point estimates for the interest rate elasticity increase substantially as the lower bound is reduced.

For reasons of space, we only report the welfare computations for the first three cases and for the US. Figure 17 presents the welfare costs for Denmark, the euro area, Sweden and Switzerland for the Selden-Latané functional form. For the case of a zero lower bound, the welfare costs are somewhat higher than for the US: between 0.4 and 0.6 percentage points of consumption. The range increases to 0.5% to 0.8% when the lower bound is assumed to be -1.2%.

In Figure 20, we report estimates for the log-log specification when we assume a lower bound equal to -1.2%. In this case, we obtain substantially higher numbers: close to 0.6% for Denmark and Sweden, around 0.8% for the euro area, and close to 1% for the United States.

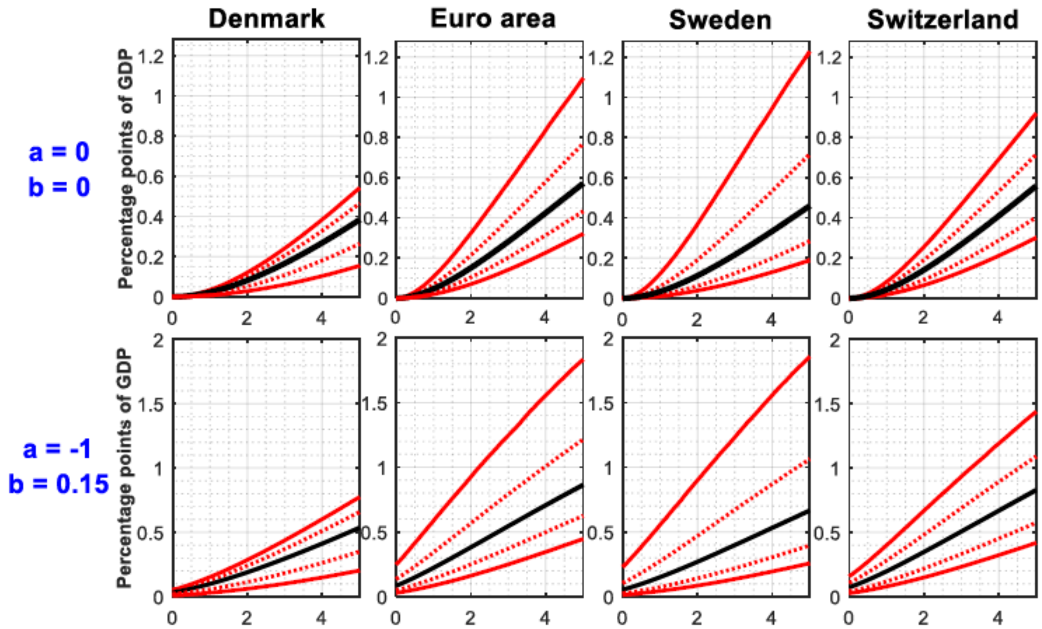

Figure 17: Estimated welfare cost functions based on the Selden-Latané specification: Point estimates of the lower and upper bounds, 5th and 16th percentiles of the lower bounds, and 84th and 95th percentiles of the upper bounds of the bootstrapped distributions

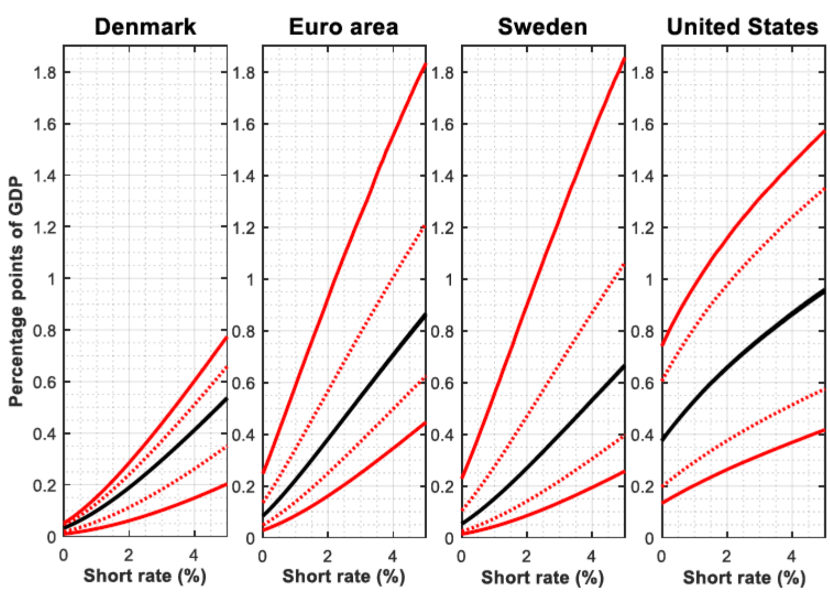

Figure 18: Estimated welfare cost functions based on the log-log specification, $a = -1$ and $b = 0.15$: Point estimates of the lower and upper bounds, 5th and 16th percentiles of the lower bounds, and 84th and 95th percentiles of the upper bounds

7 Conclusion

How large is the cost of deviation from the Friedman rule if the nominal interest rate is set at 5% in the steady state? A well established tradition, started by Bailey (1956) and Friedman (1969), estimates those costs by computing the area under the real money demand curve. Lucas (2000) follows this tradition and, arguing that a log-log specification is a good fit for the US data during the 20th century, computes that cost to be 1.2% of lifetime consumption.

However, Ireland (2009) argued that a specification with a satiation point at the lower bound provides a much better fit. A distinct feature of the finite satiation point when the opportunity cost of money is zero implies that the integral under the real money demand is not as large as with the log-log. He also argues that the elasticity is much lower than the one used by Lucas. When both things are considered, Ireland estimates the welfare cost to be a mere 0.036% of consumption.

We use new data for the US and also study the behavior of real money demand for several other developed countries. Our analysis quite strongly supports Lucas's estimates. When using the full support, a functional form with a finite satiation point performs better with the data. However, the analysis for very low values of the opportunity cost works better with the log-log specification. Finally, our sample contains countries that experienced negative policy rates, suggesting the possibility of a negative lower bound on the opportunity cost of money.

These considerations provide the two most extreme scenarios. Our lowest set of estimates is obtained with a finite satiation point, corresponding to the Selden-Latané specification, and assuming the lower bound is zero. This case delivers a welfare cost of a 5% nominal interest rate of about 0.35% percent of permanent consumption for the US. For the log-log case and a negative lower bound compatible with the experience of the countries in our sample, the welfare cost is about 1% of permanent consumption.

References

Afrouzi, H., S. Bhattarai, and E. Wu (2024):"The Welfare Cost of Inflation in Production Networks," Unpublished manuscript.

Alvarez, F., and F. Lippi (2009): "Financial Innovation and the Transactions Demand for Cash," *Econometrica*, 77(2), 363-402.

Alvarez, F, F. Lippi, and R. Robatto (2019): "Cost of Inflation in Inventory Theoretical Models,"*Review of Economic Dynamics*, 32, 206-226.

Bailey, M. J. (1956): "The Welfare Cost of Inflationary Finance," *Journal of Political Economy*, 64(2), 93-110.

Baumol, W. J. (1952): "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics*, 66(4), 545-556.

Benati, L. (2008): "Investigating Inflation Persistence across Monetary Regimes," *Quarterly Journal of Economics*, 123(3), 1005-1060.

Benati, L. (2020): "Money Velocity and the Natural Rate of Interest," *Journal of Monetary Economics*, 116, 117-134

Benati, L. (2021): "The Monetary Dynamics of Hyperinflations, Reconsidered" Unpublished manuscript.

Benati, L., R. E. Lucas Jr., J. P. Nicolini, and W. Weber (2021): "International Evidence on Long-Run Money Demand," *Journal of Monetary Economics*, 117, 43-63.

Blanchard, O.J., G. Dell'Ariccia, and P. Mauro (2010): "Rethinking Macroeconomic Policy," Staff Position Note 10/03, International Monetary Fund.

Cavaliere, G., A. Rahbek, and A. M. R. Taylor (2012): "Bootstrap Determination of the Co-integration Rank in Vector Autoregressive Models," *Econometrica*, 80(4), 1721-1740.

Cochrane, J.H. (1994): "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics*, 109(1), 241-265.

Coibion, O., Y. Gorodnichenko, and J. Wieland (2012): "The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise their Inflation Targets in Light of the

Zero Lower Bound?," *Review of Economic Studies*, 79, 1371-1406.

Cynamon, B.Z., D.H. Dutkowsky, and B.E. Jones (2006): "Redefining the Monetary Aggregates: A Clean Sweep," *Eastern Economic Journal*, 32(4), 661-672.

Elliot, G. (1998): "On the Robustness of Cointegration Methods When Regressors Almost Have Unit Roots," *Econometrica*, 66(1), 149-158,

Elliot, G., T. J. Rothenberg, and J. H. Stock (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64(4), 813-836.

Friedman, M. (1969): "The Optimum Quantity of Money,"in *The Optimum Quantity of Money and Other Essays*, Aldine Publishing Company, 1-50.

Ireland, P. (2009): "On the Welfare Cost of Inflation and the Recent Behavior of Money Demand," *American Economic Review*, 99(3), 1040-1052.

Khan, A., R. King , and A. Wolman (2003): "Optimal Monetary Policy," *The Review of Economic Studies* , 70(4), 825–860.

Kurlat, P. (2019): "Deposit Spreads and the Welfare Cost of Inflation," *Journal of Monetary Economics*, 106, 78-93.

Latané, H. A. (1960): "Income Velocity and Interest Rates: A Pragmatic Approach," *Review of Economics and Statistics*, 42(4), 445-449.

Lucas, R.E., Jr. (1988): "Money Demand in the United States: A Quantitative Review," *Carnegie-Rochester Conference Series on Public Policy*, 29, 137-167.

Lucas, R.E., Jr. (2000): "Inflation and Welfare," *Econometrica*, 68(2), 247-274.

Lucas, R.E., Jr., and J. P. Nicolini (2015): "On the Stability of Money Demand," *Journal of Monetary Economics*, 73, 48-65.

Luetkepohl, H. (1993): *Introduction to Multiple Time Series Analysis*, (2nd ed.), Springer-Verlag.

Michaelis, H. (2022) : "Going below Zero - How Do Banks React?," *Discussion Paper no.33/2022, Deutsche Bundesbank.*

Nakamura, E., J. Steinsson, P. Sun, and D. Villar (2018). "The Elusive Costs of Inflation:

Price Dispersion during the U.S. Great Inflation," *Quarterly Journal of Economics* 133(4), 1933–1980.

Prescott, E. C. (1987): "A Multiple Means-of-Payment Model," in W. A. Barnett and K. J. Singleton, eds., *New Approaches to Monetary Economics*, Cambridge University Press, 42-51.

Selden, R. T. (1956): "Monetary Velocity in the United States," in M. Friedman, ed., *Studies in the Quantity Theory of Money*, University of Chicago Press, pp. 179-257.

Stock, J. H., and M. W. Watson (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica*, 61(4), 783-820.

Tobin, J. (1956): "The Interest-Elasticity of Transactions Demand for Cash," *Review of Economics and Statistics*, 38(3), 241-247.

Wright, J. H. (2000): "Confidence Sets for Cointegrating Coefficients Based on Stationarity Tests," *Journal of Business & Economic Statistics*, 18(2), 211-222.