Journal of Economic Dynamics & Control 162 (2024) 104842

Contents lists available at ScienceDirect

Journal of Economic Dynamics and Control

journal homepage: www.elsevier.com/locate/jedc

The simple macroeconometrics of the quantity theory and the welfare cost of inflation

Kenneth G. Stewart

Department of Economics, University of Victoria, Victoria, British Columbia, V8W 2Y2, Canada

ARTICLE INFO

JEL classification:
E31
E4

Keywords:
Quantity theory of money
Money demand
Bounds tests
Indicator saturation
Welfare cost of inflation

ABSTRACT

The quantity theory of money hypothesizes that the price level is determined through the equilibration of money supply and demand. Predicated on this causal structure, a single-equation error correction model decomposes from a larger vector autoregressive system so as to make available bounds tests for a levels relationship that are robust to the univariate integration properties of the variables. This model is estimated using three monetary aggregates and two money demand specifications, for U.S. and U.K. annual data over the past century and quarterly post-WWII data. The classic homogeneity propositions of the quantity theory are testable, and are found to be most compatible with U.S. annual M2 using log-log money demand with structural change permitted. Nevertheless, the resulting welfare costs are similar to those yielded by the U.K. annual data, being less than one percent of GDP at interest rates experienced during the past century.

1. Introduction

Setting aside hyperinflations, the truth of Milton Friedman's famous phrase that "inflation is always and everywhere a monetary phenomenon" has seldom been so clearly revealed as during the Covid-19 pandemic and its aftermath. As the pandemic developed in 2020, major central banks wisely expanded their money supplies to prevent the accompanying economic downturn from developing into a full-blown crisis, and were successful in doing so. To moderate the inflation that followed, a counter-balancing constraining of monetary expansion has been required, the success of which is also becoming apparent.

These events can only serve to remind economists that, at some level, we all believe in the quantity theory of money: the level of prices is determined by the amount of money in circulation relative to people's desire to hold it for facilitating transactions. Going hand-in-hand with this is the proposition that the demand for nominal money balances should be related one-for-one to nominal transactions: a doubling of nominal transactions should double the demand for money. Equivalently, money demand is a demand for *real* money balances as a homogeneous-of-degree-one function of real transactions.

Yet, as is perhaps not unusual with things many of us believe about economics, when we turn to formal empirical analysis it can sometimes be surprisingly difficult to find the quantity theory in the data. One reason for this is that, although these "homogeneity propositions" are compelling, there is no reason why they should necessarily be satisfied for *all* money stock measures, however defined. Instead we can legitimately ask which monetary aggregates are actually consistent with these properties.

In this paper I argue that the quantity theory of money (QTM) can be captured by the simplest of dynamic econometric frameworks, a single-equation autoregressive distributed lag (ADL) model and its error correction mechanism (ECM) reparameterization,

E-mail address: kstewart@uvic.ca.

https://doi.org/10.1016/j.jedc.2024.104842
Received 12 November 2021; Received in revised form 15 February 2024; Accepted 26 February 2024

Available online 4 March 2024

and that there are many advantages to doing so. These advantages include the ability to test the homogeneity propositions just mentioned within a cointegration methodology, due to Pesaran et al. (2001), that is robust to the univariate integration properties of the series. The equilibrium error of the ECM is the cointegrating relationship between money, prices, income, and (possibly) interest rates that is predicted by the QTM. Because the univariate integration properties of interest rates are in dispute, the PSS bounds tests are a natural tool for studying the existence of an equilibrium levels relationship.

That cointegration is a natural framework for studying the quantity theory has long been recognized, although it has often been cast in terms of estimating money demand. Seminal early contributions include Hendry and Ericsson (1991), Baba et al. (1992), and Stock and Watson (1993); while Hoffman et al. (1995), Carlson et al. (2000), and Hendrickson (2014) are examples of prominent subsequent work. However, to my knowledge the present paper is the first to recognize that the PSS methodology is particularly well-suited to studying the QTM, for reasons I will argue.

Because the QTM and the welfare cost of inflation are inherently long run phenomena, I conjecture that they are most likely to be revealed by long-span data. This work is therefore in the spirit of other recent interest in bringing long-span data to bear on fundamental questions in macroeconomics. Notable examples include Jorda et al. (2017), Benati et al. (2021), and Rogoff et al. (2022). It is well established that the properties of unit root and cointegration tests improve with the span of the data, as opposed to merely having the larger number of observations provided by higher frequency data over a shorter span.

Macroeconomic data for substantial parts of the past century are available for many countries, but consistently-constructed uninterrupted series for the aggregates of the quantity theory that go as far back as World War I and continue to be updated are available for only two major countries: the United States and United Kingdom. For these two countries, results from the annual data are compared with analogous analyses of post-WWII quarterly series of the kind that have most often been used in empirical money demand studies.

Many of the findings are intuitive. In terms of cointegration, the homogeneity propositions, and plausible welfare costs, the QTM is most strongly revealed by the long-span annual data rather than the shorter-span quarterly data, even though the latter have more observations. It is also usually better revealed by a broader monetary aggregate than a narrower one. The value of long-span data means that empirical tests of the homogeneity propositions, and other aspects of the analysis, can sometimes hinge on permitting structural change. For this purpose I use the recently-developed indicator saturation methodology of Castle et al. (2015), where structural breaks are located by the data rather than being conjectured by the researcher. The important 20th century structural breaks are found to be quite different between the U.S. and U.K. Even so, the welfare costs of inflation turn out to be fairly comparable across the two countries, by most measures amounting to less than one percent of GDP at nominal interest rates experienced over the past century.

Studying the QTM within an empirical framework that is robust to the integration properties of the series, and in which the homogeneity propositions are testable, contrasts with other recent approaches to related issues. Benati et al. (2021) (henceforth BLNW) use an extensive international dataset to study long run M1 demand for 38 countries. They interpret the relevant cointegrating relationship to be between velocity and the interest rate. This has the advantage of requiring only nominal GDP for each country (as opposed to both nominal and real GDP, and the implied GDP deflator as a measure of the price level), because nominal GDP is all that is needed to obtain velocity. This increases the number of countries for which longer-span annual time series are available. But it has two disadvantages: first, that the homogeneity propositions are adopted as maintained hypotheses; and, second, that velocity and the interest rate must be I(1) in their univariate behavior. A similar view is taken by Benati (2020). In contrast, I argue that nothing in the QTM requires or implies that velocity or interest rates must be integrated individually, or that they be cointegrated. Because this may elicit some controversy, it is useful to set the stage and establish notation by reminding ourselves of the elementary essentials of the QTM, in order to understand how those essentials motivate my simple methodological approach.

2. Background

Were our interest in the short-term dynamics of the variables of the quantity theory, quite complicated specifications for money supply and demand could be entertained. Money supply M_t^S would be specified as interest-elastic, because commercial banks reduce excess reserves in response to higher interest rates. And, similarly, much could be said about the demand for money M_t^D, on which there is of course an enormous and longstanding literature. Indeed, in contemporary models of the short term dynamics of interest rates and prices, as surveyed by McCallum and Nelson (2011, sec. 8.1), the dominant factors are such that money may be viewed as not playing an important role at all.

2.1. The causal structure of the quantity theory

At its most essential, however, the QTM is a theory of price level determination in the long run, because it attributes inflation to money supply growth, and inflation is by its nature a long run phenomenon. By definition inflation is an ongoing and sustained growth in prices over an extended period of time, not merely a one-time or temporary jump in prices. So too is the welfare cost of inflation an inherently long run phenomenon, because—as the literature conventionally defines it—that welfare cost arises from people economizing on their money holdings in response to expectations of inflation. And for expectations of inflation to develop, an inflation must be sustained.

This emphasis on the long run is consistent with the evidence. Stock and Watson (1993, p. 811) found that “The empirical analysis suggests that the precise estimation of long-run M1 demand requires a long span of data: estimates over the full 90 years are considerably more precise than over the first half of the century alone, and when used in isolation the data since 1946 contain quite limited information about long-run M1 demand."

In this conception of the QTM, the longer time frame to which it applies simplifies rather than complicates its empirical implementation. With respect to the money supply, the quantity of money in circulation can be taken to be determined by the policies of the monetary authority: $M_t^S = M_t$. With respect to money demand, people desire to hold money, first and foremost, to facilitate their nominal transactions $P_t Y_t$. As a secondary consideration, they are influenced in this willingness by the opportunity cost of holding wealth in the form of money, given by the nominal interest rate r_t. The relevant interest rate is a short term one, because liquid interest-bearing assets like treasury bills or commercial paper are the most direct substitute for money.

This essential money demand function may be denoted $M_t^D = f(P_t Y_t, r_t)$. The equilibration of supply and demand determines the price level P_t, which adjusts to bring people's willingness to hold money into conformity with the quantity available to be held, $M_t^S = M_t^D$. In this time frame, P_t is the variable determined by the equilibrium condition

$$M_t = f(P_t Y_t, r_t). \tag{1}$$

The notion of a dichotomy between the real and nominal sectors is taken to be a reasonable approximation to reality: transactions Y_t are determined by real economic activity. The nominal interest rate r_t is determined by flows of saving and investment in financial markets, which reflect expectations of inflation based on past inflation.

Traditionally, empirical implementations of the equilibrium condition (1) have been cast in terms of the estimation of money demand, so that the quantity of money demanded is viewed as the endogenous variable explained by the theory. This may be appropriate when shorter-run dynamics are being studied. But that, in the long run, it is the price level P_t that is being determined, has important implications for the empirical implementation of the QTM. Vector error correction systems in which the variables of the levels relationship decompose according to such a causal structure have been studied rigorously by Pesaran et al. (2000, 2001; henceforth PSS). PSS (2001, Assumption 3) focuses on the case in which one variable—in our case P_t—has, in its level (not necessarily its change)—no contemporaneous feedback to the others. The other variables are, in this sense, "long run forcing" with respect to P_t. In this case, the equation describing their relationship can be decomposed from the rest of the system. So decomposed, PSS (2001) present bounds tests for the existence of a levels relationship among the variables, tests that are robust to the univariate integration properties of the variables.

Notice that this does not rule out non-contemporaneous effects of money on the other variables, of which a couple of possibilities suggest themselves even under a classical dichotomy between the real and monetary sectors. Most obviously, past money determines inflation and therefore the expectations of inflation embedded in the nominal interest rate r_t. Secondly, if the central bank follows an inflation target, then past prices can affect current money. But neither of these violates the assumption of an absence of contemporaneous feedback from P_t to the other variables of the levels relationship. In this respect the applicability of the PSS framework to the QTM is not just plausible, but compelling.

Let us now consider how this levels relationship relates to the traditional literature on the demand for money.

2.2. Empirical implementation

Given the central role of money demand in the quantity theory, empirical implementations of the QTM typically proceed using assumed specifications for the money demand function. A common assumption is to specify money demand as homogeneous in prices and income, in general not necessarily of the same degree, so that the equilibrium condition (1) is

$$M_t = k(r_t) P_t^{\theta_p} Y_t^{\theta_y}. \tag{2}$$

This describes prices as being determined according to

$$P_t = k(r_t)^{-1/\theta_p} M_t^{1/\theta_p} Y_t^{-\theta_y/\theta_p}.$$

In this case, the long run equilibrium relationship determining prices is loglinear in money and income,

$$p_t = \frac{1}{\theta_p} m_t - \frac{\theta_y}{\theta_p} y_t - \frac{1}{\theta_p} \log k(r_t),$$

where lower case p_t, m_t, and y_t denote the logarithms of the variables. In a stochastic setting this long run equilibrium relationship translates into a stationary equilibrium error

$$p_t + \lambda_m m_t + \lambda_y y_t + \lambda_k \log k(r_t), \tag{3}$$

where $\lambda_m = -1/\theta_p < 0$, $\lambda_y = \theta_y/\theta_p > 0$, and $\lambda_k = 1/\theta_p = -\lambda_m$. So the hypothesis that money demand is a demand for real rather than merely nominal money balances, $\theta_p = 1$, is the restriction $\lambda_m = -1$. The hypothesis of a unitary transactions elasticity of money demand, $\theta_y = 1$, is the restriction $-\lambda_m = \lambda_y$. These are the homogeneity propositions of the Introduction. Jointly, they are that $-\lambda_m = \lambda_y = 1$, in which case monetary equilibrium (2) is described by

$$\frac{M_t}{P_t Y_t} = k(r_t), \tag{4}$$

the maintained hypothesis of BLNW. If all four of these variables, including $\log k(r_t)$, are I(1), then the vector $\lambda = [1, \lambda_m, \lambda_y, \lambda_k]$ is the cointegrating vector. However this framework is not limited to that circumstance. It could be that just two or three of the variables are I(1), money and prices, say, or money, prices, and income, so that it is this subset that is cointegrated. In this respect the only necessary (although not sufficient) precondition for the equilibrium error (3) to be stationary is that if any of the variables is I(1), at least one other must be as well.

Which of these possibilities actually holds in practice will depend on the historical time period. With the demise of the gold standard in the early decades of the last century, prices p_t have tended to increase with the money supply m_t. So too have aggregate transactions y_t increased with economic growth. All three variables are therefore plausibly I(1) with drift. The appropriate specification for nominal interest rates is less clear. They increased during the 1970s, incorporating the rising inflation of that era, and so during that period would most likely be best approximated as I(1) with drift. But in the long run to which the quantity theory applies "... interest rates are almost certainly stationary in levels. Interest rates were about 6% in ancient Babylon; they are about 6% now. The chances of a process with a random walk component displaying this behavior are infinitesimal." (Cochrane, 1991, p. 208). Rogoff et al. (2022, section 2) provide a useful brief survey of the debate over the univariate properties of interest rates. Consistent with Cochrane's very long run view, they find that "In contrast to existing consensus, which has overwhelmingly concentrated on short samples of short-maturity rates, we find that long-maturity real interest rates across advanced economies are in fact trend stationary, and exhibit a persistent downward trend since the Renaissance." If this is the right way to think about the short-maturity rate r_t over the past century, the implication is that, in the equilibrium error (3) describing the QTM, any cointegrating relationship is between p_t, m_t, and y_t. To prevent any misunderstanding that all four variables in the equilibrium error (3) are necessarily assumed to be integrated, I call λ the *equilibrium vector* rather than the cointegrating vector.

Whereas BLNW adopt both homogeneity restrictions as a maintained hypothesis, empirical money demand studies using shorter-span data typically allow for more general possibilities, especially with respect to the income elasticity θ_y. Hence Ball (2001, equ. (1)) and McCallum and Nelson (2011, equ. (3)) describe the "canonical" or "standard" empirical money demand specification as

$$m_t - p_t = \beta + \theta_y y_t - \xi r_t. \tag{5}$$

This permits a non-unitary transactions elasticity θ_y, and specifies a semi-logarithmic function for the reciprocal of velocity: $\log k(r_t) = \beta - \xi r_t$.

As one example of work along these lines, Ball (2001) found that, for U.S. post-WWII annual data through 1996, the income elasticity θ_y is about 0.5 while the interest semi-elasticity $-\xi$ is about -0.05 (for interest rates expressed as percentages). Estimates like these are often compared against the theoretical benchmark of the classic Baumol-Tobin inventory-theoretic square root rule for money holding, which predicts an interest elasticity of money demand of -0.5. But, given aggregation issues, there is no particular reason to believe that micro-theoretic results will translate to the aggregate. At historical values for the interest rate, Ball's interest semi-elasticity of -0.05 is not particularly consistent with the square root rule.

Whether the homogeneity restrictions $-\lambda_m = \lambda_y = 1$ hold in practice will depend on the specifics of the empirical implementation, such as the data span and the choice of monetary aggregate. They are unlikely to hold in the short run, as is reflected in the canonical money demand function (5) and Ball's θ_y estimate of around 0.5. But they may hold in the long run. For example, for U.S. M1, 1900–1989, Stock and Watson (1993, p. 811) found that "Overall, the evidence is consistent with there being a single stable long-run demand for money, with an income elasticity near one ... " In short, if the homogeneity restrictions are to be employed, they should be subject to test. Let us now consider why it might be of special interest to work with a model where the homogeneity restrictions are compatible with the data.

2.3. *The welfare cost of inflation*

As we have emphasized, inflation and its welfare cost are inherently long run phenomena. And, as just argued, the homogeneity propositions are compelling as long run restrictions on the QTM. Combining these ideas, Lucas (2000) showed that closed-form expressions for the welfare cost of inflation can be derived when money demand is homogeneity-restricted and so monetary equilibrium is described by (4). He derived welfare cost formulas for two specifications of $k(r_t)$ that are commonly used in empirical work, a log-log functional form[1]

$$\log k(r_t) = \alpha - \eta \log r_t = \log(A) - \eta \log r_t \tag{6}$$

and a semi-log one,

$$\log k(r_t) = \beta - \xi r_t = \log(B) - \xi r_t. \tag{7}$$

The respective consumers' surplus welfare costs are[2]

[1] In periods of very low interest rates BLNW find that a third specification, the Selden-Latané functional form, fits the data best. Because its empirical relevance seems limited to these circumstances, I do not consider this variation on these functional forms. As well, it gives rise to a nonlinear-in-parameters form for the equilibrium error to which the PSS critical values do not strictly apply. For recent work that uses all three money demand specifications, see Benati (2020).

[2] Lucas (2000) contrasted the consumer's surplus definition of welfare cost with a compensating variation approach, and was able to derive a closed-form expression for the latter in the case of log-log money demand. Serletis and Yavari (2004) compare the two using post-WWII Canadian and U.S. data, and find the difference to be negligible.

$$w(r) = A\frac{\eta}{1-\eta}r^{1-\eta} \qquad \text{and} \qquad w(r) = \frac{B}{\xi}[1-(1+\xi r)e^{-\xi r}],$$

where "... $w(r)$ has the interpretation ... as the fraction of income people would require as compensation in order to make them indifferent between living in a steady state with an interest rate constant at r and an otherwise identical steady state with an interest rate of (or near) zero." (Lucas, 2000, pp. 250–251.)

These welfare costs assume that money pays no interest, so that the interest rate r on competing assets is the opportunity cost of holding money. This is only literally true of currency and coin and very limited categories of deposits. When money, however defined, pays interest, this opportunity cost is to some extent mitigated. But that, empirically, money demand is interest-elastic regardless of the broadness of the aggregate, indicates that this mitigation is incomplete: inflation causes people to economize on all categories of deposits, including interest-bearing ones. The partial mitigation of the welfare cost of inflation, when money pays interest, means that Lucas's formulas are best thought of as providing upper bounds on the true welfare costs, especially for broader aggregates.[3]

3. Data and preliminary evidence

Following BLNW, U.S. annual data are available from 1915; annual U.K. series begin in 1922, when the U.K. took on its modern borders of Great Britain plus Northern Ireland. For each country, three monetary aggregates are available for the century through 2019. In the case of the U.S., these are M1, M2, and an intermediate measure advocated by Lucas and Nicolini (2015, henceforth LN) that they call NewM1, which consists of M1 plus money market deposit account (MMDA) balances. For the U.K., the three aggregates are M1, M4, and an intermediate measure of "broad money" that I will denote Mbroad. Comparable quarterly series are available post-WWII, beginning in 1948 for the U.S. and in 1955 for the U.K. Detailed data descriptions and sources are given in the Data Appendix. The samples end in 2019 for several reasons. As of the time of writing, subsequent values are still subject to revision, and the price level effects of the monetary stimulation during the covid pandemic are still working themselves out, as noted in the opening paragraph of the Introduction. As well, in part owing to LN's contribution, beginning in 2020 U.S. M1 was redefined to include MMDA balances, and so is no longer comparable with earlier M1.

These monetary aggregates are typical of those that have traditionally been used in money demand studies. However, as simple-sum aggregates, they are subject to the Barnett (1980) critique that they fail to reflect the lack of perfect substitutability between different categories of deposits.[4] It would be preferable to use alternative money stock measures obtained as Divisia indexes. Unfortunately, Divisia monetary aggregates have been constructed beginning only in 1967 for the U.S. (from the Center for Financial Stability), and in 1977 for the U.K. (panel Q3 of the Bank of England's *Millennium* spreadsheet). They are therefore not yet available for the data spans ideal for studying the quantity theory, although Hendrickson (2014) is an admirable attempt.

3.1. Descriptive behavior

Key features of these data are shown in Figs. 1 (U.S. annual), 2 (U.S. quarterly), 3 (U.K. annual), and 4 (U.K. quarterly). Each figure has two panels: the three money stock measures; and the respective velocities and the short-term interest rate. In Figs. 1 and 2, money market deposit accounts were introduced in 1982, so M1 and NewM1 are coincident until then.

In the case of the velocities, the most obvious regularity is the extent to which those for the narrower aggregates (M1 and NewM1 in the case of the U.S., M1 in the case of the U.K.) are related to interest rates. Indeed, the historical peaks of the velocities for U.S. NewM1 (denoted V1new) and U.K. M1 (denoted V1) correspond virtually exactly with the historical peak in interest rates around 1980. Just as the reasoning behind the interest elasticity of money demand suggests, money circulates more rapidly the higher the opportunity cost of holding it, especially the more narrowly-defined aggregates that are most heavily used in exchange.

Another obvious regularity in these velocities is that they share an important feature: they have returned to values that today are not greatly different from what they were a century ago. Perhaps surprisingly, technological advance in the banking and financial sector has not systematically increased velocity—although one could easily have been misled into believing the contrary had one focused on U.S. NewM1 or U.K. M1 prior to the early 1980s, or U.S. M1 prior to the financial crisis.

The other variables exhibiting mean reversion are the interest rates which, from their highs around 1980, had returned to levels of the 1930s prior to the pandemic. This mean reversion is consistent with Cochrane's (1991) view, cited in section 2.2, that in the long run "... interest rates are almost certainly stationary ...," as well as the evidence in Rogoff et al. for long-maturity rates.

3.2. Univariate tests

Let us consider whether more rigorous inferential testing supports these descriptive interpretations. That growing macroeconomic variables are typically best described as unit root nonstationary is now so well established that formal testing might be viewed as superfluous. But a brief summary of some unit root evidence on the variables of the QTM usefully anticipates some of the issues that arise in the subsequent multivariate analysis.

Table 1 reports unit root tests of both the null hypothesis of a unit root, and the null of the absence of a unit root. For the former, from among the many tests of a unit root null that have been developed, I follow BLNW and Rogoff et al. in using the DF-GLS test of

[3] I thank an anonymous referee for alerting me to this qualification.

[4] The Barnett critique may be less applicable to NewM1, which Lucas and Nicolini (2015) motivate with a model that predicts that it can be legitimately constructed as a simple-sum aggregate.

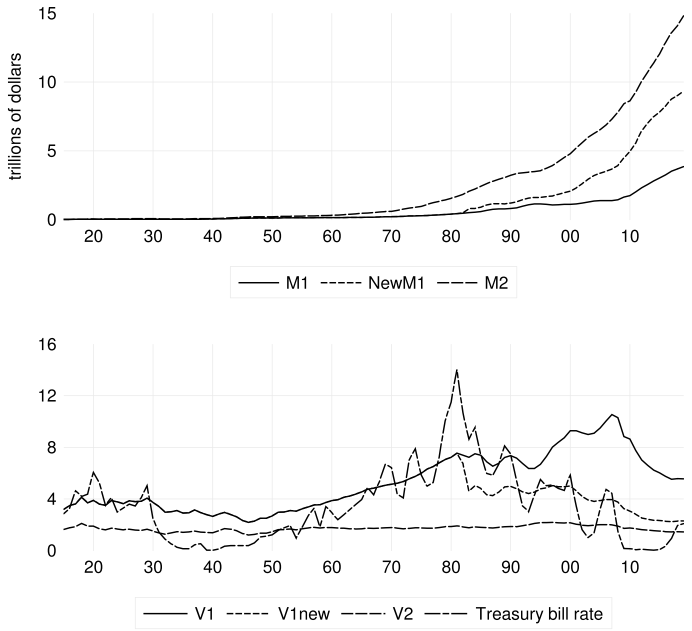

Fig. 1. U.S. monetary aggregates, velocities, and the treasury bill rate, annually 1915–2019.

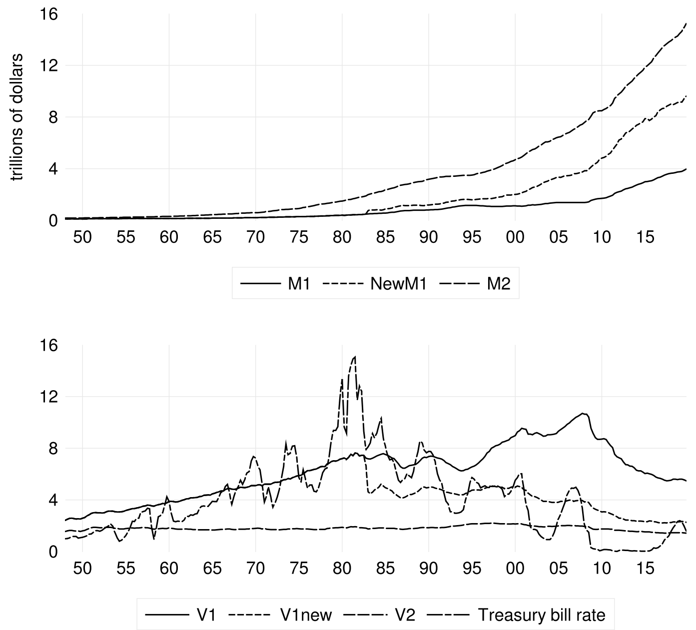

Fig. 2. U.S. monetary aggregates, velocities, and the treasury bill rate, quarterly 1948–2019.

Elliot et al. (1996). For the null of the absence of a unit root, the standard test is that of Kwiatkowski et al. (1992) (henceforth KPSS). Bolded values of the test statistics in Table 1 indicate a rejection at a 10 percent significance level; that is, the weakest evidence that a researcher using conventional significance levels would interpret as rejecting the null. (Critical values for other conventional significance levels are given in the table footnotes.) For variables that appear in both log and non-log form (the interest rate and velocities), it is reassuring that test decisions are insensitive to the log transformation, in both the annual and quarterly data of both countries.

Given that annual data are essentially smoothed versions of the quarterly series, we might expect to find stronger unit root evidence in the quarterly data, and indeed this is so: in the quarterly data the DF-GLS tests almost uniformly fail to reject a unit root, while the KPSS tests almost uniformly reject the absence of a unit root.

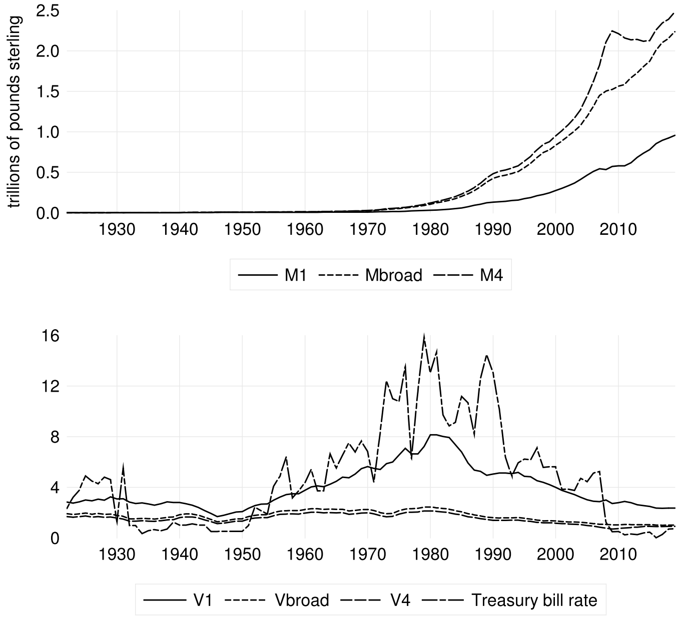

Fig. 3. U.K. monetary aggregates, velocities, and the treasury bill rate, annually 1922–2019.

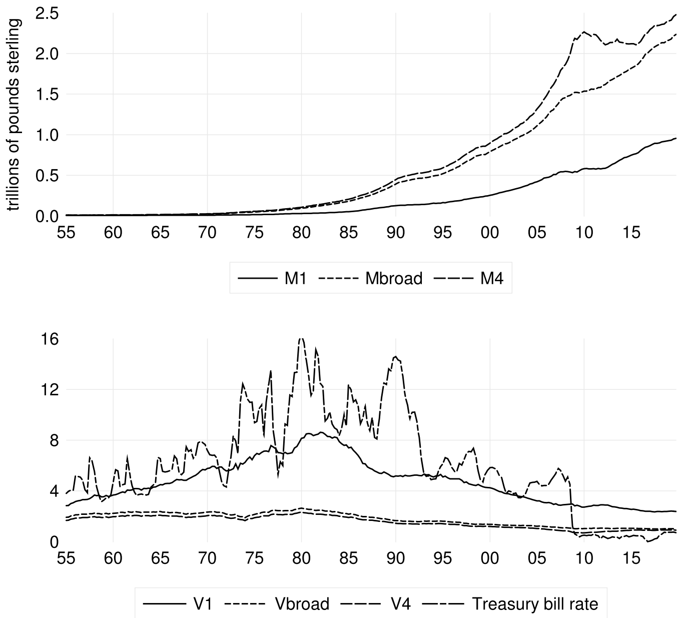

Fig. 4. U.K. monetary aggregates, velocities, and the treasury bill rate, quarterly 1955–2019.

In particular, both the U.S. and U.K. quarterly series uniformly support the view that the growing variables p_t, y_t, and the monetary aggregates should have unit roots, and so are best described as difference stationary rather than trend stationary. The same tendency dominates the annual data, the principal exceptions being U.S. M1, where a unit root is rejected, and y_t for both the U.S. and U.K.

A priori expectations are less obvious for the non-growing variables, the interest rate and velocity. For interest rates, in the case of the annual data the evidence is consistent with Cochrane's conjecture and Rogoff et al., less so for the quarterly data. For the velocities, despite the apparent mean reversion of the figures, there is considerable evidence of unit root behavior in both the annual and quarterly data, especially in the U.K. velocities. Unit roots in the U.S. velocities are most clearly evident in the narrower aggregates M1 and NewM1, a finding consistent with BLNW and Benati (2020). However, this changes with the annual M2 velocity, where a unit root is rejected.

Table 1
Univariate unit root tests.

	Annual data		Quarterly data	
	DF-GLS statistic[a]	KPSS statistic[b]	DF-GLS statistic[a]	KPSS statistic[b]
United States				
log GDP deflator, p_t	− 2.6094	**0.1817**	− 1.1598	**0.3088**
log real GDP, y_t	**−2.8848**	**0.1843**	− 0.7697	**0.4034**
interest rate, r_t	**−1.9186**	0.2251	− 1.2212	**0.4503**
log r_t	**−2.0087**	0.1520	− 1.3586	**0.5584**
Money stock, m_t:				
log M1	**−3.0392**	0.0418	− 0.7144	**0.1782**
log NewM1	− 2.2246	**0.1766**	− 0.8160	**0.3327**
log M2	− 2.7115	0.1065	− 1.0403	**0.3486**
Velocity, $V_t = P_tY_t/M_t$:				
V1	− 0.9498	**0.8904**	− 0.4514	**1.4693**
NewV1	− 1.3609	0.2731	− 0.7805	**0.4722**
V2	**−2.1404**	**0.4330**	− 1.0389	0.2696
Log velocity, $\log k_t = -\log V_t$:				
log K1	− 0.6278	**0.8910**	0.0040	**1.5451**
log NewK1	− 1.1271	0.2489	− 0.6408	**0.4992**
log K2	**−2.3265**	**0.4300**	− 0.9799	0.2585
United Kingdom				
log GDP deflator, p_t	− 1.9534	**0.1617**	− 0.8286	**0.4477**
log real GDP, y_t	**−3.3565**	0.1189	− 1.3717	**0.2731**
interest rate, r_t	**−2.1625**	0.3216	**−1.6515**	**0.7793**
log r_t	**−2.5618**	0.2289	**−1.6828**	**1.0115**
Money stock, m_t:				
log M1	− 1.2683	**0.2586**	− 0.9944	**0.2321**
log Mbroad	− 1.2119	**0.2310**	− 0.8076	**0.3966**
log M4	− 1.3243	**0.2361**	− 0.8115	**0.3676**
Velocity, $V_t = P_tY_t/M_t$:				
V1	− 1.1673	**0.3486**	− 0.7101	**0.7902**
Vbroad	− 1.0486	**0.4724**	− 0.0707	**1.8170**
V4	− 1.1316	**0.4865**	− 0.2277	**1.8007**
Log velocity, $\log k_t = -\log V_t$:				
log K1	− 1.0415	**0.3638**	− 0.5099	**0.9184**
log Kbroad	− 0.6922	**0.5426**	− 0.4876	**1.8524**
log K4	− 1.0052	**0.5541**	− 0.0907	**1.7945**

Notes: Growing variables (p_t, y_t, m_t) use an intercept-and-trend specification; non-growing variables (r_t, V_t, and $\log k_t$) use an intercept-no-trend specification. Bolded values indicate a rejection at a 10% significance level.

[a] DF-GLS statistics test the null of a unit root, and use a BIC-selected lag length. Critical values vary with the sample size and lag length, but are approximately −1.6 (10%), −1.9 (5%), and −2.6 (1%) in a constant-no-trend specification, and −2.7 (10%), −3.0 (5%), and −3.6 in a constant-plus-trend specification. Bolded values indicate 10% rejections based on actual sample sizes and lag lengths.

[b] KPSS statistics test the null of stationarity (in a constant-no-trend specification) or trend stationarity (in a constant-plus-trend specification), and use a Bartlett kernel and a Newey-West selected bandwidth. Constant-no-trend critical values are 0.347 (10%), 0.463 (5%), and 0.739 (1%). Constant-plus-trend critical values are 0.119 (10%), 0.146 (5%), and 0.216 (1%).

These univariate unit root results are instructive in suggesting what might be reasonable to believe about the nature of the cointegration underlying these data.

If one agrees with Cochrane and Rogoff et al. that interest rates are, in the long run, best thought of as stationary, then the QTM requires that cointegration be among p_t, m_t, and y_t, so that log velocity is stationary. On the basis of the preliminary evidence of Table 1, it is the U.S. annual data for the broader aggregates NewM1 and M2 that come closest to conforming to this story. The tests support stationarity of the interest rate; the KPSS test fails to reject stationarity of NewM1 velocity (although, at the same time, the DF-GLS test fails to reject a unit root); and the DF-GLS test rejects a unit root in M2 velocity (although the KPSS test rejects stationarity).

If, on the other hand, one takes the BLNW view that, at least during the past century, interest rates are best described as integrated, then the QTM requires that cointegration be between the interest rate and velocity, so that velocity is integrated. The U.S. quarterly data for M1 and NewM1 conform with this story: according to both the DF-GLS and KPSS tests, the interest rate and both velocities have unit roots. U.S. quarterly M2 might conform: the DF-GLS test does not reject a unit root in M2 velocity (although the KPSS test does not reject stationarity).

The U.K. annual data contradict either of these stories: they suggest a stationary interest rate, but unit root velocities. The U.K. quarterly data are closer to the BLNW view: the velocities are unambiguously unit root, and the KPSS test rejects stationarity of the interest rate (although the DF-GLS test rejects a unit root).

Table 2
Cointegration bounds tests on AIC and BIC-selected ADL/ECM: annual data.

ADL/ECM(·, ·, ·, ·)	LM(4)[a] (*p*-value)	*F* statistic[b]	*t* statistic[c]
United States 1915–2019			
Loglinear money demand			
M1			
AIC(5,4,0,3)	0.0023	6.2364	**−3.7314**
BIC(1,1,0,1)	0.0116	9.4937	**−3.1205**
NewM1			
AIC(5,1,0,1)	0.0005	10.0403	**−3.0224**
BIC(2,1,0,0)	0.0736	3.5044	**−3.0191**
M2			
AIC(2,4,0,4)	0.4570	10.2910	**−5.4620**
BIC(2,2,0,3)	0.0056	6.4335	**−4.5170**
Semi-logarithmic money demand			
M1			
AIC(5,4,0,0)	0.0009	8.4049	**−4.4549**
BIC(1,1,0,0)	0.0085	8.8785	**−3.5058**
NewM1			
AIC(6,1,0,0)	0.0000	17.6598	**−6.2567**
BIC(1,1,0,0)	0.0036	9.4371	**−4.8316**
M2			
AIC(3,2,0,0)	0.0357	6.3293	**−4.0756**
BIC(3,2,0,0)	0.0357	6.3293	**−4.0756**
United Kingdom 1922–2019			
Loglinear money demand			
M1			
AIC(5,4,4,2)	0.7322	3.6800	−1.3434
BIC(5,0,0,0)	0.3476	3.5216	−1.9927
Mbroad			
AIC(6,0,2,1)	0.9917	7.0221	**−5.3995**
BIC(5,0,0,0)	0.1793	5.3886	**−3.5827**
M4			
AIC(6,0,2,1)	0.9918	5.7481	**−4.1124**
BIC(5,0,0,0)	0.2335	4.8199	**−3.2812**
Semi-logarithmic money demand			
M1			
AIC(5,3,2,1)	0.3913	9.5287	**−5.2232**
BIC(5,0,0,1)	0.1584	9.9746	**−5.1426**
Mbroad			
AIC(6,3,2,2)	0.3287	14.0601	**−6.9637**
BIC(5,0,0,1)	0.2118	14.5267	**−6.3464**
M4			
AIC(6,1,2,1)	0.5989	14.7641	**−7.0417**
BIC(5,0,0,1)	0.2972	12.8918	**−5.9050**

Notes: F and t tests are of the null of absence of cointegration in Pesaran-Shin-Smith Case III of unrestricted intercept and no trend.

[a] Breusch-Godfrey Lagrange multiplier test for residual autocorrelation in the ADL/ECM, four lags.

[b] F test 10% critical values for PSS Case III: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)

[c] t test 10% critical values for PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)

To this extent, the U.S. and U.K. quarterly series are most consistent with the BLNW view. But whether cointegration and, in turn, estimation results that yield sound inferences about the QTM, emerge from these data may be compromised by their shorter span than the annual data.

Given the notorious size and power weaknesses of unit root tests, it would be a mistake to read too much into these univariate results. Yet they turn out to anticipate much of what is revealed by the ADL/ECM models of Tables 2 through 9. The strongest results, in terms of compatibility with the homogeneity restrictions and, in turn, a basis for credible welfare cost calculations, emerge from the U.S. annual data using M2 as the monetary aggregate, where the univariate tests indicate both the interest rate and M2 velocity to be plausibly stationary. The other data and model permutations yield less convincing findings, but the U.K. annual and U.S. quarterly data still yield results for some monetary aggregates that can be usefully compared with the U.S. annual results. The U.K. quarterly data turn out to yield little that is persuasive, something that is predicted by the preliminary cointegration tests that we are about to consider.

Table 3
Cointegration bounds tests on AIC and BIC-selected ADL/ECM: quarterly data.

ADL/ECM(·, ·, ·, ·)	LM(4)[a] (*p*-value)	*F* statistic[b]	*t* statistic[c]
United States 1948–2019			
Loglinear money demand			
M1			
AIC(4,1,2,0)	0.1325	3.5916	−1.8434
BIC(3,0,0,0)	0.0763	3.6486	−1.6069
NewM1			
AIC(4,0,2,1)	0.0774	4.1787	−0.6153
BIC(1,2,5,1)	0.1251	4.8770	−0.5251
M2			
AIC(3,0,4,0)	0.0515	3.6515	**−2.7503**
BIC(3,0,0,0)	0.0634	4.0894	**−2.7839**
Semi-logarithmic money demand			
M1			
AIC(4,1,2,0)	0.1640	6.7805	**−3.0630**
BIC(3,0,0,0)	0.1882	7.5740	**−3.0737**
NewM1			
AIC(4,0,2,0)	0.0730	6.1639	−2.3353
BIC(3,0,0,0)	0.1841	7.4182	−2.4109
M2			
AIC(3,0,4,0)	0.1237	7.4313	**−3.3756**
BIC(3,0,0,0)	0.1643	8.0452	**−3.4065**
United Kingdom 1955–2019			
Loglinear money demand			
M1			
AIC(5,3,2,0)	0.8176	7.4114	−0.6222
BIC(4,0,1,0)	0.4459	6.9885	0.1902
Mbroad			
AIC(4,0,1,0)	0.4142	5.8834	1.1036
BIC(5,0,1,0)	0.4677	5.8045	0.9033
M4			
AIC(5,0,1,0)	0.4139	5.5142	0.3013
BIC(4,0,1,0)	0.4410	5.5351	0.5007
Semi-logarithmic money demand			
M1			
AIC(5,4,1,0)	0.8214	11.4160	**−2.8658**
BIC(4,0,1,0)	0.5836	11.5405	−2.2863
Mbroad			
AIC(5,4,1,0)	0.2249	9.1259	−1.2216
BIC(4,0,1,0)	0.6028	11.2479	−1.0436
M4			
AIC(5,4,1,0)	0.1944	10.3356	−2.0941
BIC(4,0,1,0)	0.6222	11.1665	−1.7307

Notes: F and t tests are of the null of absence of cointegration in Pesaran-Shin-Smith Case III of unrestricted intercept and no trend.
[a] Breusch-Godfrey Lagrange multiplier test for residual autocorrelation in the ADL/ECM, four lags.
[b] F test 10% critical values for PSS Case III: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[c] t test 10% critical values for PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)

Another reason why it would be a mistake to dwell on the univariate velocity findings is that doing so accepts the homogeneity restrictions as a maintained hypothesis. Instead we should turn to an empirical framework in which these are testable, in order to focus on monetary aggregates for which they are most likely to be satisfied.

The other key lesson from the unit root tests of Table 1 is the desirability of using methodology that is robust to the univariate properties of the series, rather than being dependent on particular assumptions about them. Once again, this motivates the use of PSS bounds tests, which are agnostic on how many of the variables in the hypothesized cointegrating relationship are I(1). By contrast, more conventional Johansen testing within a VAR system assumes that all variables are I(1).[5]

[5] One reference on this point is Johansen (1995). The assumption/requirement that all variables in the system are I(1) appears in both the specification of the model (Chapter 4) and the analysis of the distributional properties of the statistics (Chapter 11). In specifying the ECM for the vector X_t, Theorem 4.2 requires that "… X_t is a cointegrated I(1) process … "; that is, the individual components of X_t are all I(1). Later, in Chapter 11, Johansen develops the sampling properties of the trace and maximum eigenvalue test statistics for cointegrating rank. He states explicitly (p. 156) that, in the derivation of these sampling properties, "… we assume throughout that … the process X_t is I(1)."

Table 4
U.S. annual 1915–2019 ADL/ECM under log-log money demand: $\log k(r_t) = \log A - \eta \log r_t$.

ECM(·, ·, ·, ·) equilibrium error: $p_t + \lambda_m m_t + \lambda_y y_t - \lambda_m(\alpha - \eta \log r_t)$	M1 (2,4,0,3)	NewM1 (2,4,0,3)	M2(2,5,0,3) no structural change	M2(2,5,0,3) structural change[g]
Parameters of the error correction term:				
speed-of-adjustment coefficient[a], ψ	−0.0674	−0.0706	−0.1472	−0.0731
	(0.0158)	(0.0289)	(0.0302)	(0.0261)
interest elasticity[b], $-\eta$	−0.3382	−0.5529	−0.0840	−0.0956
	(0.1649)	(0.6438)	(0.0161)	(0.0249)
real transactions, λ_y	−0.1826	−0.5982	0.7068	0.8568
	(0.3835)	(0.4910)	(0.1570)	(0.3827)
nominal money, $-\lambda_m$	0.5058	0.1913	0.8857	0.9316
	(0.2412)	(0.2515)	(0.0833)	(0.1838)
PSS tests of null of no levels relationship:				
F statistic[c]	8.6157	5.3705	7.5720	4.3281
t statistic[d]	−4.2650	−2.4470	−4.8805	−2.7978
LM test for residual autocorrelation[e] (p-value)	0.9099	0.6077	0.2947	0.0015
Wald tests of homogeneity restrictions (p-values):				
(1) Real money demand ($\theta_p = 1$): $\lambda_m = -1$	0.0435	0.0018	0.1737	0.7109
(2) Unitary income elasticity ($\theta_y = 1$): $\lambda_y = -\lambda_m$	0.0000	0.0016	0.0190	0.7103
(3) Joint test of (1) and (2): $-\lambda_m = \lambda_y = 1$	0.0000	0.0065	0.0001	0.9325
Estimate of λ_y when $\lambda_m = -1$ imposed	0.5809	1.5291	0.9993	0.9190
	(0.0483)	(1.7572)	(0.0398)	(0.0182)
p-value for $\lambda_y = 1$	0.0000	0.7640	0.9869	0.0000
Restricted equilibrium error ($p_t - m_t + y_t + \log A - \eta \log r_t$) imposes $-\lambda_m = \lambda_y = 1$:				
speed-of-adjustment coefficient[a], ψ	−0.0017	0.0135	−0.0799	−0.0692
	(0.0079)	(0.0234)	(0.0265)	(0.0203)
interest elasticity[b], $-\eta$	−4.4596	0.0715	−0.1051	−0.0955
	(19.9016)	(0.4713)	(0.0291)	(0.0219)
intercept parameter[f], A	0.0000	0.4802	0.4327	0.5136
	(0.0000)	(1.1193)	(0.0509)	(0.0525)
Naive money demand estimates from static estimation of $\log k_t = \log A - \eta \log r_t$:				
interest elasticity[b], $-\eta$	−0.0932	−0.1520	−0.0501	
	(0.0262)	(0.0126)	(0.0070)	
intercept parameter[f], A	0.1428	0.1440	0.4840	
	(0.0160)	(0.0078)	(0.0145)	
Credible welfare cost estimates yielded by restricted ECM (percent of income): $100 \times w(r)$				
at $r = 0.03$			0.2205	0.2274
at $r = 0.05$			0.3482	0.3609
at $r = 0.06$			0.4099	0.4256
at $r = 0.13$			0.8188	0.8565
at $r = 0.14$			0.8750	0.9158

Notes: Standard errors are in parentheses.
[a] Convergence to equilibrium requires $-1 < \psi < 0$.
[b] For comparison, the Baumol-Tobin square-root rule for money holding predicts an interest elasticity of −0.5.
[c] 10% critical values for PSS Case III are: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[d] 10% critical values are PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)
[e] Reported tests are based on 4 lags, but the test decisions are generally robust to variations on the lag length.
[f] Obtained by nonlinear least squares estimation of the ECM.
[g] Indicator saturation finds impulse indicators in 1920 and 1921, and step indicators in 1931, 1933, 1938, and 1939.

3.3. Preliminary cointegration testing

As a first step toward models in which the homogeneity restrictions can be tested, Tables 2 and 3 present PSS cointegration tests from ADL/ECM models selected mechanically by information criteria. As is well known, Akaike's information criterion (AIC) tends to select models with longer lag lengths, while Schwarz's Bayesian information criterion (BIC) tends to select more parsimonious ones.

Intuitively, the PSS methodology tests whether variables of the ECM are guided in their evolution by an equilibrium relationship in their levels, as opposed to being influenced by nothing more than short run interactions among their first differences. This has two testable implications for the coefficients of the ECM. The first is that, in the hypothesized equilibrium error, the coefficients on the levels of the variables should jointly be non-zero; this is tested with an F or Wald statistic. The second implication is that the speed-of-adjustment coefficient should be non-zero; this is tested with a t statistic. Although these F and t statistics are standard, their distributions are not. PSS (2001) tabulate critical value bounds on each, for the polar extremes of all I(0) variables versus all I(1) variables. They recommend that the two statistics be interpreted sequentially: "... we suggest the following procedure for ascertaining the existence of a level relationship ...: test H_0 ... using the bounds procedure based on the Wald or F-statistic ...: (a) if H_0 is not rejected, proceed no further; (b) if H_0 is rejected, test ... using the bounds procedure based on the t-statistic ... " (PSS

Table 5
U.S. annual 1915–2019 ADL/ECM under semi-log money demand: $\log k(r_t) = \log B - \xi r_t$.

			M2(2,4,0,1)	
ECM(·, ·, ·, ·) equilibrium error: $p_t + \lambda_m m_t + \lambda_y y_t - \lambda_m(\beta - \xi r_t)$	M1 (2,4,0,1)	NewM1 (2,4,0,1)	no structural change	structural change[g]
Parameters of the error correction term:				
speed-of-adjustment coefficient[a], ψ	−0.0667	−0.1022	−0.1225	−0.0701
	(0.0157)	(0.0274)	(0.0268)	(0.0242)
interest semielasticity[b], $-\xi$	−14.3398	−16.0113	−4.4792	−6.1700
	(6.2611)	(4.1651)	(1.2706)	(1.8706)
real transactions, λ_y	−0.0833	−0.1185	0.5167	0.5105
	(0.3811)	(0.2302)	(0.1916)	(0.3508)
nominal money, $-\lambda_m$	0.5617	0.4323	0.7878	0.7707
	(0.2395)	(0.1187)	(0.1018)	(0.1678)
PSS tests of null of no levels relationship:				
F statistic[c]	7.9116	6.7672	7.8533	5.2519
t statistic[d]	−4.2550	−3.7370	−5.6974	−2.8930
LM test for residual autocorrelation[e] (p-value)	0.6224	0.6684	0.1810	0.0231
Wald tests of homogeneity restrictions (p-values):				
(1) Real money demand ($\theta_p = 1$): $\lambda_m = -1$	0.0706	0.0000	0.0399	0.1757
(2) Unitary income elasticity ($\theta_y = 1$): $\lambda_y = -\lambda_m$	0.0000	0.0000	0.0039	0.1635
(3) Joint test of (1) and (2): $-\lambda_m = \lambda_y = 1$	0.0000	0.0000	0.0003	0.3763
Estimate of λ_y when $\lambda_m = -1$ imposed	0.5960	0.8034	0.9040	0.9939
	(0.0494)	(0.2365)	(0.0264)	(0.0527)
p-value for $\lambda_y = 1$	0.0000	0.4080	0.0004	0.9088
Restricted equilibrium error $(p_t - m_t + y_t + \log B - \xi r_t)$ imposes $-\lambda_m = \lambda_y = 1$:				
speed-of-adjustment coefficient[a], ψ	−0.0036	−0.0187	−0.0697	−0.0525
	(0.0082)	(0.0228)	(0.0256)	(0.0175)
interest semielasticity[b], $-\xi$	−98.4903	−20.5553	−5.4431	−6.0972
	(211.9886)	(15.4556)	(1.8637)	(1.7328)
intercept parameter[f], B	4.6314	0.4242	0.7544	2.1005
	(31.7724)	(0.1796)	(0.0792)	(1.3725)
Naive money demand estimates from static estimation of $\log k_t = \log B - \xi r_t$:				
interest elasticity[b], $-\xi$	−6.8774	−8.9435	−2.5399	
	(1.2911)	(0.5304)	(0.3757)	
intercept parameter[f], B	0.2635	0.3611	0.6460	
	(0.0153)	(0.0086)	(0.0109)	
Credible welfare cost estimates yielded by restricted ECM (percent of income): $100 \times w(r)$				
at $r = 0.03$			0.1658	
at $r = 0.05$			0.4290	
at $r = 0.06$			0.5963	
at $r = 0.13$			2.1960	
at $r = 0.14$			2.4618	

Notes: Standard errors are in parentheses.
[a] Convergence to equilibrium requires $-1 < \psi < 0$.
[b] For comparison, Ball (2001) reports a U.S. interest semielasticity $-\xi/100$ of approximately −0.05. The interpretation is that a permanent one percentage point increase in the interest rate decreases money demand by five percent and, under the homogeneity restriction that $\lambda_m = -1$, increases the price level by five percent.
[c] 10% critical values for PSS Case III are: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[d] 10% critical values are PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)
[e] Reported tests are based on 4 lags, but the test decisions are generally robust to variations on the lag length.
[f] Obtained by nonlinear least squares estimation of the ECM.
[g] Indicator saturation finds impulse indicators in 1920, 1921, 1931, and 1932, and step indicators in 1922, 1931, 1938, 1941, and 1943.

2001, p. 304). Intuitively, begin by testing whether the variables belong in the model in their levels; if they do, then test whether short run movements in the dependent variable are driven in part by adjustment to the levels relationship.

For example, for an equilibrium error without a time trend (Case III in the PSS taxonomy), the 10% critical bounds for the F statistic are 2.72 if all variables are I(0) and 3.77 if all are I(1) (PSS Table CI(iii)). A researcher who believes that, along with the other variables, interest rates are best treated as I(1), would insist on evidence as strong as $F > 3.77$ to reject the null at a 10% significance level. One who believes that interest rates are I(0) would accept somewhat weaker evidence, but would still require something stronger than $F > 2.72$ given that the other variables—money, prices, and income—are naturally regarded as I(1). Similar bounds hold for the PSS t statistic.

For the U.S. and U.K. annual data (Table 2), all the PSS F tests pass the minimal threshold (10 percent significance, all I(0) variables) for cointegration, and many are much stronger. And almost all the t tests pass the minimum threshold. The sole exception is U.K. M1 based on log-log money demand (i.e., the interest rate appears as $\log r_t$ rather than r_t), which anticipates the weak results for this aggregate that I find in Table 6. Remarkably (considering their very different testing methodology and slightly different

Table 6
U.K. annual 1922–2019 ADL/ECM under log-log money demand: $\log k(r_t) = \log A - \eta \log r_t$.

ECM(·, ·, ·, ·) equilibrium error: $p_t + \lambda_m m_t + \lambda_y y_t - \lambda_m(\alpha - \eta \log r_t)$	M1 (4,1,2,0)	Mbroad(5,1,1,1) no structural change	Mbroad(5,1,1,1) structural change[g]	M4 (5,0,2,1)
Parameters of the error correction term:				
speed-of-adjustment coefficient[a], ψ	−0.0205	−0.1038	−0.1126	−0.0840
	(0.0139)	(0.0287)	(0.0226)	(0.0241)
interest elasticity[b], $-\eta$	0.2824	−0.1496	−0.1619	−0.1574
	(1.2485)	(0.0364)	(0.0429)	(0.0444)
real transactions, λ_y	−2.8083	−0.0935	−0.5700	−0.1023
	(2.2467)	(0.2750)	(0.2866)	(0.3309)
nominal money, $-\lambda_m$	0.2429	0.6066	0.3893	0.5976
	(0.7511)	(0.0807)	(0.1050)	(0.0959)
PSS tests of null of no levels relationship:				
F statistic[c]	2.4282	5.4726	10.7092	4.3709
t statistic[d]	−1.4680	−3.6231	−4.9780	−3.4885
LM test for residual autocorrelation[e] (p-value)	0.1830	0.2833	0.5758	0.5543
Wald tests of homogeneity restrictions (p-values):				
(1) Real money demand ($\theta_p = 1$): $\lambda_m = -1$	0.1018	0.0000	0.0000	0.0001
(2) Unitary income elasticity ($\theta_y = 1$): $\lambda_y = -\lambda_m$	0.0908	0.0006	0.0000	0.0040
(3) Joint test of (1) and (2): $-\lambda_m = \lambda_y = 1$	0.2227	0.0000	0.0000	0.0000
Estimate of λ_y when $\lambda_m = -1$ imposed	1.6191	1.4347	0.9651	1.3954
	(2.9525)	(0.3929)	(0.2254)	(0.2155)
p-value for $\lambda_y = 1$	0.8344	0.2718	0.8773	0.0703
Restricted equilibrium error ($p_t - m_t + y_t + \log A - \eta \log r_t$) imposes $-\lambda_m = \lambda_y = 1$:				
speed-of-adjustment coefficient[a], ψ	−0.0039	−0.0037	−0.0333	−0.0020
	(0.0121)	(0.0137)	(0.0218)	(0.0117)
interest elasticity[b], $-\eta$	−0.3912	−0.0256	−0.1093	−0.4324
	(0.8041)	(0.7097)	(0.0595)	(2.2201)
intercept parameter[f], A	0.0466	1.2621	0.3532	0.0529
	(0.2168)	(6.3803)	(0.1129)	(0.7920)
Naive money demand estimates from static estimation of $\log k_t = \log A - \eta \log r_t$:				
interest elasticity[b], $-\eta$	−0.2374	−0.1284		−0.1377
	(0.0186)	(0.0158)		(0.0176)
intercept parameter[f], A	0.1172	0.3795		0.4237
	(0.0084)	(0.0231)		(0.0286)
Credible welfare cost estimates yielded by restricted ECM (percent of income): $100 \times w(r)$				
at $r = 0.03$			0.1908	
at $r = 0.05$			0.3007	
at $r = 0.06$			0.3537	
at $r = 0.13$			0.7042	
at $r = 0.14$			0.7522	

Notes: Standard errors are in parentheses.
[a] Convergence to equilibrium requires $-1 < \psi < 0$.
[b] For comparison, the Baumol-Tobin square-root rule for money holding predicts an interest elasticity of −0.5.
[c] 10% critical values for PSS Case III are: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[d] 10% critical values are PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)
[e] Reported tests are based on 4 lags, but the test decisions are generally robust to variations on the lag length.
[f] Obtained by nonlinear least squares estimation of the ECM.
[g] Indicator saturation finds impulse indicators in 1975 and 1980, and a step indicator in 1974.

sample period), it is also consistent with BLNW (Table 2), who found cointegration for U.S. M1 and NewM1 based on log-log money demand, but not for U.K. M1.

For the quarterly data (Table 3) the evidence is more mixed. This is especially so for the U.K., perhaps reflecting its shorter 1955–2019 span. The minimal test threshold is passed by only a single model: a semi-log specification (i.e., the interest rate appears as r_t rather than $\log r_t$) for M1 demand embedded within the ADL/ECM yielded by the AIC. We will examine the detailed results for this model in Table 9.

The U.S. quarterly cointegration evidence in Table 3 is stronger, especially for M2. A cautionary note is the sensitivity of some results to what should be innocuous specification changes. For the U.S. M1 models of Table 3, the AIC and BIC each yield identical lag specifications, regardless of whether the interest rate appears in log form (log-log money demand) or untransformed (semi-log money demand). Yet the semi-log model yields considerably stronger evidence of cointegration. The more detailed specification searches across models and monetary aggregates, described in the next section and presented in Tables 8 and 9, find M2 to be the only U.S. quarterly aggregate yielding credible welfare costs.

Table 7
U.K. annual 1922–2019 ADL/ECM under semi-log money demand: $\log k(r_t) = \log B - \xi r_t$.

ECM(·, ·, ·, ·) equilibrium error: $p_t + \lambda_m m_t + \lambda_y y_t - \lambda_m(\beta - \xi r_t)$	M1 (4,3,0,2)	Mbroad (5,0,0,1)	M4 (5,1,2,1)
Parameters of the error correction term:			
speed-of-adjustment coefficient[a], ψ	−0.0657	−0.1359	−0.1420
	(0.0196)	(0.0214)	(0.0222)
interest semielasticity[b], $-\xi$	−14.4747	−7.2098	−7.5029
	(3.1939)	(1.0696)	(1.0178)
real transactions, λ_y	−0.4948	−0.1550	−0.0324
	(0.3519)	(0.1577)	(0.1531)
nominal money, $-\lambda_m$	0.5211	0.5806	0.6114
	(0.1149)	(0.0460)	(0.0441)
PSS tests of null of no levels relationship:			
F statistic[c]	4.2902	14.5267	12.4780
t statistic[d]	−3.3497	−6.3464	−6.3884
LM test for residual autocorrelation[e] (p-value)	0.3211	0.1621	0.1262
Wald tests of homogeneity restrictions (p-values):			
(1) Real money demand ($\theta_p = 1$): $\lambda_m = -1$	0.0001	0.0000	0.0000
(2) Unitary income elasticity ($\theta_y = 1$): $\lambda_y = -\lambda_m$	0.0001	0.0000	0.0000
(3) Joint test of (1) and (2): $-\lambda_m = \lambda_y = 1$	0.0003	0.0000	0.0000
Estimate of λ_y when $\lambda_m = -1$ imposed	1.0885	1.4410	1.3935
	(0.2468)	(0.2281)	(0.1132)
p-value for $\lambda_y = 1$	0.7208	0.0566	0.0008
Restricted equilibrium error ($p_t - m_t + y_t + \log B - \xi r_t$) imposes $-\lambda_m = \lambda_y = 1$:			
speed-of-adjustment coefficient[a], ψ	−0.0192	−0.0009	−0.0028
	(0.0157)	(0.0116)	(0.0107)
interest semielasticity[b], $-\xi$	−11.7179	−138.5004	−24.7839
	(4.4225)	(1816.3330)	(87.4552)
intercept parameter[f], B	0.3489	0.0136	1.6960
	(0.0926)	(0.7279)	(4.8523)
Naive money demand estimates from static estimation of $\log k_t = \log B - \xi r_t$:			
interest elasticity[b], $-\xi$	−8.4946	−3.5449	−3.8063
	(0.4660)	(0.5540)	(0.6125)
intercept parameter[f], B	0.4144	0.7149	0.8356
	(0.0121)	(0.0248)	(0.0320)
Credible welfare cost estimates yielded by restricted ECM (percent of income): $100 \times w(r)$			
at $r = 0.03$	0.1460		
at $r = 0.05$	0.3493		
at $r = 0.06$	0.4671		
at $r = 0.13$	1.3396		
at $r = 0.14$	1.4531		

Notes: Standard errors are in parentheses.
[a] Convergence to equilibrium requires $-1 < \psi < 0$.
[b] For comparison, Ball (2001) reports a U.S. interest semielasticity $-\xi/100$ of approximately −0.05. The interpretation is that a permanent one percentage point increase in the interest rate decreases money demand by five percent and, under the homogeneity restriction that $\lambda_m = -1$, increases the price level by five percent.
[c] 10% critical values for PSS Case III are: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[d] 10% critical values are PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)
[e] Reported tests are based on 4 lags, but the test decisions are generally robust to variations on the lag length.
[f] Obtained by nonlinear least squares estimation of the ECM.

These findings for the annual versus quarterly data are therefore consistent with the a priori intuition that, for the purpose of studying long run relationships between variables, data span is more important than frequency. This tendency continues to be borne out in the ADL/ECM models in which the homogeneity restrictions are tested, to which we now turn.

4. Estimation results

Although the PSS tests of Tables 2 and 3 are useful as preliminary evidence on cointegration, models yielded by the mechanical application of the AIC and BIC may be unsatisfactory in other respects. One indication of this is that the LM autocorrelation tests show that some of those models—especially for the U.S. annual data—do not adequately capture the temporal dependence in the data.

This motivates more careful specification searches, typically beginning with the model yielded by the AIC and eliminating insignificant lags. Even so, many of the preliminary results of Tables 2 and 3 turn out to presage broader findings. First, for the U.S. annual data, two lags on the dependent variable were generally adequate, while the U.K. data needed four or five. Shorter lags on

Table 8
U.S. quarterly 1948–2019 ADL/ECM under log-log money demand: $\log k(r_t) = \log A - \eta \log r_t$.

	M2(3,0,2,0)	
ECM(·, ·, ·, ·) equilibrium error: $p_t + \lambda_m m_t + \lambda_y y_t - \lambda_m(\alpha - \eta \log r_t)$	no structural change	structural change[g]
Parameters of the error correction term:		
speed-of-adjustment coefficient[a], ψ	−0.0074	−0.0124
	(0.0028)	(0.0025)
interest elasticity[b], $-\eta$	−0.1243	−0.1322
	(0.0653)	(0.1020)
real transactions, λ_y	0.1244	−0.5121
	(0.4757)	(0.2630)
nominal money, $-\lambda_m$	0.6040	0.1811
	(0.2263)	(0.1382)
PSS tests of null of no levels relationship:		
F statistic[c]	3.6450	8.5857
t statistic[d]	−2.7437	−5.0366
LM test for residual autocorrelation[e] (p-value)	0.2746	0.0050
Wald tests of homogeneity restrictions (p-values):		
(1) Real money demand ($\theta_p = 1$): $\lambda_m = -1$	0.0813	0.0000
(2) Unitary income elasticity ($\theta_y = 1$): $\lambda_y = -\lambda_m$	0.0576	0.0000
(3) Joint test of (1) and (2): $-\lambda_m = \lambda_y = 1$	0.1338	0.0000
Estimate of λ_y when $\lambda_m = -1$ imposed	0.9451	1.0284
	(0.0621)	(0.1423)
p-value for $\lambda_y = 1$	0.3775	0.8419
Restricted equilibrium error ($p_t - m_t + y_t + \log A - \eta \log r_t$) imposes $-\lambda_m = \lambda_y = 1$:		
speed-of-adjustment coefficient[a], ψ	−0.0052	−0.0043
	(0.0025)	(0.0020)
interest elasticity[b], $-\eta$	−0.0994	−0.0849
	(0.0455)	(0.0369)
intercept parameter[f], A	0.2984	3.3222
	(0.0871)	(3.2115)
Naive money demand estimates from static simple regression of $\log k$ on $\log r$:		
interest elasticity[b], $-\eta$	−0.0313	
	(0.0036)	
intercept parameter[f], A	0.4972	
	(0.0071)	
Credible welfare cost estimates yielded by restricted ECM (percent of income): $100 \times w(r)$		
at $r = 0.03$	0.1401	
at $r = 0.05$	0.2219	
at $r = 0.06$	0.2615	
at $r = 0.13$	0.5246	
at $r = 0.14$	0.5608	

Notes: Standard errors are in parentheses.
[a] Convergence to equilibrium requires $-1 < \psi < 0$.
[b] For comparison, the Baumol-Tobin square-root rule for money holding predicts an interest elasticity of −0.5.
[c] 10% critical values for PSS Case III are: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[d] 10% critical values are PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)
[e] Reported tests are based on 4 lags, but the test decisions are generally robust to variations on the lag length.
[f] Obtained by nonlinear least squares estimation of the ECM.
[g] Indicator saturation finds impulse indicators in 1948:3, 1950:3, 1951:1, 1951:2, 1951:3, 1952:3; and step indicators in 1949:3, 1952:1, 1972:3.

the dependent variable in the U.K. models always left residual autocorrelation, regardless of the number of lags on the explanatory variables.[6] Second, the weak evidence of cointegration in Table 3 anticipates other problems with many of the quarterly models, such as incompatibility with the homogeneity restrictions, implausible and poorly-estimated money demand coefficients, and even incorrect dynamics (such as positive speed-of-adjustment coefficients in a few instances).

Consequently, we begin with the results from the annual data sets. Tables 4 and 5 present results based on the U.S. annual data, Tables 6 and 7 for the U.K. annual data. Tables 8 and 9 then turn to the quarterly data. All these models share the objective of finding parsimonious ADL/ECM's that treat residual autocorrelation, show evidence of cointegration where the coefficients of the equilibrium error are compatible with the homogeneity restrictions, and yield well-estimated money demand parameters that provide a credible basis for welfare cost calculations.

[6] Among other variations on the model specifications, exploratory work showed that time trends in the cointegrating error are invariably statistically insignificant.

Table 9
Quarterly ADL/ECM under semi-log money demand: $\log k(r_t) = \log B - \xi r_t$.

ECM(·, ·, ·, ·) equilibrium error: $p_t + \lambda_m m_t + \lambda_y y_t - \lambda_m(\beta - \xi r_t)$	U.S. 1948–2019 M2(3,0,2,0) no structural change	structural change[g]	U.K. 1955–2019 M1(5,4,1,0)
Parameters of the error correction term:			
speed-of-adjustment coefficient[a], ψ	−0.0086	−0.0048	−0.0112
	(0.0026)	(0.0030)	(0.0039)
interest semielasticity[b], $-\xi$	−10.4870	−26.4694	−59.0625
	(5.7664)	(24.4908)	(113.7967)
real transactions, λ_y	−0.0813	−0.5295	−3.0990
	(0.3986)	(0.5921)	(1.4585)
nominal money, $-\lambda_m$	0.4858	0.3690	0.1944
	(0.1896)	(0.2933)	(0.4019)
PSS tests of null of no levels relationship:			
F statistic[c]	6.0136	8.0969	11.4160
t statistic[d]	−3.2644	−1.5321	−2.8658
LM test for residual autocorrelation[e] (p-value)	0.5239	0.0092	0.8214
Wald tests of homogeneity restrictions (p-values):			
(1) Real money demand ($\theta_p = 1$): $\lambda_m = -1$	0.0071	0.0324	0.0033
(2) Unitary income elasticity ($\theta_y = 1$): $\lambda_y = -\lambda_m$	0.0075	0.0180	0.0066
(3) Joint test of (1) and (2): $-\lambda_m = \lambda_y = 1$	0.0263	0.0370	0.0050
Estimate of λ_y when $\lambda_m = -1$ imposed	1.0037	−0.1474	1.7438
	(0.0616)	(1.2090)	(0.6527)
p-value for $\lambda_y = 1$	0.9524	0.3435	0.2556
Restricted equilibrium error ($p_t - m_t + y_t + \log B - \xi r_t$) imposes $-\lambda_m = \lambda_y = 1$:			
speed-of-adjustment coefficient[a], ψ	−0.0055	−0.0028	−0.0016
	(0.0024)	(0.0017)	(0.0036)
interest semielasticity[b], $-\xi$	−7.1262	−12.9828	−52.1372
	(2.9053)	(7.2596)	(101.6827)
intercept parameter[f], B	0.5343	0.3471	0.2354
	(0.0465)	(0.3520)	(0.3226)
Naive money demand estimates from static simple regression of $\log k$ on r:			
interest elasticity[b], $-\xi$	−1.0820		−8.0020
	(0.1767)		(0.3020)
intercept parameter[f], B	0.5841		1.4760
	(0.0053)		(0.0325)
Credible welfare cost estimates yielded by restricted ECM (percent of income): $100 \times w(r)$			
at $r = 0.03$	0.1488	0.1571	
at $r = 0.05$	0.3767	0.3699	
at $r = 0.06$	0.5181	0.4911	
at $r = 0.13$	1.7784	1.3448	
at $r = 0.14$	1.9748	1.4502	

Notes: Standard errors are in parentheses.
[a] Convergence to equilibrium requires $-1 < \psi < 0$.
[b] For comparison, Ball (2001) reports a U.S. interest semielasticity $-\xi/100$ of approximately −0.05. The interpretation is that a permanent one percentage point increase in the interest rate decreases money demand by five percent and, under the homogeneity restriction that $\lambda_m = -1$, increases the price level by five percent.
[c] 10% critical values for PSS Case III are: 2.72 if all variables are I(0), 3.77 if all variables are I(1). (PSS Table CI(iii).)
[d] 10% critical values are PSS Case III: −2.57 if all variables are I(0), −3.46 if all variables are I(1). (PSS Table CII(iii).)
[e] Reported tests are based on 4 lags, but the test decisions are generally robust to variations on the lag length.
[f] Obtained by nonlinear least squares estimation of the ECM.
[g] Indicator saturation finds impulse indicators in 1948:3, 1950:3, 1951:1, 1951:2, 1951:4, 1952:3; and step indicators in 1948:4, 1949:3, 1950:4, 1951:2, 1952:1, 1952:4, 1972:3, 1981:2.

4.1. U.S. annual data

The results for the U.S. annual data come closest to satisfying these criteria. Tables 4 and 5 present results for ADL/ECM's based on log-log and semi-log money demand, respectively, for the three aggregates M1, NewM1, and M2.

The unrestricted ECM's show that the narrower aggregates M1 and NewM1 often yield poorly estimated values for key parameters, including estimates of the transactions coefficient λ_y of the wrong sign (although statistically insignificant). Even with the often large standard errors, the homogeneity restrictions are typically rejected at conventional significance levels. This is more surprising for NewM1 than it is for M1, given that LN constructed NewM1 for the specific purpose of addressing the apparent absence of stable M1 demand.

Unsurprisingly, then, imposing these rejected restrictions on the M1 and NewM1 log-log (Table 4) or semi-log (Table 5) models yields restricted ECM's that are not a basis for credible welfare costs, with money demand coefficients that are poorly estimated and, in the case of the NewM1 log-log model, a speed-of-adjustment coefficient and interest elasticity of the wrong signs. That the estimates of the M1 and NewM1 money demand parameters are not a basis for credible welfare cost calculations is emphasized by a comparison with the useful benchmark of naive static simple regression estimates of the money demand functions, shown toward the bottom of Tables 4 and 5.

These weak results for the narrower aggregates are not rectified by varying lag lengths or incorporating structural breaks. At the same time, the results shown do reveal a few constructive findings. First, in their comprehensive international study using long-span annual data, BLNW found typical M1 interest elasticities in the range of (negative) 0.3 to 0.6, a range that includes the Baumol-Tobin square root rule of $\eta = 0.5$. The Table 4 unrestricted ADL/ECM η estimates of 0.3382 (M1) and 0.5529 (NewM1) are in this range. Second, the PSS tests reveal some evidence of cointegration, especially for M1. Third, in both the log-log (Table 4) and semi-log (Table 5) models based on NewM1, the homogeneity restrictions $\lambda_m = -1$ and $\lambda_y = -\lambda_m$ are rejected individually and jointly. However, if money demand is required to be a demand for real money balances by imposing the first of these restrictions, NewM1 yields estimates of λ_y not significantly different from $\lambda_y = 1$: 1.5291 in the case of the log-log model, 0.8034 for the semi-log. So imposing the canonical money demand function (5) yields results consistent with a unitary NewM1 income elasticity, especially for the semi-log model. This is a marked difference with the model based on unmodified M1, the traditional aggregate of most U.S. money demand studies.

Turning to M2, the results are much stronger, as the univariate unit root tests of section 3.2 suggested might be the case. The coefficients of the unrestricted log-log (Table 4) and semi-log (Table 5) ECM's are well-estimated and correctly signed, with interest rate sensitivities weaker than those for M1 or NewM1, as is often the case for a broader aggregate. There is strong evidence of cointegration. The point estimates of the homogeneity coefficients λ_m and λ_y are roughly compatible with the homogeneity restrictions, although they are precise enough that the homogeneity restrictions tend to be formally rejected, especially $\lambda_y = 1$. As with NewM1, if we add to the model the requirement that $\lambda_m = -1$, then estimates of λ_y much closer to $\lambda_y = 1$ are obtained—remarkably so in the case of log-log money demand, where λ_y is estimated to be 0.9993 with a standard error of 0.0398. Fully imposing the homogeneity restrictions yields estimates of the money demand parameters that are broadly comparable with the benchmarks of the naive static simple regressions.

For the purpose of estimating welfare costs, the major deficiency of these baseline M2 models is the formal rejection of the homogeneity restrictions. Unlike the M1 and NewM1 models, this changes dramatically if structural change is treated, as shown in the final column of Tables 4 and 5. Although the exact structural change dates identified by indicator saturation vary slightly with the model, they make intuitive sense, being associated with the Spanish flu epidemic (1920 and 1921), the Great Depression of the 1930s, and (in the semi-log model of Table 5) World War II. The main substantive difference between the resulting log-log and semi-log M2 estimations is that the log-log model yields more precise estimates of the money demand parameters, which are also more stable in the face of the treatment of structural change. In turn, the log-log model yields highly plausible welfare costs. By contrast, the only semi-log model that yields even remotely credible welfare cost estimates is that for M2 without structural change, but even these are at the high end of plausible numbers. In this model, treating structural change alters the estimate of the intercept parameter B from 0.7544 to 2.1005, the latter with the fairly large standard error of 1.3725, which in turn yields implausibly large welfare costs, and so are not shown at the bottom of Table 5. (For example, for $r = 0.14$ the log-lin M2 model with structural change yields an obviously preposterous welfare cost of 7.255 percent of GDP.)

In conclusion, the model that emerges from the U.S. annual data that offers the most credible basis for welfare cost calculations is the log-log M2 model with structural change treated. Although, in that model, the treatment of structural change is important for obtaining coefficient estimates consistent with homogeneity, the welfare cost estimates are little affected.

4.2. U.K. annual data

Results for annual ADL/ECM's based on the three U.K. monetary aggregates M1, Mbroad, and M4 are shown in Tables 6 and 7 for log-log and semi-log money demand, paralleling the construction of the annual U.S. tables.

Like the U.S., there is much in the unrestricted models that makes sense: speed-of-adjustment coefficients are small negative fractions, interest rate effects are negative and well-estimated (with the exception of log-log M1), and there is much evidence of cointegration, especially for Mbroad and M4. But, like the U.S., in the absence of giving the model the a priori information that money demand is of the canonical form (5), so that the restriction $\lambda_m = -1$ is imposed, the data have difficulty estimating λ_m and λ_y separately: the transactions elasticity λ_y, which should be positive, is even of the wrong sign, as well as being poorly estimated. In turn, the homogeneity restrictions are typically rejected.

This changes if $\lambda_m = -1$ is imposed. For both functional forms and most aggregates (the exception being log-log M1, where there is an absence of cointegration), this restricted form of the ECM dramatically improves the estimation of λ_y, yielding estimates in the range of 0.9651–1.4410, with small standard errors that often do not reject $\lambda_y = 1$. To this extent, requiring money demand to be a demand for *real* money balances yields models that are at least qualitatively consistent with a unitary transactions elasticity.

What happens in these U.K. annual models when the homogeneity restrictions are fully imposed? More so than for the U.S., the models yielding credible welfare costs vary with the functional form and aggregate. In the log-log model, Mbroad works best: with structural change permitted, it yields well-estimated money demand parameters. In a curious contrast with the U.S., indicator saturation finds that the years of structural change in the U.K. were not those of the Great Depression or World War II, but the economic tumult of the 1970s.

The alternatives of the M1 and M4 models in Table 6 yield poorly estimated money demand parameters in the restricted equilibrium error, a deficiency that was not rectified by varying lag lengths or permitting structural change. And, as already noted, the log-log M1 model shows little evidence of cointegration, consistent with the models yielded by the AIC and SIC in Table 2.

The outcome of imposing the homogeneity restrictions is quite different in the semi-log model (Table 7), where it is M1 that yields well-estimated money demand parameters. (Again, it turns out that there is no benefit to introducing structural change; in fact, the precision of the money demand estimates deteriorates.) The semi-log M1 model also yields better evidence of cointegration than the log-log M1 model of Table 6, consistent with the models yielded by the AIC and SIC in Table 2.

In conclusion, for the U.K. annual data the most plausible welfare costs come from the log-log Mbroad model with structural change, and the semi-log M1 model. At lower interest rates these two models yield similar welfare costs—less than 0.2 percent of GDP at $r = 0.03$. To the extent that the welfare costs differ at higher interest rates, it is unclear to what extent the difference is due to the different functional forms or the different aggregates. In terms of an apples-with-apples comparison with the U.S., the natural comparators are the U.S. log-log M2 model of Table 4 and the U.K. log-log Mbroad model of Table 6, both with structural change treated. These welfare costs are remarkable in their consistency, ranging from 0.2 percent of GDP at an interest rate of 3 percent to, at an interest rate of 14 percent, 0.9 percent for the U.S. and 0.75 percent for the U.K.

4.3. Quarterly models

Recall from the preliminary cointegration tests of Table 3 that evidence of cointegration is weaker in the quarterly data than in the annual, especially for the U.K., possibly because of its shorter span. For the most part, this remains the case for a broader range of model specifications beyond those yielded mechanically by the AIC and BIC. As with the annual data, weak evidence of cointegration tends to presage other deficiencies in the estimated models, such as poorly estimated money demand parameters. Consequently the quarterly data yield fewer estimated models that offer useful insights; they are presented in Tables 8 and 9 for log-log and semi-log money demand.

U.S. quarterly models

For the U.S., as with the annual data, only M2 yielded models that provide a basis for credible welfare costs. For both log-log (Table 8) and semi-log (Table 9) money demands, independent specification searches led to M2(3,0,2,0) lag lengths. Reflecting the quarterly frequency, speed-of-adjustment coefficients are much smaller negative fractions than in the annual data. In the unrestricted ECM's, interest rate effects are correctly signed but not particularly well estimated. Depending on the model and whether structural change is treated, there is some evidence of cointegration, the strongest being in the baseline semi-log model.

As with the annual data, in the absence of introducing a priori information the homogeneity coefficients λ_y and λ_m are poorly estimated and, in the case of the transactions elasticity, sometimes even of the wrong sign. This improves if $\lambda_m = -1$ is imposed: both the baseline log-log and semi-log models yield estimates of λ_y remarkably close to unity. Wald tests show the baseline log-log model to be most consistent with the homogeneity restrictions.

The fully-restricted baseline models yield money demand parameter estimates that are ballpark-comparable to those from their naive simple regression counterparts. In fact, the baseline semi-log model of Table 9 yields an interest semielasticity of $-\xi = -7.1262$—incidentally, astonishingly close to Lucas's (2000) calibrated value of -7. In turn, the baseline models yield plausible welfare costs, although different in a way that parallels the results from the annual data: welfare costs from the two models are similar at the lowest interest rates (at $r = 0.03$, 0.1401 for the log-log model of Table 8, and 0.1488 for the semi-log model of Table 9) but depart at higher rates, the semi-log model yielding larger welfare costs.

Unlike the annual data, allowing for structural change in these U.S. quarterly M2(3,0,2,0) models does little to improve these results. In fact, in the log-log model of Table 8, the intercept parameter A of the money demand function is 3.3222, much larger than the naive benchmark of 0.4972, and therefore yields implausible welfare costs (such as 5.1 percent of GDP at $r = 0.14$), and so are not shown. The results for the semi-log model in Table 9 are better in this respect: the welfare costs are fairly robust to treating structural change, especially at lower interest rates.

These results for the U.S. quarterly models are not improved by other variations on their specification, such as varying lag lengths or using the narrower aggregates M1 or NewM1.

U.K. quarterly models

Recall from Table 3 that the U.K. quarterly data showed little evidence of cointegration for any of the three aggregates. The sole (although still weak) exception was the semi-log M1(5,4,1,0) model selected by the AIC.

More extensive specification searches across the two functional forms and three aggregates, including permitting structural change, confirmed these findings more generally, as well as the other weak estimation results that tend to accompany an absence of cointegration. The final column of Table 9 illustrates this with the semi-log M1(5,4,1,0) model yielded by the AIC, about the only redeeming feature of which is that the speed-of-adjustment coefficient is sensible. Most notably, the interest semielasticity is poorly estimated regardless of whether the homogeneity restrictions are imposed, and this does not change with the usual model variations such as different lag lengths or permitting structural change.

4.4. Comparative assessment

Focusing on the estimation results that yield credible welfare costs, how do they compare across the various models, monetary aggregates, and data sets, and with previous work? Table 10 collects the credible welfare costs obtained from the annual data estimations of Tables 4–7 and compares them with Lucas (2000), who studied U.S. annual M1, 1900–1994. Table 11 collects the

Table 10
Summary of credible welfare costs: annual data (percent of income).

Illustrative interest rates (percent)	U.S. log-log M2 no structural change	U.S. log-log M2 structural change	U.S. semi-log M2 no structural change	U.K. log-log Mbroad structural change	U.K. semi-log M1 no structural change	Comparator study: Lucas (2000) U.S. M1 1900–1994[a] log-log	semi-log
3	0.2205	0.2274	0.1658	0.1908	0.1460	0.845[b]	0.097[b]
5	0.3482	0.3609	0.4290	0.3007	0.3493	1.091[b]	0.247[b]
6	0.4099	0.4256	0.5963	0.3537	0.4671	1.195[c]	0.340
13	0.8188	0.8565	2.1960	0.7042	1.3396	1.760	1.172
14	0.8750	0.9158	2.4618	0.7522	1.4531	1.826	1.302
interest parameter:	$-\eta = -0.1051$	$-\eta = -0.0955$	$-\xi = -5.4431$	$-\eta = -0.1093$	$-\xi = -11.7179$	$-\eta = -0.5$	$-\xi = -7$
intercept parameter:	$A = 0.4327$	$A = 0.5136$	$B = 0.7544$	$A = 0.3532$	$B = 0.3489$	$A = 0.0488$	$B = 0.3548$

[a] The intercept values $A = 0.0488$ and $B = 0.3548$ do not appear in Lucas (2000), but are given in Ireland (2009, p. 1040).
[b] Using Lucas's parameter values, Ireland (2009, p. 1041) remarks that if "... the steady-state real interest rate equals 3 percent, so that $r = 0.03$ prevails under a policy of zero inflation or price stability, then ... this policy costs the economy the equivalent of 0.85 percent of income when money demand is log-log, but only 0.10 percent of income when money demand has the semi-log form. Likewise, an ongoing 2 percent inflation [so $r = 0.05$] costs the economy 1.09 percent of income under [a loglinear form], but only 0.25 percent of income under [a semi-log form]."
[c] "At a six percent interest rate, for example, the log-log curve implies a welfare cost of about one percent of income ... " (Lucas, 2000, p. 251).

Table 11
Summary of credible welfare costs: quarterly data (percent of income).

Illustrative interest rates (percent)	U.S. log-log M2 no structural change	U.S. semi-log M2 no structural change	U.S. semi-log M2 structural change	Comparator study: Ireland (2009) U.S. semi-log M1 1980:I–2006:IV
3	0.1401	0.1488	0.1571	0.0131[a]
5	0.2219	0.3767	0.3699	0.0356[a]
6	0.2615	0.5181	0.4911	0.0510
13	0.5246	1.7784	1.3448	0.2192[a]
14	0.5608	1.9748	1.4502	0.2510
interest parameter:	$-\eta = -0.0994$	$-\xi = -7.1262$	$-\xi = -12.9828$	$-\xi = -1.7944$
intercept parameter:	$A = 0.2984$	$B = 0.5343$	$B = 0.3471$	$B = 0.1686$

[a] From Table 6 of Ireland (2009, p. 1050) based on static OLS estimation of semi-log money demand. Ireland also presents welfare costs based on dynamic OLS, but the values are similar.

credible quarterly results from Tables 8 and 9 and compares them with Ireland (2009), who studied U.S. quarterly M1, 1980–2006, using the semi-log model.

Despite their very different data sets, both Lucas and Ireland were led to emphasize (Ireland, p. 1040) the "... very different money demand behavior at low interest rates ... " yielded by the log-log and semi-log models. Both authors found that their semi-log models yielded very low welfare costs at low interest rates (as in the final columns of Tables 10 and 11), in comparison with higher welfare costs yielded by a log-log specification (as in the penultimate column of Table 10)). "Hence ... these competing money demand specifications also have very different implications for the welfare cost of modest departures from Friedman's ... zero nominal interest rate rule for the optimum quantity of money. ... These calculations underscore the importance of discerning the appropriate form of the money demand function before evaluating alternative monetary policies, including those that generate very low but positive rates of inflation." (Ireland, pp. 1040–41.)

My results in Tables 10 and 11 suggest the opposite: there is a remarkable order-of-magnitude consistency in the welfare costs at lower interest rates, not only across models but also across countries, monetary aggregates, data frequencies, and whether structural change is treated. At an interest rate of 3 percent, my welfare costs fall in the fairly narrow range of 0.1401–0.2274 percent of income (which come from quarterly and annual data, respectively; both are from the U.S. log-log M2 model with no structural change). This range also happens to be remarkably consistent with Benati et al. (2017, p. 26) who, for U.S. NewM1, 1915–2016, found "... welfare losses at the average short rate that has prevailed over the sample period ... " of 0.18 percent for the log-log and semi-log models.

Instead the more significant departure in my welfare costs occurs at higher interest rates, and the source of this departure is clear: it comes not from the choice of country or data frequency, but from the functional form. The semi-log model yields higher welfare costs than the log-log model, the opposite of Lucas's findings (in the final two columns of Table 10). Having said this, the magnitudes of my welfare costs are far more consistent with Lucas's numbers than Ireland's, which is unsurprising given that Ireland's quarterly data were over the fairly short span of 1980–2006, modest even in comparison with the span of my quarterly data, let alone the annual data.

5. Conclusions

Many of the findings of this analysis are not difficult to anticipate: the quantity theory emerges most clearly from longer-span data and broader monetary aggregates.

The strongest case for this is U.S. annual M2 over the past century, where both log-log and semi-log money demand specifications yield well-estimated, statistically significant parameter estimates with sensible signs and magnitudes. There is strong evidence of cointegration, most likely between prices, money, and income, with the interest rate probably best thought of as stationary. Especially with log-log money demand and structural change treated, the coefficients of the equilibrium vector are consistent with the homogeneity propositions: money demand is a demand for real money balances with a unit transactions elasticity. In this model, the welfare cost of inflation is estimated to range between 0.2274 and 0.9158 percent of income at interest rates of 3 and 14 percent, respectively—values that are fairly insensitive to the treatment of structural change.

A significant finding, less easily anticipated, is the widespread rejection of the homogeneity propositions for other aggregates and shorter data spans. Even the U.K. annual data often reject these restrictions, with corresponding implications for the estimates of money demand parameters when the restrictions are imposed. Even so, the most plausible of the U.K. welfare costs are remarkably comparable to those for the U.S., being less than one percent of GDP at interest rates experienced during the past century. If one accepts these figures, then one overarching implication is that the deadweight loss of seigniorage is probably modest in comparison with other taxes.[7]

Disclosure statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

I am indebted to the Associate Editor Luca Benati and two anonymous referees for extensive suggestions that have significantly improved the paper, and to Ryland Thomas of the Bank of England for graciously providing me with series updates through 2019 to the Bank's *A Millenium of Macroeconomic Data* spreadsheet. A preliminary version of this paper was presented at the 2022 annual meetings of the Canadian Economics Association in Ottawa and in a department seminar at the University of Victoria. I thank participants for their comments, especially Merwan Engineer and Vasco Gabriel.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jedc.2024.104842.

References

Baba, Y., Hendry, D.F., Starr, R.M., 1992. The demand for M1 in the USA: 1960–1988. Rev. Econ. Stud. 59, 25–60.
Ball, L., 2001. Another look at long-run money demand. J. Monet. Econ. 47, 31–44.
Barnett, W.A., 1980. Economic monetary aggregates: an application of index number and aggregation theory. J. Econom. 14, 11–48.
Benati, L., 2020. Money velocity and the natural rate of interest. J. Monet. Econ. 116, 117–134.
Benati, L., Lucas Jr., R.E., Nicolini, J.-P., Weber, W., 2017. Long-run money demand redux. Working Paper 737. Federal Reserve Bank of Minneapolis.
Benati, L., Lucas Jr., R.E., Nicolini, J.-P., Weber, W., 2021. International evidence on long-run money demand. J. Monet. Econ. 117, 43–63.
Carlson, J.B., Hoffman, D.L., Keen, B.D., Rasche, R.H., 2000. Results of a study of the stability of cointegrating relations comprised of broad monetary aggregates. J. Monet. Econ. 46, 345–383.
Castle, J.L., Doornik, J.A., Hendry, D.F., Pretis, F., 2015. Detecting location shifts during model selection by step-indicator saturation. Econometrics 3, 240–264.
Cochrane, J.H., 1991. Comment. In: NBER Macroeconomics Annual 1991, pp. 201–210.
Elliot, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813–836. 402–415.
Hendrickson, J.R., 2014. Redundancy or mismeasurement: a reappraisal of money. Macroecon. Dyn. 18, 1437–1465.
Hendry, D.F., Ericsson, N.R., 1991. Modeling the demand for narrow money in the United Kingdom and the United States. Eur. Econ. Rev. 35, 833–886.
Hoffman, D.L., Rasche, R.H., Tieslau, M.A., 1995. The stability of long-run money demand in five industrial countries. J. Monet. Econ. 35, 317–339.
Ireland, P.N., 2009. On the welfare cost of inflation and the recent behavior of money demand. Am. Econ. Rev. 99, 1040–1052.
Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, Oxford.
Jorda, Ò., Schularick, M., Taylor, A., 2017. Macrofinancial history and the new business cycle facts. NBER Macroecon. Annu. 31, 213–263.
Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. J. Econom. 54, 159–178.
Lucas, R.E., 2000. Inflation and welfare. Econometrica 68, 247–274.
Lucas, R.E., Nicolini, J.-P., 2015. On the stability of money demand. J. Monet. Econ. 73, 48–65.
McCallum, B.T., Nelson, E., 2011. Money and inflation: some critical issues. In: Handbook of Monetary Economics vol. 3A. Elsevier. Chap. 3.
Pesaran, M.H., Shin, Y., Smith, R.J., 2000. Structural analysis of vector error correction models with exogenous I(1) variables. J. Econom. 97, 293–343.
Pesaran, M.H., Shin, Y., Smith, R.J., 2001. Bounds testing approaches to the analysis of level relationships. J. Appl. Econom. 16, 289–326.
Rogoff, K., Rossi, B., Schmelzing, P., 2022. Long-run trends in long-maturity real rates 1311–2021. NBER Working Paper 30475.
Serletis, A., Yavari, K., 2004. The welfare cost of inflation in Canada and the United States. Econ. Lett. 84, 199–204.
Stock, J., Watson, M., 1993. A simple estimator of cointegrating vectors in higher order integrated systems. Econometrica 61, 783–820.

[7] I thank Merwan Engineer for emphasizing to me this interpretation of the substantive implication of these results.