Volume Title: Monetary Trends in the United States and United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975

Volume Author/Editor: Milton Friedman and Anna J. Schwartz

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-26409-2

Volume URL: http://www.nber.org/books/frie82-2

Publication Date: 1982

Chapter Title: Velocity and the Demand for Money

Chapter Author: Milton Friedman, Anna J. Schwartz

Chapter URL: http://www.nber.org/chapters/c11406

Chapter pages in book: (p. 205 - 304)

6 Velocity and the Demand For Money

The most intriguing finding in the preceding chapter is the extraordinary parallelism of velocity in the United States and the United Kingdom, especially since 1905 (see chart 5.5). We have attributed the deviations before 1903 primarily to the growing financial sophistication of the United States, but they may also reflect the greater inaccuracy of the earlier data for both the United States and the United Kingdom.

The parallelism presumably reflects both similar money-holding propensities in the United States and the United Kingdom for more than two-thirds of a century and a largely common set of factors determining the number of weeks of income that residents of the United States and the United Kingdom chose to hold as money. The alternative is to attribute the parallelism to an accidental offsetting of differences in propensities by differences in the factors affecting the amount of money held. The parallelism therefore suggests that we can use the data for the two countries as if they came from a single parent universe in trying to identify the factors determining the behavior of velocity. That is the task of the present chapter. The parallelism also suggests that it is worthwhile exploring the influences connecting the two countries. That is the task of chapter 7.

An analysis of the behavior of velocity is an analysis of the demand for money. As we noted in chapter 2, we can express the "real" quantity of money in various ways: most directly, as a quantity of money divided by a price index (as in equation 7 in chap. 2) or, as is frequently more meaningful in comparing different countries or widely separated periods, by the device used in chapter 5, as the number of weeks of income or consumption to which the quantity of money is equivalent. The income velocity of circulation is simply the reciprocal of the number of weeks of income held as money, which is why an analysis of velocity is equivalent to an analysis of the demand for money. Velocity is usually expressed per year rather than per week and hence, as usually expressed, is equal to the reciprocal of the number of weeks of income held as money times fifty-two. But that is simply a question of units. The percentage rate of change of weeks-of-income held as money is equal in numerical value but opposite in sign to the rate of change of velocity.

If velocity is viewed in terms of the demand for money, we must, as in any demand study, distinguish "desired" or "demanded" quantities from actual quantities and actual quantities from measured quantities. Desired and actual quantities may differ because the demand function in question involves a different level or stage of adjustment—a different Marshallian period—than corresponds to the observed quantity; because, that is, we admit that demanders may be "off" their demand curve of the kind in question. And actual quantities may differ from measured quantities because of errors of measurement.

The simplest explanation of measured velocity and its movements is that velocity is the ratio of two independent magnitudes, each determined by a separate set of forces. This explanation is consistent with the view that (1) there does not exist a stable demand for money as a function of a small number of variables; or (2) there exists a stable function but it has a special form, for example, Keynesian absolute liquidity preference, so that velocity adapts passively to the separate movements in income and money; or (3) the errors of measurement of numerator and denominator dominate the fluctuations of velocity. Section 6.1 demonstrates that our data are inconsistent with this explanation.

At the opposite extreme doctrinally is the explanation that desired velocity is a numerical constant—a simple-minded quantity theory rather than Keynesian theory. On this explanation, observed deviations from constancy reflect either errors of observation or differences between actual and desired velocity. This extreme quantity theory must be rejected. Yet it is impressive how far it carries us in explaining (1) the movements of aggregate income, if the quantity of money is regarded as exogenous; or (2) the movements of aggregate quantity of money, if income is regarded as exogenous and the quantity of money as demand-determined (i.e., an adaptive supply) (sec. 6.2).

To go beyond these simple theories, we must investigate the effect of other variables on the quantity of money demanded. In view of the findings of the previous chapter, one variable that we must allow for in some way or another is the changing financial sophistication in the United States—which we regard as the chief explanation for the difference between the United States and the United Kingdom in the level and trend of velocity before 1902 (sec. 6.3). Beyond this, we consider the possible influence of real per capita income (sec. 6.4), of population and prices (sec. 6.5), and of the cost of holding money (sec. 6.6). Finally in section 6.7, we consider the joint influence of these variables.

6.1 Velocity: A Will-o'-the Wisp?

As we indicated in chapter 2, much recent literature deriving its inspiration, though not always its details, from Keynes's *General Theory* treats nominal income as if it were determined by forces largely independent of the quantity of money, so that velocity adjusts passively.[1] Carried to its limits, this interpretation makes velocity—and its reciprocal—the ratio of two statistically independent magnitudes.[2] In that extreme case, the correlation between income and money would be zero and the variance of the logarithm of weeks of income held as money would be equal to the sum of the variances of the logarithms of income and money, and hence larger than either.

This interpretation is clearly not valid for *levels* of income, money, and weeks of income. The simple correlation between the logarithms of income and money is .992 for the United States and .991 for the United Kingdom for the period as a whole, and its lowest value for separate periods before and after 1914 is .980. As a result, the variance or standard deviation of the logarithm of weeks of income is much less than the variance or standard deviation of the logarithms of either money or income.

However, few if any proponents of this interpretation would apply it to the levels of income, money, and velocity for a long period. They grant that in the long run the desired ratio of money to income is not indefinitely malleable, that it would be impossible to multiply nominal income manyfold without something like a corresponding rise in the quantity of money. They view velocity as malleable over shorter periods—perhaps over periods even shorter than the phases that are our units of observation.

[1] See J. G. Gurley and E. S. Shaw, *Money in a Theory of Finance*, Washington, D. C.: Brookings Institution, 1959; (Radcliffe) Committee on the Working of the Monetary System, *Report*, Cmd. 827, London: HMSO, 1959; R. S. Sayers, "Monetary Thought and Monetary Policy in England," *Economic Journal* 70 (December 1960): 710–24.

[2] For example, Alvin Hansen, the leading American disciple of Keynes, wrote in 1957 (*The American Economy*, p. 50): "I think we should do well to eliminate once and for all, the phrase 'velocity of circulation' from our vocabulary. Instead, we should simply speak of the ratio of money to aggregate spending. The phrase 'velocity of circulation' is, I feel, unfortunate because those who employ it tend to make an independent entity out of it and imbue it with a soul. This little manikin is placed on the stage, and the audience is led to believe that it is endowed with the power of making decisions directing and controlling the flow of aggregate spending. In fact it is nothing of the sort. It is a mere residual. We should get on much better if we substitute the word 'ratio.' The little manikin would then be forced back into oblivion, where it properly belongs."

A more relevant application is therefore to the rates of change of money, income, and weeks of income. It could be that there is a fairly well determined relation between the trends of money and income but that deviations from the trends are largely attributable to forces affecting money and income separately, so that deviations of weeks of income from its trend are a largely passive consequence.[3] However, this application too is contradicted by the standard deviations and correlation coefficients in table 5.6. For the period as a whole, including and excluding war years, for two subperiods, and for both the United States and the United Kingdom, the correlations between the rates of change of income and money range from .73 to .95, and the rate of change of weeks of income fluctuates less than the rates of change of both nominal income and money.[4]

Taken as a whole, the evidence is decisive that movements in the rate of change of money are accompanied by sufficiently closely correlated movements in the same direction in the rate of change of income to make the rate of change of weeks of income or of velocity decidedly more stable than either of its components. Velocity is not simply or even mainly a will-o'-the-wisp, over either periods measured in decades or periods measured in phases.

6.2 Velocity: A Numerical Constant?

Velocity, as measured, is clearly not a numerical constant. However, measured velocity differs from "true," "permanent," or "desired" velocity for two reasons: errors of measurement, and deviations between actual and "desired" velocity. May these deviations not explain the failure of measured velocity to be a numerical constant?

In terms of the demand for money, this interpretation taken literally says that the demand function has a special form:

$$M = kY, \tag{1}$$

where M is the nominal quantity of money, Y is nominal income, both measured accurately, and k is a numerical constant equal to the recipro-

---

[3] The world "largely" is required in this statement because there must be some mechanism to bring the various series back to the trend lines, hence the deviations from trend cannot be wholly random. However, for short periods, the independent transitory elements in money and income could dominate the related systematic elements.

[4] A positive correlation does not alone guarantee this result:

$$g_V = g_Y - g_M$$

$$\sigma_{g_V} = \sqrt{\alpha^2_{g_Y} + \sigma^2_{g_M} - 2r_{g_Y g_M}\,\sigma_{g_Y}\,\sigma_{g_M}}$$

If $\sigma_{g_Y} = \sigma_{g_M}$, then $r_{g_Y g_M}$ must exceed 0.5 for σ_{g_V} to be less than both σ_{g_Y} and σ_{g_M}. If $\sigma_{g_Y} \neq \sigma_{g_M}$, r must exceed half the ratio of the larger to the smaller to assure that σ_{g_V} is less than both σ_{g_M} and σ_{g_Y}.

cal of velocity. It says that if M can be regarded as exogenous, and money-holders are always on the demand curve defined by equation (1), then nominal income at time T will be $\frac{1}{k}\ M(T)$. Alternatively, if nominal income is regarded as exogenous, and the quantity of money as passively adapting to the quantity demanded, then the nominal quantity of money at time T will be $kY(T)$. Note that equation (1), as the preceding sentence indicates, is itself completely neutral about the much discussed question of "direction of influence" or "causal significance." It simply says that if equation (1) is the demand function for money and if money-holders are on their demand function and if M and Y are measured accurately, then M and Y will move together in fixed proportion. The observation that M and Y do move together would leave entirely open the question which is the "cause" and which the "effect," or whether both are the common "effect" of still other variables.

The significant time trends in velocity for long periods for the United States and the United Kingdom (e.g., for the period before World War I, the United States regression has a slope of $-.024$ with a standard error of .0014; the United Kingdom regression has a slope of $+.0018$ with a standard error of .0006) is alone sufficient to rule out this simple version. Presumably, trends average out both statistical errors and deviations between actual and desired balances.

A more sophisticated version, like the more sophisticated will-o'-the-wisp explanation, allows for time trends, regarding the demand equation as

$$M\ (T) = k\ (T)\ Y\ (T), \tag{2}$$

where T is time and $k(T)$ is some simple function of time (such as a semilog trend), since if $k(T)$ were left completely free, equation (2) could be regarded as an identity defining $k(T)$.[5]

This version too can be rejected out of hand. A sufficient basis for rejecting it is the close parallelism of the deviations of velocity about trends in the United States and the United Kingdom for over two-thirds

---

[5] J. P. Gould and C. R. Nelson, "The Stochastic Structure of the Velocity of Money," *American Economic Review* 65 (June 1974): 405–18, assert that there is no significant trend in our velocity series for the United States from 1869 to 1960—that, on the contrary, it can be regarded as a random walk without drift. On their interpretation, the data are consistent with equation (1) plus stochastic disturbances. Houston H. Stokes and Hugh Neuberger, "A Note on the Stochastic Structure of the Velocity of Money: Some Reservations," *American Economist* 23 (fall 1979): 62–64, demonstrate that the Gould-Nelson result is produced by combining nonhomogeneous periods and that for the period 1879–1940 the trend is significant, and the series is not a random walk with drift, that is, neither equation (1) nor equation (2) plus stochastic disturbances is acceptable. The next paragraph of our text, plus the rest of this chapter, provides additional and in our view decisive evidence contradicting the Gould-Nelson interpretation.

of a century, and also, of the rates of change of velocity, which can be regarded as incorporating a sensitive adjustment for trend. Since the statistical bases of the estimates are completely independent for the two countries, and in each country, for money and income, errors of measurement cannot account for the parallelism. Similarly, the obvious link between the two countries—the balance of payments—seems more likely to have produced inverse than parallel deviations in the two countries between measured and desired velocity. A surplus for one country tended to be accompanied by a deficit for the other country, so whatever the effect of the balance of payments on velocity, it would be in opposite directions in the two countries.[6] We know of no links that work in the other direction.

Though we have treated the constant-velocity explanation and the will-o'-the-wisp explanation as if they were opposite extremes, it is worth noting that they have much in common. Insofar as variations in velocity reflect errors of measurement—which they undoubtedly do to some extent—such variations are consistent with both explanations. Moreover, as in all analysis involving a distinction between actual and desired, or between permanent and transitory, it is generally impossible to separate the part of the transitory component of the measured variable that is attributable to errors of measurement from the part that is attributable to other forces. Hence, at least some part of the variability of velocity attributable to deviations between measured and desired velocity is also consistent with both explanations.

Though a numerically constant velocity must be rejected as a full explanation of the relation between money and income, it should not be dismissed without recording how far it takes us. For any lengthy period, equation (1)—the simplest and most rigid form of the constant-velocity view—accounts for the great bulk of the variation in nominal income (if the nominal quantity of money is regarded as exogenous and income as adapting to the quantity of money) or in nominal quantity of money (if nominal income is regarded as exogenous and the quantity of money as adapting to income).

This is demonstrated in table 6.1. For the century as a whole, or the more than eighty years excluding phases we have designated as war

---

[6] Under fixed exchange rates, a surplus that arises from forces independent of the nominal income of the country with the surplus would tend to lower measured velocity below desired velocity in the interval between its effect on the quantity of money and the effect of the changed quantity of money on income. If the surplus arose from an autonomous rise in income that raised measured velocity above desired velocity (the sequence generally envisioned in the monetary theory of the balance of payments), it would reduce measured velocity to bring it back to desired velocity. In either case the surplus would be accompanied by a decline in velocity, though in the first there is no reason to expect it to be preceded by a rise in measured velocity; in the second there is.

phases, for the United States or the United Kingdom separately, or for the two combined, equation (1), with V set simply equal to its (geometric) average value over the period and for the country or countries in question, accounts for at least 94.5 percent of the variability in money or, alternatively, income.[7] For three separate peacetime periods the corresponding percentage exceeds 90 in nine out of eighteen observations, and 50 in fourteen out of eighteen observations.[8] The remaining four, one of which is negative and the other three between 30 and 40, are all for the briefest (eighteen years) and most turbulent subperiod—that between the wars. The most striking result is for the two countries together, for which the money and income observations have been pooled by the crude device of simply converting figures in pounds sterling to dollars at the ruling rate of exchange. Here, more than 95 percent of the variability is accounted for by equation (1) except for the pre–World War I period, where the sharp difference between the trends of velocity in the United States and the United Kingdom reduces the percentage accounted for to 61 for income and 66 for money.

The major upward trends in income and money in both countries play an important role in producing these results. The absence of wide differ-

---

[7] Taking logs of equation (1), we have

(a) $\log M = \log k + \log Y.$

If this equation were satisfied precisely, the variables could stand equally for observed or predicted values. In fact, it will not be satisfied precisely, so we must distinguish between observed and predicted values of M, k, and Y. Let the variables as written in equation (a) stand for observed values, $\hat{M}$ and $\hat{Y}$ for predicted values of M and Y, and $\overline{\log k}$ for the mean value of the observed values of $\log k$ (so its antilog is the geometric mean of the observed values of k, and the reciprocal of its antilog, the geometric mean of the observed values of velocity, or V). Then assuming Y exogenous,

(b) $\log \hat{M} = \overline{\log k} + \log Y,$

and the error of prediction is

(c) $\log M - \log \hat{M} = \log M - \log Y - \overline{\log k}$

(d) $= \log k - \overline{\log k}.$

Hence the root-mean-square error in the logs is equal to the standard deviation of $\log k$, and, since $\log k = -\log V$, is also the standard deviation of $\log V$.

If M is assumed exogenous, we have

(e) $\log \hat{Y} = \log M - \overline{\log k},$

(f) $\log Y - \log \hat{Y} = \log Y - \log M + \overline{\log k}$

(g) $= \log V - \overline{\log V},$

so the root-mean-square error is again the standard deviation of $\log V$. As noted in footnote *b* of table 6.1, these estimates are maxima because neither $\log Y$ nor $\log M$ is measured without error. See also footnote 28 of chapter 5.

[8] For reasons which will become apparent later in this chapter, we shift our periodization of nonwar phases for the rest of the book from the pre–World War I/ post–World War I dichotomy of chapter 5 to the trichotomy pre–World War I/ interwar/ post–World War II.

Table 6.1 Variability in Money and Income Accounted for by Constant-Velocity Demand Curves

			Variation (Percent or Percentage Points)[a]								
			Total			Maximum[b] Residual after Allowing for Constant Velocity[c]			Minimum Fraction of Total Variance[d] Accounted for by Constant Velocity		
	Phase Numbers										
Period	U.S.	U.K.	U.S. (1)	U.K. (2)	U.S. and U.K. (3)	U.S. (4)	U.K. (5)	U.S. and U.K. (6)	U.S. (7)	U.K. (8)	U.S. and U.K. (9)
			Level of Money: Income Assumed Exogenous								
Full period	4–52	1–37	172.6	117.1	148.6	34.0	16.7	29.2	96.1	98.0	96.1
Full period ex-wars			182.8	125.1	156.9	34.4	14.8	29.2	96.5	98.6	96.5
Pre–World War I	4–24	1–11	71.7	23.5	56.6	29.6	3.1	33.0	83.0	98.3	66.0
Interwar	28–36	15–23	11.4	7.7	61.0	13.3	6.0	12.9	−36.1	39.3	95.5
Post–World War II	40–52	27–37	44.4	40.6	109.3	8.4	16.0	14.8	96.4	84.5	98.2
			Level of Income: Money Assumed Exogenous								
Full period	4–52	1–37	145.6	120.4	137.9	34.0	16.7	29.2	94.6	98.1	95.5
Full period ex-wars			155.3	131.4	146.4	34.4	14.8	29.2	95.1	98.7	96.0
Pre–World War I	4–24	1–11	44.3	25.7	53.1	29.6	3.1	33.0	55.4	98.6	61.4
Interwar	28–36	15–23	16.4	7.2	69.3	13.3	6.0	12.9	34.2	30.6	96.5
Post–World War II	40–52	27–37	50.6	51.9	105.1	8.4	16.0	14.8	97.2	90.5	98.0

Rate of Change of Money: Rate of Change of Income Assumed Exogenous										
Full period	5–51	2–36	4.45	3.50	4.05	2.79	1.91	2.28	70.0	67.1	69.1
Full period ex-wars			3.61	2.44	3.09	2.64	1.37	1.94	62.8	70.4	65.0
Pre–World War I	5–23	2–10	1.99	0.81	2.44	3.12	0.36	1.69	−52.1	75.7	−1.2
Interwar	29–35	16–22	3.13	1.95	2.69	2.74	1.53	2.24	63.9	41.5	64.5
Post–World War II	41–51	28–36	2.05	4.51	3.39	1.45	2.96	2.33	51.0	−67.1	−63.9
Rate of Change of Income: Rate of Change of Money Assumed Exogenous											
Full period	5–51	2–36	5.09	3.33	4.10	2.79	1.91	2.28	60.7	70.2	68.3
Full period ex-wars			4.33	2.52	3.28	2.64	1.37	1.94	46.5	68.5	60.6
Pre–World War I	5–23	2–10	2.53	0.73	1.68	3.12	0.36	1.69	−145.8	80.2	52.0
Interwar	29–35	16–22	4.56	2.00	3.76	2.74	1.53	2.24	23.4	38.4	30.7
Post–World War II	41–51	28–36	1.18	2.29	1.82	1.45	2.96	2.33	50.0	56.9	52.8

Note: Based on unadjusted money and velocity for the United States.

[a] For levels, figures given are standard deviations of natural logarithms multiplied by 100, which can be regarded as estimates of the coefficients of variation of the original observations, expressed as a percentage. For rates of change, figures are standard deviations of the rates of change expressed as percentages, hence are in percentage points.

[b] "Maximum" residual because these are computed without allowing for errors of measurement of variable assumed exogenous, hence the reported residual includes such errors of measurement as part of the nonexplained part of dependent variable (see chap. 5, notes 28 and 29).

[c] For levels, equal to standard deviation of velocity; for rates of change, to the root-mean-square of the rate of change of velocity, that is, to the square root of the sum of the variance and the square of the mean rate of change. (See notes 7 and 10, pp. 211, 214.)

[d] Computed as unity minus square of ratio of residual standard deviation to total standard deviation. "Minimum" for reason explained in note b above.

[e] For levels, data for the United States and United Kingdom combined by converting United Kingdom figure in pounds to dollars at contemporaneous rate of exchange. For rates of change, data for both United States and United Kingdom are pure numbers per unit time, hence can be combined without adjustment.

ences between the trend rates of growth in money and income is consistent with equation (1), but it provides, as it were, only a single observation that receives a heavy weight in the percentages cited in the preceding paragraph.

Yet even if we eliminate trends, equation (1) carries us a long way. As noted, rates of change are a sensitive way of allowing for trends.[9] Equation (1) implies that the rate of change of money should equal the rate of change of income, that is, that the rate of change of velocity should be zero. Hence, the root-mean-square value of the rate of change of velocity is an upper limit of the residual variability of rates of change of income or money not accounted for by equation (1) (upper limit because it includes errors of observation in both M and Y).[10]

Using this estimate gives a *minimum* estimate of the fraction of variation accounted for by equation (1). As columns 7, 8, and 9 of table 6.1 show, this *minimum* estimate ranges from 46 to 69 percent for the full period, including or excluding war phases. For other periods the result is much more mixed, the minimum being negative in six out of eighteen observations. Yet even for the subperiods the minimum percentage exceeds 50 for seven of the twelve positive observations.

Nonetheless, there is much absolute variation remaining to be accounted for after allowing for a constant velocity. The maximum coefficient of variation of the residual variability in velocity ranges from 3 to 34 percent for various periods and countries. The maximum standard deviation of the residual variability in rates of change varies from 0.4 to

---

[9] Note that fitting a single straight-line trend to the logarithms of money and income would give results intermediate between those for the level figures for the period as a whole and the rate-of-change figures, since there are substantial differences between the single trends and trends fitted to subperiods.

[10] Differentiate equation (a) of footnote 7 above with respect to time:

$$\text{(h)} \qquad \frac{1}{M}\frac{dM}{dt} = \frac{1}{Y}\frac{dY}{dt},$$

or, in simpler notation,

$$\text{(j)} \qquad g_M = g_Y.$$

Hence the rate of change of either variable is to be predicted from the rate of change of the other by assuming them equal. Again distinguish predicted from observed, and we have that the error in the predicted g_M is

$$g_M - \hat{g}_M = g_M - g_Y = g_k = -g_V.$$

Hence the root-mean-square of the rate of change of velocity is the root-mean-square error. Note that the root-mean-square is *not* equal to the standard deviation of g_V because g_V may not average zero. Rather,

$$(Eg^2{}_V)^{1/2} = [\sigma^2_{g_V} + (\bar{g}_V)^2]^{1/2}.$$

As for levels, the root-mean-square value is the same for g_M and g_Y, and both are maximum errors because of errors of measurement.

3.1 percentage points. These are appreciable magnitudes, and we shall have a highly imperfect understanding of monetary relations until we can account for them.[11] Yet they are generally much less than the total variation in levels or rates of change of money and income.

Table 6.1 makes it clear that a numerically constant velocity does not deserve the sneering condescension that has become the conventional stance of economists. It is an impressive first approximation that by almost any measure accounts for a good deal more than half of the phase-to-phase movements in money or income. Almost certainly, measurement errors aside, it accounts for a far larger part of such movements than the other extreme hypothesis—that velocity is a will-o'-the-wisp reflecting independent changes in money and income. Yet, for most of the period since the mid-1930s, the will-o'-the-wisp extreme has been nearly the orthodox view among economists!

Our ultimate objective is an explanation of the behavior of velocity, which is to say, of the quantity of money demanded, that takes account simultaneously of all the variables affecting velocity. Nonetheless, we believe that this ultimate objective is better approached indirectly, by examining variables one or two at a time, than by what has become the prevailing fashion in econometric work, the immediate computation of multiple regressions including all variables that can reasonably be regarded as relevant. We believe that the indirect approach yields insights that cannot be obtained from the more sweeping approach—that multiple correlations with many variables are almost impossible to interpret correctly unless they are backed by more intensive investigations of smaller sets of variables. Accordingly, we shall proceed in the following sections to consider variables one or two at a time, and reserve to section 6.7 estimating their simultaneous effect.[12]

---

[11] In light of the contrast that we and others have drawn between Keynesian and quantity theories, it is fascinating that it was precisely with respect to this point that Keynes made his famous remark about the long run. Having noted "that the quantity theory is often stated" in a form in which velocity is a constant and output is determined independently of the quantity of money, he went on to say, "Now 'in the long run' this is probably true. If, after the American Civil War, the American dollar had been stabilised and defined by law at 10 per cent below its present value, it would be safe to assume that . . . [the price level] would now be just 10 per cent greater than [it actually is]. . . . But this *long run* is a misleading guide to current affairs. *In the long run* we are all dead. Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is long past, the ocean is flat again."

Like much of Keynes's own subsequent research, our monetary studies and those of other scholars are an attempt to respond to the challenge stated so colorfully by Keynes. J. M. Keynes, *A Tract on Monetary Reform* [1923], Royal Economic Society edition (London: Macmillan, 1971), p. 65.

[12] The indirect approach played a critical role in the formulation of the multiple regressions that we calculate in section 6.7. At the same time, the partial results in sections

6.3 Effect of Financial Sophistication

The character of the financial system clearly affects velocity—indeed, it affects what items we designate as money. The development of banks of issue led students of money to add bank notes to coins and specie; the development of government money issue, to add government fiduciary or fiat currency; and the development of commercial banking, to add deposits, though with much disagreement about the kinds of deposits.

The financial structures of the United States and the United Kingdom have had enough in common and have remained sufficiently constant for the past century that we have found it feasible to use the same basic definition of money for both countries and the whole period, with only one minor exception.[13] However, while the general character of the financial structures has remained the same, there have been substantial changes in detail.

We concluded in chapter 5 (sec. 5.1.2) that the more rapid spread of financial institutions in the United States than in the United Kingdom after 1880 was probably the main reason for the near elimination by 1903 of the wide difference in velocity that prevailed in 1876–77—a difference of 2.8 to 1. Clearly, we must somehow or other allow for this effect.

A full analysis would require identifying the features of financial organization that are most directly relevant to the demand for money and measuring the separate influence of each—such features as number of bank offices per capita, average minimum distance of a bank office from residents in a specified area, fraction of the population having deposit accounts, detailed costs of and returns from deposit accounts, variables connected with the "quality" of deposits as judged by depositors, and so on. More broadly, and particularly for countries other than the United States and the United Kingdom, a major feature is the use of giro rather than checking systems.[14]

Such a full analysis, though it would be extremely illuminating, is clearly beyond the scope of this study. In lieu thereof, we tried to see whether a few simple variables (such as the deposit-currency ratio as a proxy for the quality of deposits) could capture the major effects of the

---

6.4, 6.5, and 6.6 are similar to the results from the final regressions. Hence the reader who is interested primarily in the results and not their derivation may want to skip these sections. This comment does not apply to section 6.3.

[13] That exception is the exclusion of large negotiable certificates of deposit from the United States money stock since 1961. See Friedman and Schwartz, *Monetary Statistics*, pp. 80–81.

[14] A giro system was introduced in the United Kingdom in 1968. Although deposits grew from £10 to £145 million in March 1975, they were only 0.5 percent of United Kingdom commercial bank deposits at the later date. Bank of England, *Quarterly Bulletin* 15 (June 1975), table 11/1, p. 198.

spread of banking. These experiments were uniformly unsuccessful. We have therefore resorted to a simple statistical expedient to correct United States data for what, in chapter 5, we called the changing financial sophistication of the United States relative to the United Kingdom.

1. On the basis of chart 5.5, showing velocity (or number of weeks of income) in the United States and the United Kingdom, we conclude that the change was largely completed by the year 1903.

2. We assume that the change affected the amount of money held and therefore velocity by the same percentage each year from 1869 to 1903, increasing money holdings and decreasing velocity each year by that percentage, over and above any change that can be explained by other variables.

3. One way we estimate the magnitude of the change is by using a dummy variable. For statistical regressions in which the dependent variable is the *level* of money or income or number of weeks of income held as money, the dummy variable is taken equal to zero for the United States phase centered on 1904.0 and later phases, and to the number of years elapsing to 1903 for earlier phases. For regressions in which the dependent variable is the *rate of change* of one of the indicated magnitudes, the dummy variable is taken equal to zero for phases centered on 1906 and later phases and equal to one for earlier phases.

4. Different regressions yield different estimates, yet all are within a fairly narrow range. For one set of forty-eight regressions, for example, for different periods, for the United States separately, and for the United States and the United Kingdom, the estimates range from a low of .022 to a high of .032; that is, increasing financial sophistication added between 2.2 and 3.0 percent per year to United States cash balances from 1876 to 1903. Accordingly, a second, and main, way we allow for the change in financial sophistication is to construct an "adjusted" set of money estimates for the United States for years before 1903 by raising the logarithm of the actual money stock by a constant times the number of years elapsing to 1903. The constant we have used is .0250; that is, we treat increasing financial sophistication as adding 2.5 percent per year to desired United States money balances from 1876 to 1903. The adjusted rates of change of money, and of number of weeks of income held as money, are therefore 2.5 percentage points lower than the unadjusted, and the adjusted rates of change of velocity, 2.5 percentage points higher. The effect is to eliminate most of the sharp early decline in velocity that is so prominent a feature of the raw data.

The second way has the advantage that it keeps the allowance for financial sophistication the same while other variables are altered. The first way does not; in practice, other variables with definite trends, and particularly steeper trends before 1903 than after, tend to absorb the effect of financial sophistication in their coefficients.[15] Of course, this consideration is not decisive. It may be that our adjustment for financial sophistication is in fact absorbing part of the effect of these other variables. However, a variety of considerations lead us to conclude the opposite.

Allowing for the changing degree of financial sophistication by adjusting our money estimates has a substantial effect (table 6.2). As we have interpreted it, this factor operates for 30 percent of the total period and for three-quarters of the period before World War I. Yet for the United States the adjustment reduces the variance of the aggregate level of money by 16 percent, of real per capita money balances by 35 percent, and of velocity by over 70 percent. To put it differently, table 6.1 showed that constant velocity alone, with income treated as exogenous, accounts for at least 96.1 percent of the variance of money and leaves a maximum residual variation of 34.0 percent. For the adjusted money series, constant velocity alone, with income again treated as exogenous, leaves a residual variation of 18 percent (see table 6.3, line 1, col. 2). In consequence, constant velocity plus the adjustment for financial sophistication account for no less than 99 percent of the total variance of the initial money series.

Put the other way around, treating money as exogenous, the initial money series plus constant velocity account for at least 94.5 percent of the variance of nominal income; substituting the adjusted money series accounts for more than two-thirds of the residual variation, raising the percentage of the variance accounted for to at least 98.5.

The effect of adjusting the money series is much the same for the United States for the period excluding wars and is smaller, though still very substantial indeed, for the United States plus the United Kingdom, for either the total period or the period excluding wars. The effect is, of course, even larger for the pre–World War I period, where the whole of the adjustment is concentrated.

For rates of change, the adjustment has little effect on money—indeed, the adjusted variance is somewhat higher than the unadjusted. The unadjusted rates of change for the earlier period are not far from the average for the period as a whole, thanks to the high rates of change during World Wars I and II, so lowering them does not much affect the standard deviation of the rates of change of money. If the war phases are excluded, the adjustment does lower the variance. More important, the adjustment reduces very appreciably the variation of per capita real

---

[15] A specially clear example is provided by population. In one set of level regressions, for example, the inclusion of population yields coefficients of the dummy variable ranging from .0011 to .0096 in absolute magnitude, far smaller than in other regressions.

Table 6.2 **Effect of Changing Financial Sophistication on Variability in Money and Velocity**

	Variation (Percent or Percentage Points)[a]						Fraction of Variance Accounted for by Increasing Financial Sophistication in U.S.		
	Full Period		Full Period Ex-Wars		Pre–WW I				
	Original	Adjusted	Original	Adjusted	Original	Adjusted	Full Period	Full Period Ex-Wars	Pre–WW I
				United States					
Level									
Money	172.6	158.6	182.8	167.9	71.7	48.9	15.56	15.64	53.49
Per capita real money balances	77.8	62.5	80.8	64.3	47.6	23.9	35.46	36.67	74.80
Velocity	34.0	18.0	34.4	16.7	29.6	7.2	71.66	76.43	94.08
				United States and United Kingdom					
Level									
Money	148.6	142.2	156.9	149.9	56.6	47.2	8.43	8.72	30.46
Per capita real money balances	60.6	51.3	62.6	52.4	39.6	19.8	28.34	29.93	75.00
Velocity	29.2	18.2	29.2	16.6	33.0	13.3	61.15	67.68	83.76
				United States					
Rate of Change									
Money	4.45	4.53	3.61	3.26	1.99	2.12	−3.63	18.45	−13.49
Per capita real money balances	2.57	2.25	2.36	1.58	1.72	1.25	76.65	44.82	52.82
Velocity	2.59	2.31	2.23	1.85	1.74	1.54	20.45	31.18	21.67
				United States and United Kingdom					
Rate of Change									
Money	4.04	3.88	3.09	2.81	2.44	1.75	7.76	17.30	48.56
Per capita real money balances	2.30	2.04	1.98	1.49	1.98	1.15	76.66	56.63	33.73
Velocity	2.27	2.08	1.92	1.60	1.59	0.92	16.04	30.56	66.52

[a] For levels, figures given are standard deviations of natural logarithms multiplied by 100, which can be regarded as estimates of the coefficients of variation of the original observations, expressed as a percentage. For rates of change, figures are standard deviations of the rates of change expressed as percentages, hence are in percentage points.

Table 6.3 **Standard Deviation of Logarithms of United States and United Kingdom Velocities: United States Velocity Based on Money Adjusted and Unadjusted for Changing Financial Sophistication**

			Standard Deviation of Velocity (Percentage)				
	Phases		United States			United States and United Kingdom	
Period	United States	United Kingdom	Unadjusted (1)	Adjusted (2)	United Kingdom (3)	Unadjusted (4)	Adjusted (5)
1. Full	4–52	1–37	34.0	18.1	16.7	29.2	18.2
2. Full ex-wars			34.4	16.7	14.8	29.2	16.6
3. Pre–World War I	4–24	1–11	29.6	7.2	3.1	33.0	13.3
4. Equal financial sophistication	18–52	8–37	15.4		19.4	17.5	

balances for both the United States and the United States and the United Kingdom combined, for all periods considered.

Our adjustment for this factor treats it as no longer operative after 1903. That cannot be strictly valid. Many and sizable changes in financial organization have occurred since then in both the United States and the United Kingdom, particularly rapidly in recent decades. The apparent effect of such changes on United States money-holding propensities before 1903 raises the question whether they have not been an important factor in both countries since then. That possibility adds to the desirability of a full study of the effect of financial structure on the demand for money. However, comparison of the variability of velocity, adjusted and unadjusted, for different periods (table 6.3) suggests that the adjustment we have made for the United States before 1903 is of much greater magnitude than any further adjustment for this factor and renders the data for the different periods relatively homogeneous. The adjusted standard deviations for the periods including the pre-1903 period for the United States are much more like those for the United States for subsequent periods—the whole period after 1903 (line 4 of table 6.3), the interwar and post–World War II periods—than are the unadjusted standard deviations. In addition, the adjusted United States standard deviations are much closer to the standard deviations for the United Kingdom and for the two countries combined than the unadjusted—even though some of the adjusted standard deviations do differ significantly both between periods and countries.

Throughout the rest of the book, all references to United States money figures, unless otherwise noted, are to the figures adjusted for the changing degree of financial sophistication by the second of the two methods outlined above.[16]

6.4 Effect of Real per Capita Income

6.4.1 Levels

Equation (1), corresponding to a constant numerical velocity, is entirely in nominal terms. But it can be regarded as a special case of a more general function that separates out real per capita income and money from population and prices, for example, the function

$$\left(\frac{M}{NP}\right)(N)(P) = k\left(\frac{Y}{NP}\right)^{\alpha} N^{\beta} P^{\gamma} \tag{3}$$

---

[16] To make sure that this procedure did not introduce a bias, for almost all our calculations we made parallel estimates for the period after 1903, which we treated as corresponding to equal financial sophistication in the two countries. The results conformed so closely to those for the period as a whole based on adjusted observations that we have not thought it worthwhile to include the results for that period in the tables that follow.

where N is population and P is the price level so that M/NP, which we shall hereafter designate m, is real per capita cash balances and Y/NP, which we shall designate as y, is real per capita income. Equation (1) is a special case of equation (3) for $\alpha = \beta = \gamma = 1$.

In general, economists have tended to take $\beta = \gamma = 1$, on the grounds, for γ, that doubling all prices is equivalent simply to a change in units and will therefore double nominal money demanded for double the nominal income; for β, that the relevant function is for the individual economic unit—household or business—and that doubling the number of units would therefore double the nominal money demanded for double the nominal income.

This reasoning is unobjectionable for different levels of prices, all characterized by the same rate of change of prices (generally implicitly taken to be zero). However, it takes time for a change from one level to another to have its full effect. Hence the observed γ may not equal unity for a particular set of empirical observations. If it does not, it presumably would reflect incomplete adjustment to price changes and therefore be less than unity.

For population, the reasoning is unobjectionable if different sizes of population are obtained by changing the unit of observation—going from an individual state in the United States to a collection of states or to the country as a whole. However, the effect of a change of population is not at all clear for time-series observations for the same geographical area, such as those with which we are primarily concerned. Larger population means a higher density of population, which reduces the average distance between persons and enables different patterns of transactions to develop. It would be preferable to allow for such factors directly. However, if that is not done, and population is used as a proxy, it may well be that the observed value of β may not equal unity.

It is not clear a priori whether β will tend to exceed unity or be less than unity: the greater ease of transactions in a denser market would lead to a β less than unity; the greater division of labor and the more extensive chain of intermediaries between the provision of factor services and the ultimate consumer would lead to a β greater than unity.

For the rest of this section, we shall assume $\beta = \gamma = 1$, but in the next section we shall investigate this assumption to see whether it is indeed supported by our data.

For real per capita income, there is no a priori reason to expect α to equal unity or to have a constant value over time. Our earlier work has led us to the empirical conclusion that α is generally greater than unity; that is, that money is, in the terminology of consumption theory, a "luxury" rather than a "necessity."[17] As we noted in chapter 5, the

---

[17] M. Friedman, *The Demand for Money: Some Theoretical and Empirical Results*, Occasional Paper 68 (New York: NBER, 1959), reprinted in M. Friedman, *The Optimum*

comparisons with the United Kingdom persuade us that our earlier estimates of the income elasticity of demand for real balances for the United States may be too high, biased upward by our failure to allow for changes in the degree of financial sophistication. Hence a reexamination of this issue is clearly called for.

With $\beta = \gamma = 1$, equation (3) can be rewritten

$$m = ky^{\alpha} . \tag{4}$$

If both sides are divided by y, we have an equation for desired number of time units (e.g., weeks or years) of income held as money, or

$$\frac{m}{y} = ky^{\alpha - 1}. \tag{5}$$

Taking the reciprocal of both sides gives

$$V = \frac{1}{k} y^{1 - \alpha} , \tag{6}$$

where V is velocity, turnovers per time unit.

If we take logarithms of equation (4), we have

$$\log m = \log k + \alpha \log y. \tag{7}$$

This is the form in which we shall fit the equation statistically, both because it gives a linear equation to fit and because the error term that needs to be added to equation (7) to convert it into a stochastic function is more nearly homoskedastic in the logarithms than in the original values.

The use of per capita real balances and per capita real income is correct on economic grounds if $\beta = \gamma = 1$. However, it does introduce a purely statistical bias into equation (7) and similar equations. The basic data consist of separate estimates of Y, M, N, and P. We then calculate m and y from these estimates. The estimates of M, N, and P are statistically independent—the estimates of M coming from the books of financial institutions, of N from population censuses, and of P from price indexes or national income statistics. There is some dependence between the estimates of Y and P, since the price level we use is the price level implicit in the calculation of national income at constant prices. However, the national income calculations are mostly based on data for nominal income, and real income in individual sectors is derived by dividing by an independent price index. Hence, even for Y and P, the two magnitudes can be regarded as largely independent.

Quantity of Money and Other Essays (Chicago: Aldine, 1969); M. Friedman and A. J. Schwartz, "Money and Business Cycles," *Review of Economics and Statistics*, 45, no. 1, part 2, suppl. (February 1963): pp. 43–45, reprinted in M. Friedman, *Optimum Quantity of Money*; Friedman and Schwartz, *A Monetary History*, pp. 639, 679–82.

But if Y, M, N, and P are statistically independent magnitudes, m and y are not, since

$$\log m = \log M - \log N - \log P$$

$$\log y = \log Y - \log N - \log P.$$

On economic grounds, use of the same values of N and P to calculate both m and y is desirable, since in effect it enables us to impose the economic restriction that $\beta = \gamma = 1$. It would raise no statistical problem if N and P were measured without error. But obviously they are not. And errors of measurement in N and P introduce common errors into m and y, since the same numerical estimates are subtracted from $\log M$ and $\log Y$ in calculating $\log m$ and $\log y$. These common errors tend to bias the statistical estimate of α toward unity.[18] We know no simple way to estimate the size of this bias, so we shall have to allow for it only qualitatively.

[18] Express the logarithm of each variable as the sum of a "true" value and an error of measurement, say.

$$\log M = (\log M)^* + e_M$$

$$\log Y = (\log Y)^* + e_Y$$

$$\log NP = (\log NP)^* + e_{NP},$$

where an asterisk designates a true value and we have combined N and P because what matters for our purpose is only their product. We then have

$$\log m = (\log m)^* + e_M - e_{NP}$$

$$\log y = (\log y)^* + e_Y - e_{NP}.$$

Assume that e_M, e_Y, and e_{NP} are independent of the asterisked variables and of one another, and that all variables are expressed as deviations from their means. We have two "true" estimates of α, one from the regression of $(\log m)^*$ on $(\log y)^*$, the other from the regression of $(\log y)^*$ on $(\log m)^*$, say

$$\alpha^*_{my} = \frac{E[(\log m)^* (\log y)^*]}{\sigma^2_{(\log y)^*}},$$

$$\alpha^*_{ym} = \frac{\sigma^2_{(\log m)^*}}{E[(\log m)^* (\log y)^*]},$$

where E stands for expected value, and, if the coefficients are positive, $\alpha^*_{my} < \alpha^*_{ym}$.

We have two corresponding statistical estimates of α, say

$$\alpha_{my} = \frac{E[(\log m)(\log y)]}{\sigma^2_{\log y}} = \frac{E[(\log m)^* (\log y)^*] + \sigma^2_{e_{NP}}}{\sigma^2_{(\log y)^*} + \sigma^2_{e_Y} + \sigma^2_{e_{NP}}}$$

$$= \frac{\alpha^*_{my} + \dfrac{\sigma^2_{e_{NP}}}{\sigma^2_{(\log y)^*}}}{1 + \dfrac{\sigma^2_{e_Y} + \sigma^2_{e_{NP}}}{\sigma^2_{(\log y)^*}}},$$

The Separate Countries

Chart 6.1 gives scatter diagrams of log *m* and log *y*, panel A for the United States and panel B for the United Kingdom. In each panel three lines are drawn through the point corresponding to the mean values of log *m* and log *y*: a line with a slope of unity, which is the counterpart of

---

$$\alpha_{ym} = \frac{\sigma^2_{\log m}}{E[(\log m)(\log y)]} = \frac{\sigma^2_{(\log m)*} + \sigma^2_{e_M} + \sigma^2_{e_{NP}}}{E[(\log m)^*(\log y)^*] + \sigma^2_{e_{NP}}}$$

$$= \frac{\alpha^*_{ym} + \dfrac{\sigma^2_{e_M} + \sigma^2_{e_{NP}}}{E[(\log m)^*(\log y)^*]}}{1 + \dfrac{\sigma^2_{e_{NP}}}{E[(\log m)^*(\log y)^*]}}.$$

The statistical estimates differ from the true values because of errors in measuring (a) *Y* and *M* and (b) *NP*.

With respect to (a), this is the standard regression effect, as can be seen by supposing *NP* to be measured without error, so that $\sigma_{e_{NP}} = 0$. We then have:

$$\alpha_{my} = \frac{\alpha^*_{my}}{1 + \dfrac{\sigma^2_{e_Y}}{\sigma^2_{(\log y)^*}}} < \alpha^*_{my} < \alpha^*_{ym} < \alpha^*_{ym} + \frac{\sigma^2_{e_M}}{E(\log m)^*\ (\log y)^*} = \alpha_{ym},$$

that is, the calculated regression coefficients will cover a wider range than the "true" ones.

To isolate the effect of (b), namely, errors in *NP*, suppose *Y* and *M* measured without error, so $\sigma_{e_Y} = \sigma_{e_M} = 0$. We then have:

$$\alpha_{my} = \frac{\alpha^*_{my} + \dfrac{\sigma^2_{e_{NP}}}{\sigma^2_{(\log y)^*}}}{1 + \dfrac{\sigma^2_{e_{NP}}}{\sigma^2_{(\log y)^*}}}$$

$$\alpha_{ym} = \frac{\alpha^*_{ym} + \dfrac{\sigma^2_{e_{NP}}}{E[(\log m)^*(\log y)^*]}}{1 + \dfrac{\sigma^2_{e_{NP}}}{E[(\log m)^*(\log y)^*]}}.$$

Both these expressions are of the form of weighted averages of the "true" coefficient and unity, the weights being 1 and $\sigma^2_{e_{NP}}/\sigma^2_{(\log y)^*}$ for α_{my} and 1 and $\sigma^2_{e_{NP}}/E[(\log m)^*(\log y)^*]$ for α_{ym}. Accordingly, the effect of errors in *NP* is to bias the computed coefficients toward unity by comparison with the true coefficients. This conclusion remains valid when the assumption that $\sigma_{e_M} = \sigma_{e_Y} = 0$ is dropped; that is, computed coefficients are biased toward unity compared with hypothetical coefficients computed under the assumption that $\sigma_{NP} = 0$ but σ_{e_M} and σ_{e_Y} are whatever they in fact are.

The possible magnitude of the regression effect can be judged by computing estimates of α from both regressions and regarding these as upper and lower limits of the correct values. There is no similar simple way that we know of to judge the size of the bias arising from the errors in *N* and *P*, since that depends on the size of the measurement errors in *N* and *P* compared with the variance of log *Y* and log *M*.

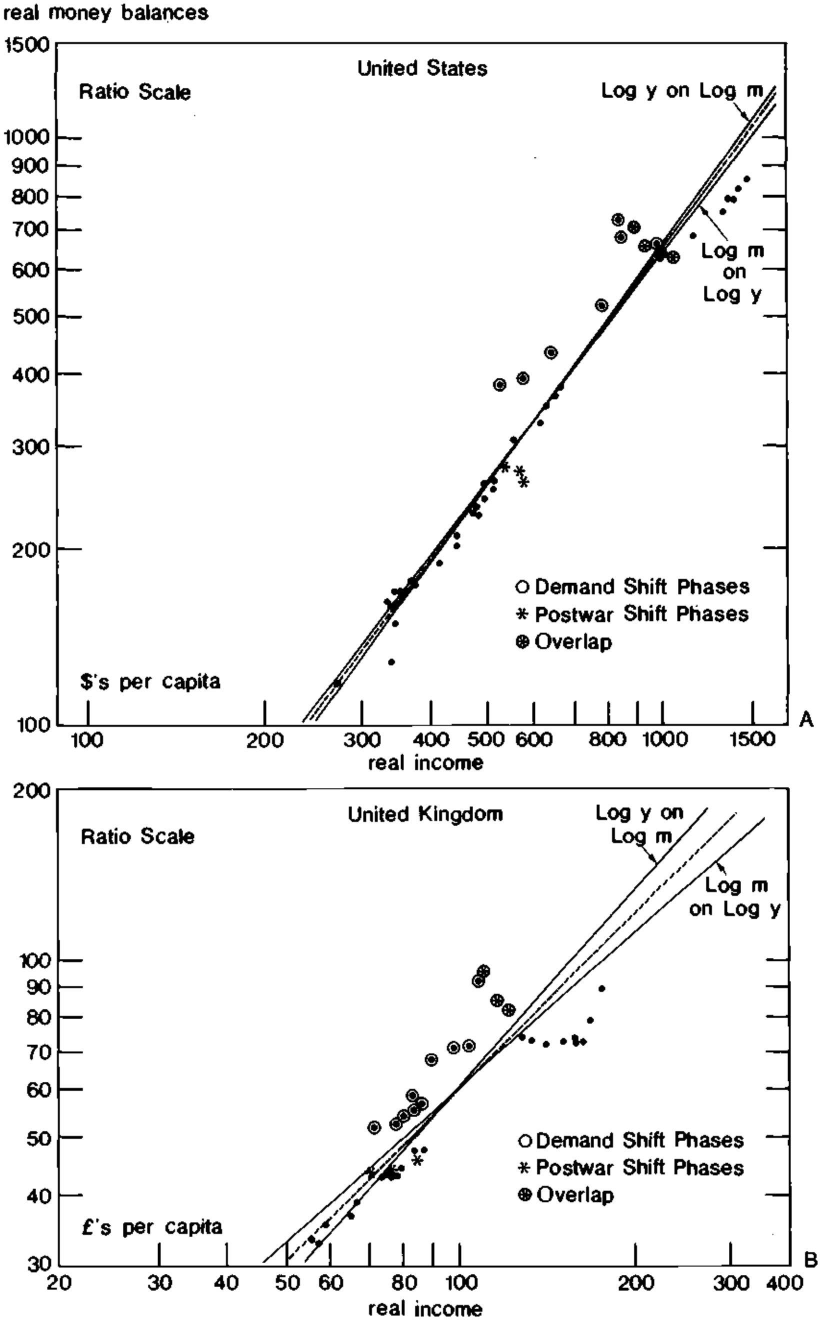

Chart 6.1 Scatter diagram of levels of real per capita income and real per capita money and regression lines.

Table 6.4 **Income Elasticities and Correlation Coefficients, Calculated from Levels of Real per Capita Money and Income**

Period and Country	Phase Numbers	Income Elasticities: Lower Limit (1)	Income Elasticities: Upper Limit (2)	Correlation Coefficient r (3)
United States				
Full period	4–52	1.26	1.31	.978
Full ex-wars		1.23	1.27	.983
Pre–World War I	4–24	1.20	1.28	.966
Interwar	28–36	0.37	4.32	.296
Postwar	40–52	0.59	0.75	.886
United Kingdom				
Full period	1–37	0.84	1.10	.874
Full ex-wars		0.78	0.93	.918
Pre–World War I	1–11	0.85	0.88	.984
Interwar	15–23	1.26	1.47	.926
Postwar	27–37	−0.02	−13.31	.080
United States and United Kingdom				
Full period	4–52, 1–37	1.13	1.27	.948
Full ex-wars		1.09	1.20	.957
Pre–World War I	4–24, 1–11	1.09	1.19	.961
Interwar	28–36, 15–23	0.85	2.04	.844
Postwar	40–52, 27–37	0.36	1.06	.976

Note: For United States plus United Kingdom, allowance is made for a difference in level between the two countries after converting the United Kingdom figures from pounds to dollars at the exchange rate of $4.862.

equation (1); and two regressions, one of log m on log y, the other of log y on log m, both computed from the phase average data for 1873 to 1975 for the United States, and from 1874 to 1975 for the United Kingdom. The slopes of these lines can be regarded as upper and lower estimates of α.[19] Table 6.4 gives, in columns 1 and 2, the numerical elasticities for these regressions, and also for regressions for the period before World War I, the interwar period, and the post–World War II period.

For the United States, for the period as a whole, the income elasticity is clearly greater than unity. The line with a slope of unity is above most observations for income and money below the mean and below most observations for income and money above the mean (indeed there are only three observations that do not conform to this pattern). Both regressions are steeper than the line with unit slope.

The points that are most out of line are for the 1930s, the World War II years, and the immediate post-World War II years. These indicate much

---

[19] See the preceding note.

larger money holdings (i.e., lower velocity) than the others. We are inclined to attribute this result to the effects of the great contraction of 1929–33 and the war. Declining prices during the contraction meant that money balances had a substantial positive yield, and great uncertainty raised desired liquidity. These effects were partly offset by bank failures, which lowered the "quality" of deposits. The uncertainty lasted long after the contraction itself and was then reinforced by the war; in addition, during the war, the unavailability of goods reinforced uncertainty in raising desired money balances.

These discrepancies are reflected in the much less satisfactory results for the interwar and postwar periods than for the period as a whole or the pre–World War I period. The correlation for the interwar period is close to zero, and the elasticities for the postwar period are far lower than for the period as a whole.

The United Kingdom results differ in two respects. First, only one of the regression lines has a slope greater than unity. The scatter diagram as a whole does not suggest an elasticity greater than unity: about half the points for incomes below the mean are above the line with a slope of unity, and about the same fraction of points for incomes above the mean are below the line. Second, the scatter is much looser, with a correlation coefficient of .87, compared with .98 for the United States.

However, there are also some similarities, most strikingly the division of the points into two sets, one at a higher level than the other. The points at the higher level (circled in the charts) include the same period as those for the United States (1939–54) and in addition all the phases for the 1920s from 1921 on.[20] The additional phases are no less significant than the common ones. The United Kingdom's time of serious economic trouble started in the early 1920s and recovery began in 1932, before the Great Contraction ended in the United States. The forces we adduced to explain the discrepant points for the United States—namely, price deflation, uncertainty arising out of serious cyclical fluctuation reinforced by the outbreak of war, and the unavailability of goods during the war—all operated at least as strongly in the United Kingdom as in the United States but started nearly a decade earlier. It therefore supports this interpretation that higher-level-points for the United Kingdom span a period beginning earlier but ending at roughly the same date as the analogous United States points.

A second similarity between the two countries came to light when we studied the relations between rates of change (see sec. 6.4.2), namely, an immediate postwar adjustment that produced idiosyncratic rates of change after World War I and World War II in both the United States and the United Kingdom—in both countries, in opposite directions after the

---

[20] In addition, because of the difference in reference dates, the final phase at the higher level ends in 1955 rather than 1954.

two wars but by roughly the same amount in the two countries. We had missed these perturbations when we first studied the level data because even a sizable deviation in a rate of change for a phase or two produces only a small effect on the levels.

However, once alerted to them, it is easy to see their effect. For the United States for World War I they are reflected in a substantial deviation of the phase centered on 1919 (designated by the lowest of the three uncircled *'s) below the lower of the two regression lines in chart 6.1, panel A. Indeed, it deviates more from the line than all but one other observation (for the phase centered on 1884). The level of the phase centered on 1920 is also below the line (the second lowest of the three uncircled *'s), the level of the phase centered on 1921 (the third of the uncircled *'s) is nearly on the line. The postwar adjustment reaction apparently carried too far, bringing the phase level centered on 1922.5 well above the line. The same pattern is repeated, but in the other direction, for World War II. The 1947.5 point is well above the upper line (the highest of the circled *'s marked on the chart), the largest deviation with one exception in the upper set. The 1949 point (the second highest circled *) is a trifle above the line, that for the 1952 phase (the third of the circled *'s) well below.

For the United Kingdom a similar, though initially less clear, pattern is present after World War I—the tendency to overshoot carrying the observations into the beginning of the period we isolated as reflecting an upward shift in liquidity. After World War II, the pattern of points for the United Kingdom around the upper regression line in chart 6.1, panel B is almost identical with that for the United States, except it comes later in time. The phase centered on 1949 (the highest circled * marked on the chart) is well above the line, with 1952, 1954, and 1957 moving toward and then below it, and the 1957 phase reflecting both the overshooting and the end of high liquidity preferences.

We shall refer to these postwar perturbations as the postwar shift and to the upward shift in liquidity preference as the demand shift.

When we analyze the effect of other variables influencing the quantity of money demanded later in this chapter, we shall see whether their fluctuations are such as to account for the two shifts: the demand shift and the postwar shift. In the meantime, in order to continue examining the effect of per capita real income by itself, we shall use the device of introducing two shift variables.

To allow for the apparent increase in liquidity preference in the interwar, war, and early postwar periods, we introduce a dummy variable that has the value of one for the phases in the upper cluster, of zero for the other phases. This procedure is equivalent to fitting separate regressions to the two clusters of points but imposing the requirement that they have the same slope, that is, income elasticity.

To allow for the postwar shift, we also introduce a dummy variable for the affected phases, but this one equal in absolute value to the time span beteen the midpont of the affected phase and the midpoint of the first normal postwar phase, and negative after World War I, positive after World War II. The reason, as explained below in section 6.4.2, is that we interpret the adjustment as reflecting a maximum change in the rate of change that holders of money are willing to undertake in these circumstances.[21]

Consistent with our visual impression, the postwar adjustment dummy produces an appreciable minor improvement in results for the level equations, if a less substantial improvement than for the rate of change equations or than the demand shift dummy does for the level equations.[22] Panels A and B of chart 6.2 repeat the scatters from chart 6.1 and add two pairs of parallel lines, one pair corresponding to the regression of log m on log y, the other to the regression of log y on log m, the difference between the lines in each pair reflecting the demand shift and also allowing for the postwar readjustment.

The improvement in fit is striking. Indeed, for the United States the correlation is so high that the log m on log y and log y on log m regressions are difficult to distinguish on a chart the size of chart 6.2, panel A.

The imposition of a common income elasticity on the two clusters seems to do no serious violence to the observed data. The larger number and wider scatter of observations in the lower cluster means that that cluster largely determines the numerical size of the slope. Yet even for the upper cluster the slope seems appropriate—though there is some indication for the United Kingdom that a slightly steeper slope would be preferable.

The remarkable feature of these charts is how closely a single pattern fits such widely separated observations as those before World War I and after World War II. Allowance for the demand and postwar shifts drastically reduces the differences between regressions for the period before World War I, when no shift variables are used, and similar regressions for the period thereafter, including shift variables. Once the shift variables are included, little is gained by distinguishing war phases from the other phases, or by any of the other subgroups we have used.[23]

---

[21] See below, following equation (8), for a more detailed listing of the values of the dummy variables.

[22] The t-value for the coefficient of the postwar adjustment dummy is 2.8 for the United States, 3.6 for the United Kingdom, and 2.1 for the two countries combined, compared with corresponding t-values for the demand shift dummy of 9.7, 10.8, and 9.9. The numerical size of the estimated coefficient of the postwar adjustment dummy is about 2 percent per year for both the United States and the United Kingdom, about 1.5 percent per year for the two countries combined, which is less than one-third of the estimate derived from the rate of change data.

[23] We have consistently made calculations reported in the rest of this chapter for the

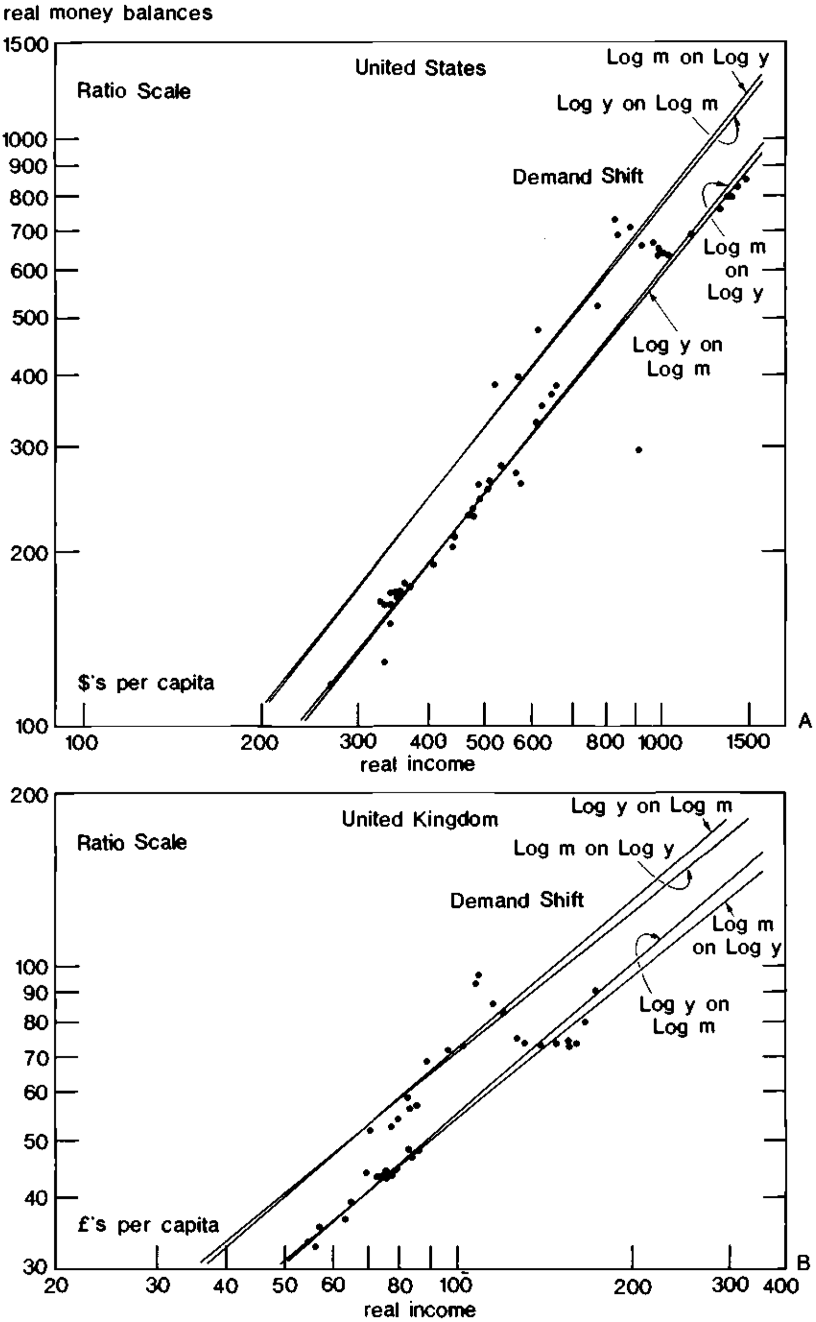

Chart 6.2 Scatter diagram of levels of real per capita income and real per capita money and regression lines isolating effect of demand shift.

Table 6.5 **Effect of Allowing for Postwar Adjustment and Shift Dummies on Difference between Variances and Regressions**

	Simple Regressons, Dummies		Multiple Regressions, Dummies	
Source of Variation[a]	Absent	Present	Absent	Present
	Mean Sum of Squared Residuals			
United States				
Within first period (22)	.0038		.0028	
Within second period (27)(26)[b]	.0196	.0031	.0058	.0020
Two periods combined	.0127	.0034	.0045	.0024
Between two periods	.1254	.0222	.0443	.0048
United Kingdom				
Within first period (12)	.0005		.0004	
Within second period (25) (24)[b]	.0247	.0061	.0049	.0027
Two periods combined	.0153	.0038	.0032	.0018
Between two periods	.1968	.0060	.0507	.0114
	F Values and Significance of Difference			
United States				
Between two periods in				
Residual variances	5.2**	0.8	2.1	0.7
Regressions	9.9**	6.5**	9.8**	2.0
United Kingdom				
Between two periods in				
Residual variances	49.4**	12.2**	12.2**	6.8**
Regressions	12.9**	1.6	15.8**	6.3**

[a] And number of observations, *not* degrees of freedom.
[b] Number of observations one greater for simple regressions than for multiple regressions.
*Significant at .05 level.
**Significant at .01 or more stringent level.

Table 6.5 summarizes the effect of allowing for the demand and postwar shift dummies on the difference between the period through 1918 (first period) and the period thereafter (second period) with respect to the residual variances within the two periods and also the regression equations for the two periods. We present these calculations for the simple regression equations that, aside from the dummies, include only per capita real money balances and per capita real income and also for the multiple regression equations that we finally settle on in section 6.7, which include in addition two independent variables measuring the yield on alternative assets.

The residual variance within the second period is reduced sharply by allowing for the dummies, by between 45 percent and 84 percent. For the

---

subperiods as well as for the period as a whole. However, we have found them for the most part to add so little to the single calculation including one or more shift variables that it has not seemed worth including them in the tables.

United States, the resulting residual variance for the second period does not differ significantly from that for the first: the set of data becomes homogeneous in that respect. For the United Kingdom there remains a significant difference, which reflects the abnormally low variance for the first period, in our opinion a statistical artifact arising from the extensive interpolation entering into the early United Kingdom data.

The variance of the difference between the two periods is reduced even more sharply, by between 78 and 97 percent. For the United States, the difference between the two periods remains statistically significant for the simple regression but not the multiple—apparently the time difference reflects primarily differences in average interest rates. For the United Kingdom the situation is reversed: there is no significant difference for the simple regression; there is for the multiple. However, in light of the major reduction in both the variance between periods and the F values, and our doubts about the statistical validity of the variance within the first period, we do not regard this one result as contradicting the validity of treating the whole period as homogeneous after allowing for the two shifts.

The main difference between the two countries is the steeper slope of the regression for the United States than for the United Kingdom—a higher than unit elasticity for the United States, a lower than unit elasticity for the United Kingdom.

Table 6.6 summarizes the numerical effect of allowing for the demand and postwar shifts. Remarkably, the size of both shifts is almost identical in the two countries. For the demand shift, desired money balances were apparently raised by about 30 percent at each level of per capita income by the factors associated with Depression and its aftermath and World War II. For the postwar shift, desired money balances were lowered after World War I and raised after World War II by the limited rate at which balances readjusted from their abnormal wartime levels to peacetime levels. The percentage discrepancy on this account narrowed at a rate equal to roughly two percentage points per year; that is, the percentage discrepancy was equal in absolute value to twice the time interval between each of the relevant phases and the first postwar phase we treated as normal (phases centered on 1922.5 and 1954.0 for the United States and on 1923.0 and 1957.0 for the United Kingdom).

The estimated income elasticity is 1.2 for the United States, about 0.8 for the United Kingdom. This difference raises a real puzzle. Does it reflect a fundamental difference between the two countries in the response of the quantity of money demanded to a change in real income? Or may it be a disguised reflection of a differential behavior in the two countries of other variables affecting the quantity of money demanded? We shall pursue that question in the later sections of this chapter. Here we may only note that the difference in computed elasticities is large

Table 6.6 **Relation between Real per Capita Money and Real per Capita Income: Effect of Allowing for Upward Shifts in Money Demand during Interwar, World War II and Early Postwar Period, and for Postwar Shifts**

	United States	United Kingdom	United States and United Kingdom
1a. Period of upward demand shift	1929–54	1921–55	
b. Period of postwar shift	1918–21, 1946–53	1918–21, 1946–55	
2a. Size of demand shift (percentage)[a]	27.7 (27.2)	29.6 (29.6)	30.9 (30.0)
b. Size of postwar shift (percentage)[a,b]	2.3 (2.3)	2.2 (2.0)	1.6 (1.4)
c. Difference in level between United States and United Kingdom (at market exchange rate, as percentage of United States level)[a]			10.7 (12.8)
3. Income elasticity, lower limit	1.19	0.79	1.07
4. Income elasticity, upper limit	1.21	0.83	1.12
5a. Simple correlation (money and income)	0.978	0.874	0.945
b. Partial correlation (allowing for shifts)	0.994	0.974	0.976
6. Multiple correlation, including shifts[a]	0.995 (0.994)	0.983 (0.975)	0.981 (0.980)

	Standard Error of Estimate (%)		
Real per capita money			
7. Total	62.5	32.6	51.3
Maximum residual after allowing for:			
8. Constant velocity	18.1	16.7	18.2
9. Constant elasticity	13.2	16.0	16.6
10. Constant velocity plus shifts	11.7	9.8	11.5
11. Constant elasticity plus shifts	6.5	6.3	10.1
Real per capita income			
12. Total	48.7	33.7	45.0
Maximum residual after allowing for:			
13. Constant velocity	18.1	16.7	18.2
14. Constant elasticity	10.3	16.6	13.9
15. Constant velocity plus shifts	11.7	9.8	11.5
16. Constant elasticity plus shifts	5.4	7.8	9.2

[a] Figures in parentheses from regressions of y on m and dummy (or dummies) solved for m; others from regression of m on y and dummy (or dummies).

[b] Note that the postwar shift is positive after World War I, negative after World War II.

enough, and the residual variation small enough, so that the elasticity difference is highly significant statistically, while the differences in shift effects are not.[24]

The constant velocity assumption of section 6.2 is equivalent to assuming an elasticity of unity between money and income. In view of the closeness of the estimated elasticities to unity, it is not surprising that

---

24. The relevant analysis of variance is as shown in table 6.N.1.

The difference in slope and in level are highly significant, the difference in shifts is not. Similarly, both the common demand shift and the postwar shift are highly significant, and of course the common slope is highly significant.

It should be noted that the partition of the sums of squares among the individual degrees of freedom is not invariant to the order in which the various effects are allowed for. We have generally followed the principle of allowing for the effects in the order of their significance, as indicated by *t*-values.

Table 6.N.1 **Analysis of Variance of Differences between Demand for Money Regressions for the United States and the United Kingdom Based on Levels**

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*
Within United States	45	.1900	.0042(.0039)[a]	
Within United Kingdom	33	.1310	.0040(.0044)[a]	
Within United States and United Kingdom	78	.3197	.0041	
Difference between United States and United Kingdom				
In slope	1	.4943	.4943	120.6
In shifts	2	.0170	.0085	2.1
In level, at market exchange rates	1	.4590	.4590	112.0
Subtotal	82	1.2900		
Attributable to				
Demand shift	1	1.4090	1.4090	343.7
Postwar shift	1	0.0450	0.0450	11.0
Slope	1	19.6680	19.6680	4797.1
Total	85	22.4120		

[a] The total sum of squares does not equal the sum of within country sums because of the need to allow for the effect of unequal weighting of observations. The mean squares within parentheses are estimates of the common mean square after adjustment for difference of weights. The common mean is the weighted average of the parenthetical figures, the weights being the number of degrees of freedom. The two parenthetical mean squares do not differ at a .05 level.

For 1 and 78 degrees of freedom, .05 *F* value is 4.0; .01 *F* value, 7.0; .001 *F* value, 11.8. For 2 and 78 degrees of freedom, the corresponding *F* values are 3.1, 4.9, and 7.6.

allowing for a constant elasticity instead of an elasticity of unity produces only a rather small reduction in the computed standard error of estimate for either real per capita money, with real per capita income taken as exogenous, or real per capita income, with real per capita money taken as exogenous (compare lines 8 and 9 and 13 and 14 in table 6.6). For reasons explained in chapter 5, these two standard errors cannot both be simultaneously valid measures of the variation in each of the variables independent of the variation in the other, because each assumes that the other variable is perfectly known. The entries in lines 9 and 14 of table 6.6, like the entries in lines 8 and 13, are therefore to be regarded as estimates of the *maximum* residual standard error. Simultaneously valid estimates of the independent variation in each variable would necessarily be lower.

Allowing for the demand and postwar shifts in liquidity preference in three out of four cases reduces the maximum residual standard error more drastically than allowing for the difference between constant elasticity and constant velocity; and allowing for both the shifts and the differences between constant elasticity and constant velocity does even better. Allowing for the shifts generally enhances the value of substituting constant elasticity for constant velocity. Together the two roughly cut the standard error in half and in addition, for real per capita money, though not income, bring the standard errors for the United States and the United Kingdom closer together. All these residual standard errors are below 10 percent—and these continue to be maximum estimates.

The constant velocity standard errors are constrained to be the same for money and income. For the United States, all of the rest are higher for money than for income, but for the United Kingdom they are higher for income than for money. The reason is linked to an elasticity higher than unity for the United States and lower than unity for the United Kingdom. The total variability of money is higher in the United States than the total variability of income—which helps to produce the higher than unit elasticity.[25] The reverse is the case in the United Kingdom, which helps to produce the lower than unit elasticity. But the usual convention regards the same fraction of the variance of each variable as "accounted for" by the regression, namely the square of the correlation coefficient given in line 5 of table 6.6. As a result, if the initial standard error is higher (or lower) for log m than for log y, so also must be the residual standard error. If, as we suspect, pure errors of measurement are higher for income than for money, then the "correct" fraction of variance accounted for may well be higher for money than for income, which

---

[25] Because the lower limit is $r\,\frac{\sigma_m}{\sigma_y}$, and the upper limit, $\frac{1}{r}\frac{\sigma_m}{\sigma_y}$. Since r is necessarily less than unity, $^{\sigma}m > {}^{\sigma}y$ is a necessary but not sufficient condition for the lower limit to exceed unity, and a sufficient but not necessary condition for the upper limit to do so.

might well reverse the relation now shown in table 6.6 for the United States and increase the difference for the United Kingdom. We saw in the preceding chapter that the standard deviation of deviations about trend was higher for income than for money, and we shall see in the next section that the standard deviation of rates of change is also higher for income than for money, both of which suggest that the purely random component is higher for income than for money and that table 6.6 may show the opposite for the United States solely because of a statistical convention.

The figures in table 6.6 overstate the case for constant elasticity as compared with constant velocity. We are here dealing with per capita real magnitudes. These have lower variability than the nominal aggregates. A constant velocity (i.e., unit elasticity) demand function would enable us to predict total income from total money, if we assume money exogenous, or total money from total income, if we assume income exogenous. The constant elasticity per capita real demand function would not; we would need to know in addition what is happening to prices and population. Additional "explanation" is therefore bought at the expense of requiring more information.

The Two Countries Combined

Despite the statistically significant difference between the income elasticities for the United States and the United Kingdom derived from the regressions that allow for the demand and postwar shifts, the similarity of the regressions in other respects and the small magnitude of the elasticity difference suggest that it may be worth getting a single estimate by combining the data for the two countries, if only to explore some of the problems raised by such a combination.

We cannot combine the data directly, because the data for the United States are in dollars and those for the United Kingdom are in pounds sterling. One way to combine the data is to convert the United Kingdom figures to dollars at the exchange rate ($4.862) ruling in 1929 (the year that was the base for the price indexes used to estimate real money balances). That is the way we have combined them to get the standard deviations for the two countries in lines 7 and 12 of table 6.6. However, it is far from clear that the market rate of exchange is the relevant one for the purpose of calculating a single regression.

Both for this reason and also to allow for a possible difference in the level of money demand between the two countries, we include in all regressions for the two countries combined a dummy variable (Z) distinguishing the United Kingdom from the United States. With this addition plus dummy variables to allow for the demand (S) and postwar (W) shifts, equation (7) becomes:

(8) $$\log m = \log k + \alpha \log y + \lambda_1 W + \lambda_2 S + \lambda_3 Z ,$$

where $W =$ $-T_i$ for the phases centered on 1919, 1920, and 1921 for the United States and the United Kingdom

$+T_i$ for the phases centered on 1947.5, 1949, and 1951.5 for the United States and 1949, 1952, and 1954 for the United Kingdom

0 for all other phases

$S =$ 1 for the phases from 1929 to 1954 in the United States and from 1921 to 1955 in the United Kingdom

0 for all other phases

$Z =$ 1 if the observation is for the United Kingdom

0 if the observation is for the United States

and where $T_i =$ time interval between the center of phase i and the center of the first postwar phase treated as normal (1922.5 and 1954.0 for the United States, and 1923.0 and 1957.0 for the United Kingdom.).

The parameter λ_3 encompasses two effects: a difference in level produced by the exchange rate required to convert pounds to dollars and any difference in the level of money demand for the two countries at that exchange rate. To demonstrate, distinguish by subscripts the observations for the United States and the United Kingdom. Let EX be the rate of exchange expressed in number of dollars per pound; let the original observations for the United States be in dollars and for the United Kingdom in pounds; and assume that α, λ_1, and λ_2 are the same for the two countries. We then have

(9) $$\log m_{us} = \log k_{us} + \alpha \log y_{us} + \lambda_1 W + \lambda_2 S$$

(10) $$\log EXm_{uk} = \log k_{uk} + \alpha \log EXy_{uk} + \lambda_1 W + \lambda_2 S$$

(11) or $$\log m_{uk} = \log k_{uk} + (\alpha - 1) \log EX + \alpha \log y_{uk} + \lambda_1 W + \lambda_2 S$$

(12) or $$\log m_{uk} = \log k_{us} + \alpha \log y_{uk} + \lambda_1 W + \lambda_2 S + (\log k_{uk} - \log k_{us}) + (\alpha - 1) \log EX.$$

If we now fit the combined data for the United States and the United Kingdom with equation (8), keeping the United States data in dollars and the United Kingdom data in pounds, it is clear that

(13) $$\lambda_3 = (\log k_{uk} - \log k_{us}) + (\alpha - 1) \log EX.$$

The first term reflects any difference in the level of money demand between the two countries at the exchange rate of *EX*; the second reflects the difference in apparent level arising from the exchange rate.[26] If $\alpha = 1$, that is, income elasticity is unity, then no difference in apparent level arises from the exchange rate: velocity is then a constant and hence is independent of the level of income and equally of the monetary units in which income and money are expressed, provided the same rate of exchange is used for both money and income, which is why this problem did not arise in section 6.2, where we analyzed constant velocity.

If $\alpha \neq 1$, it is impossible to decompose the computed value of λ_3 into its two parts without a further assumption. (1) If we assume that the first term is zero, that is, that there is no difference in level, we can use equation (13) to estimate the implied value of *EX*, that is, the exchange rate that would produce identical relations for the United States and the United Kingdom (given that the slopes or elasticities and the size of the shifts are forced to be identical and that the same exchange rate is used for money and income). We shall consider later how to interpret such an implicit exchange rate. (2) Alternatively, if we assume a particular value for *EX*, say $4.862, which was the market value in 1929, the base year for both the United States and United Kingdom real magnitudes, we can use equation (13) to estimate the difference in level between the United States and the United Kingdom relations. The result will of course be precisely the same as if we had calculated equation (8), using United Kingdom figures converted to dollars at that exchange rate.[27]

The entries in table 6.4 for the United States and United Kingdom combined were calculated using this dummy variable procedure but not allowing for the shifts (i.e., treating λ_1 and λ_2 as zero). The entries in table 6.6 for the two countries combined were calculated the same way except allowing for the shifts (i.e., from equation 8). They gave us estimates of the income elasticity and of the upward shifts that are free from any assumption about the exchange rate that is appropriate or about the size of the constant terms in equations (9) and (10).

---

[26] If, instead of computing regression (8) from United States figures in dollars and United Kingdom figures in pounds, we were to use United Kingdom figures converted to dollars at the exchange rate *EX'*, equation (13) would become:

$$(13') \qquad \lambda_3 = (\log k_{uk} - \log k_{us}) + (\alpha - 1)(\log EX - \log EX')$$

In effect, using pound figures is equivalent to using an $EX' = 1$, in which case equation (13') reduces to equation (13). Note that in equation (13') k_{uk} implicitly gives the level of the United Kingdom relation in dollars at the exchange rate of *EX*, not *EX'*, just as it does in equation (13).

[27] See preceding note.

For the 1929 market exchange rate of \$4.862, the difference in level between the United States and the United Kingdom is 11 to 13 percent (line 2*c* of table 6.6). That is, at each level of real income per capita, real balances per capita tend, according to this combined estimate, to be 11 to 13 percent higher in the United Kingdom than in the United States.

Because the average elasticity (1.07–1.12) is so close to unity, it is clear that this estimate will not be highly sensitive to the precise exchange rate used to convert United Kingdom figures to dollars. For example, for reasons discussed in Appendix A to this chapter (sec. 6.8), the purchasing power parity exchange rate has a strong claim to be the relevant exchange rate to use in combining the data. An estimate of the purchasing power parity rate for various years can be constructed from the most recent independent estimate for 1970. That estimate sets it at \$5.50 for 1929.[28] At this exchange rate, the difference in level, as computed from equation (13), would have been 10 to 12 percent instead of 11 to 13 percent. Clearly, no plausible adjustment of the market exchange rate can account for the difference in level.

A far more serious problem about the estimated difference in level arises from neglecting the effect of the significant difference in income elasticity between the two countries. This problem is illustrated by chart 6.3, which plots for the two countries the money on income regressions from equation (8)—these are the two parallel lines—and the regressions calculated for each country separately from the same equation, eliminating, of course, the term Z.[29] In all cases the regressions are for values of the shift dummy variables (W and S) set equal to zero (i.e., they correspond to the lower of the two parallel lines in chart 6.2).

The constraint to a common elasticity imposed by equation (8) fitted to pooled data for the two countries forces parallelism between the resulting lines for the separate countries and so gives a single estimate for the difference between the two countries for all levels of income—the 11 to 13 percent difference recorded in line 2*c* of table 6.6. But this parallelism clearly misrepresents the data—as the other two lines on the chart show. Because the United States line has a higher elasticity than the United Kingdom line, it starts below and ends above the United Kingdom line. At the roughly equal per capita income in the two countries at the beginning of our period (1876 or 1877), the United Kingdom regression gives a level of real per capita money 40 percent higher than does the United States regression. A century later, at the prevailing United Kingdom income, the United Kingdom level is 13 percent below the United States; at the prevailing United States income, now much higher than in

---

[28] See appendix to this chapter, section 6.8.

[29] For simplicity, we show only the money on income regressions. Because of the high correlations, the income on money regressions differ only trivially.

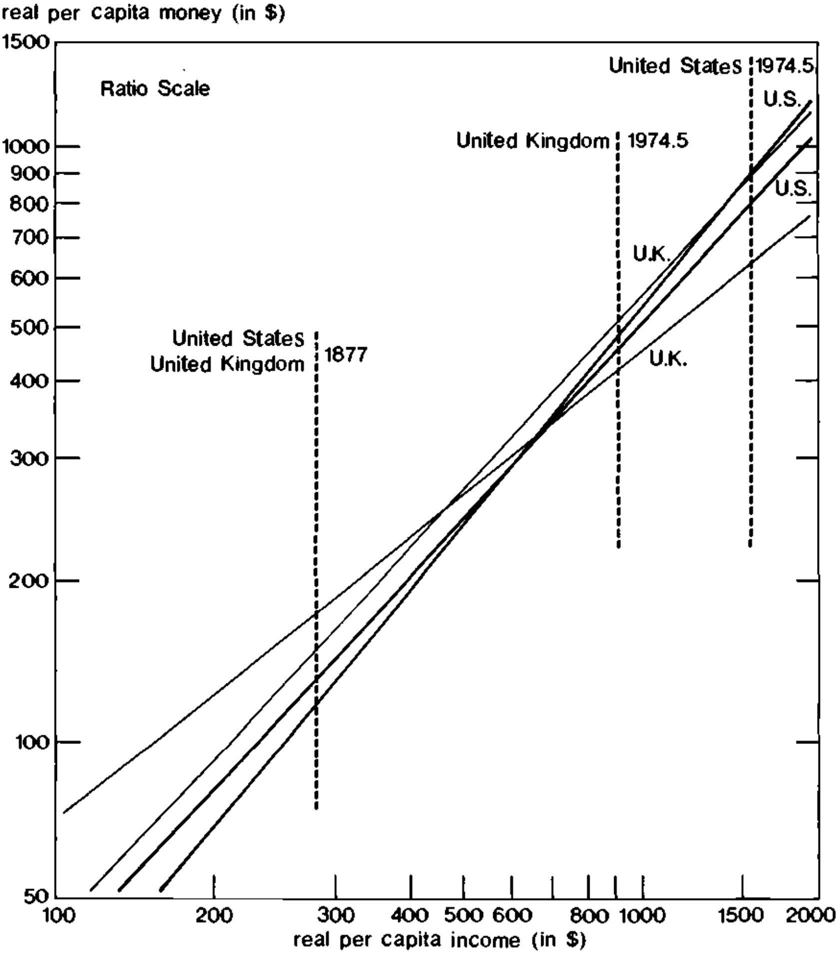

Chart 6.3 Regressions of real money per capita on real income per capita for United States and United Kingdom, elasticities different and constrained to be equal, eliminating common shift.

the United Kingdom, 30 percent. The single estimate of 11 percent is implicitly an average of these very different figures.

This chart puts a strong spotlight on the puzzle referred to above—Why the difference in elasticities? Why, at a comparable level of real income early in the century we cover, should United Kingdom real balances per capita have averaged close to 40 percent higher than United States balances, but at the end of the century, 13 to 30 percent lower? Does this difference reflect a basic discrepancy between the countries in the effect of higher real income, other major variables held constant? Or does it reflect the effect of a differential change between the two countries in those variables?

Any ambignity about the exchange rate to use in pooling the data for the two countries has little to offer to the explanation of the differences we have found between the two countries. At most the use of a wrong exchange rate could help explain the difference in average level; it would contribute nothing to explaining the difference in slopes, which is derived from data for each country separately. As a result we have relegated a full examination of the problem of the exchange rate to use in pooling the data to Appendix A of this chapter (sec. 6.8).

Summary of Results for Levels of Money and Income

The results for the levels of real money per capita are fairly straightforward. For the whole century we cover, and for each country separately, the elasticity of real money per capita with respect to real income per capita appears to be a constant, numerically greater than unity (about 1.2) for the United States, less than unity (about 0.81) for the United Kingdom. There appears to have been a single upward shift in liquidity preference of about the same size (30 percent) in both countries, connected with economic depression and its aftermath and World War II and its aftermath. We estimate that the shift spanned the years 1929 to 1954 in the United States and 1921 to 1955 in the United Kingdom. There appears also to have been a postwar readjustment to bring demand back to its desired level—after the particularly low level of money demand during World War I and the high level during World War II—that was also of the same size in both countries. The postwar adjustment was less important in both size and duration than the demand shift. Together, constant elasticity plus the shifts account for at least 99 percent of the variation in real per capita money for the United States, if real per capita income is taken as exogenous, and 98 percent for the United Kingdom; and for 99 and 98 percent for the United States and the United Kingdom, respectively, of the variation in real per capita income, if real per capita money is taken as exogenous. The maximum residual standard error is less than 6.5 percent for money for the United States, 6.3 percent for the United Kingdom, less than 5.4 percent for income for the United States, and 7.8 percent for the United Kingdom.

We have constructed a single equation for the two countries, constraining the elasticities to be equal. The calculated common elasticity is about 1.1, and the estimated level of real per capita money for given real capita income is about 11 to 13 percent higher in the United Kingdom than in the United States. However, these estimates deserve little confidence because of the difference between the two countries in elasticity, which means that for real incomes as low as those in the early part of the period, United Kingdom real per capita money for given real per capita income is something like 40 percent higher than in the United States, and for levels of real income as high as those at the end of the period, some 13 to 30 percent lower.

This difference in elasticity is the chief puzzle. Moreover, since errors in population and prices bias these estimates toward unity, unbiased estimates would be higher for the United States and lower for the United Kingdom, giving a still larger difference to explain.

A major task for later sections of this chapter is to see whether changes over time in the other variables that we explore can account for the upward shift in both countries, and whether differential changes over time in these other variables for the two countries can account for the difference in elasticity.

6.4.2 Rates of Change

The situation is somewhat different for rates of change of real per capita money and real per capita income, as is clear from chart 6.4, panels A and B, which gives scatter diagrams for rates of change like those in chart 6.1, panels A and B, for levels.

If we differentiate equation (7) with respect to time, we have

$$\frac{1}{m}\frac{dm}{dt} = \alpha\frac{1}{y}\frac{dy}{dt} \tag{14}$$

or

$$g_m = \alpha\, g_y\ , \tag{15}$$

where g designates the rate of change of the variable designated by the subscript.

A value of $\alpha = 1$ corresponds to a constant velocity. Hence, if velocity is constant,

$$g_m = g_y\ , \tag{16}$$

that is, money and income change at the same rate. The center line in each graph corresponds to equation (16). The other two lines are estimates of equation (15) obtained by regressing g_m on g_y and g_y on g_m, forcing the constant term to be zero. The computed values of α are given in table 6.7.

Clearly the scatters are very loose—more so for the United Kingdom than for the United States. The two limiting regressions are reasonably close together for the United States, a good deal farther apart for the United Kingdom.

This impression is confirmed by table 6.7. Columns 1, 2, and 3 are for equation (15). The correlations are clearly much lower than for the levels, but still most are appreciable. The limits on the income elasticity are much wider apart, but in general they include the narrower range in tables 6.4 and 6.6. However, much of the correlation, especially for the United Kingdom, comes from a common trend, as can be seen from columns 4, 5, and 6, which are for equation (15) modified by adding a constant term, that is,

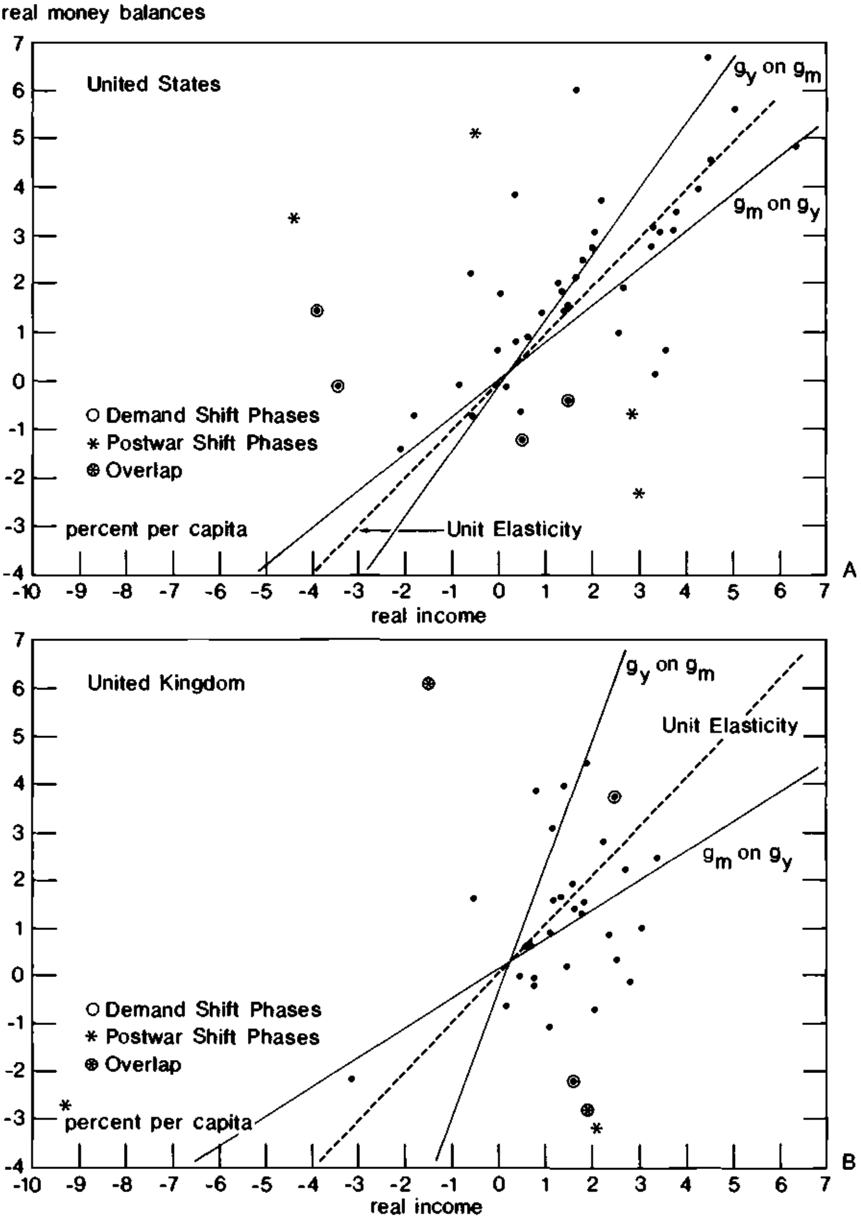

Chart 6.4 Scatter diagram of rates of change of real per capita income and real per capita money and regression lines.

Table 6.7 **Income Elasticities and Correlation Coefficients Calculated from Rates of Change of Real per Capita Money and Real per Capita Income, Various Periods, without and with Allowance for Trend by Inclusion of Nonzero Constant Term**

	Zero Constant Term, No Allowance for Trend			Nonzero Constant Term, Allowance for Trend		
	Income Elasticity			Income Elasticity		
	Lower Limit (1)	Upper Limit (2)	Correlation Coefficient (3)	Lower Limit (4)	Upper Limit (5)	Correlation Coefficient (6)
United States						
Full period	0.76	1.36	.751	0.48	1.60	.549
Full ex-wars	0.66	1.28	.721	0.41	1.18	.590
Pre–World War I	0.90	1.32	.826	0.37	1.44	.504
Interwar	0.30	1.49	.445	0.45	0.67	.822
Postwar	0.66	1.04	.798	1.02	2.10	.699
United Kingdom						
Full period	0.63	2.66	.487	0.14	23.42	.078
Full ex-wars	0.70	1.76	.633	0.50	7.51	.257
Pre–World War I	0.93	1.02	.955	1.19	1.51	.889
Interwar	1.15	1.59	.853	−0.09	−7.56	.110
Postwar	−0.05	−22.17	.049	−0.18	−78.74	.045
United States and United Kingdom[a]						
Full period	0.73	1.68	.657	0.47	2.80	.411
Full ex-wars	0.68	1.46	.682	0.42	2.38	.422
Pre-World War I	0.91	1.21	.868	0.70	1.61	.659
Interwar	0.57	1.55	.607	0.45	0.74	.782
Postwar	0.38	2.35	.400	0.87	6.07	.377

[a] Since basic observations are free of units, being rates of change, United States and United Kingdom observations were simply pooled for these regressions.

$$g_m = \kappa + \alpha g_y. \tag{17}$$

Equation (17) is obtained by adding the term κT, with T for time, to equation (7) and then differentiating with respect to time, which is why we speak of it as allowing for a trend. The correlations for the United States in this part of the table are still respectable, but for the United Kingdom they are close to zero except for the pre–World War I period.

To some extent these rather unsatisfactory results simply reflect the statistical characteristics of our data. Our rates of change are comparable to first differences between equally spaced data, differing only in being based on observations that are unequally spaced and on three rather than two observations. Correlations between first differences typically tend to be much lower than between original values because the random component plays so much larger a role compared with the systematic component. Similarly, we would expect the correlations between the rates of change to be much lower than between the levels, especially given the strong secular trends in our variables. On this ground the lower but still respectable correlations for the United States are fully in accord with expectations. However, the sharp drop in the United Kingdom correlations, to a level that does not differ from zero by an amount that is statistically significant, is disturbing.

The contrast between the prewar period for the United Kingdom and the other periods gives a clue to the results for the latter. For both the United States and the United Kingdom, the results for periods other than the prewar period are distorted by two phenomena: (1) immediate postwar perturbations, which account for the most extreme observations for both the United States and the United Kingdom, and (2) the upward shift in liquidity preference discussed in the preceding section.

1. The points in chart 6.4, panels A and B marked * are the relevant postwar phases centered on 1920 and 1921 for both countries after World War I; on 1949 and 1951.5 for the United States; and 1952 and 1954 for the United Kingdom after World War II. In both countries the two sets of postwar points are extreme, though in opposite directions: the rate of growth of real per capita money is unusually high after World War I and unusually low after World War II. The opposite directions are the counterpart of the difference between the two wars that we discussed earlier (sec. 5.2): during World War I velocity rose, so after the war real money balances rose sharply to restore them to their usual relation to real income; during World War II velocity fell, so after the war real money balances had to fall sharply to restore the usual relation.

Two other features stand out. First, the deviations in the opposite directions from the central line are of roughly the same magnitude both for the two wars and for the two countries. At the risk of reading more into the data than may be there, this curious and, at least to us, unexpected result suggests that there may be a fairly constant maximum *rate* at which money holders are willing, under the kind of conditions prevailing at the end of the two wars, to readjust their money holdings from one desired *level* to another. Second, the postwar readjustment comes at the same time in the two countries after World War I but later in the United Kingdom after World War II. This again is a familiar phenomenon that we have referred to before.

2. Because our rates of change are calculated from three phase levels, the upward shift in liquidity preference isolated in our examination of levels affects two rates of change at the onset and two at its termination. These observations are circled in chart 6.4. For the United States, the circled observations do not overlap the ones marked *. For the United Kingdom, because the upward shift came earlier and the post-World War II readjustment later, there is an overlap, so that two observations are both *'s and circled.

The upward shift tends to make the rate of change of real per capita money unusually high when it began and unusually low when it ended. Seven out of eight circled points clearly conform to that expectation; the eighth, the United Kingdom one for 1923, is somewhat ambiguous, falling between the two regressions.

To estimate the quantitative effect of allowing for these two phenomena, we have added two dummy variables to equations (15) and (17), so that they become

$$g_m = \alpha g_y + \lambda_1 W_g + \lambda_2 S_g \tag{18}$$

$$g_m = \kappa + \alpha g_y + \lambda_1 W_g + \lambda_2 S_g \, , \tag{19}$$

$$\text{where } W_g = \begin{cases} +1 & \text{for the phases centered on 1920 and 1921 for the United States and the United Kingdom} \\ -1 & \text{for the phases centered on 1949 and 1951.5 for the United States and 1952 and 1954 for the United Kingdom} \\ 0 & \text{for all other phases} \end{cases}$$

$$S_g = \begin{cases} +T_{g_i} & \text{for the phases centered on 1928.5 and 1931 for the United States and 1921 and 1923 for the United Kingdom} \\ -T_{g_i} & \text{for the phases centered on 1954 and 1956 for the United States and 1954 and 1957 for the United Kingdom} \\ 0 & \text{for all other phases} \end{cases}$$

where T_{g_i} is the time interval between the centered dates of phase $i + 1$ and phase $i - 1$. Note that W_g and S_g and T_{g_i} are equivalent to the time derivative of the corresponding variables, W, S, and T_i for levels.

The first dummy variable, W_g, allows for the postwar readjustment, the second, S_g, for the upward shift. This way of allowing for these phenomena constrains the postwar readjustment to have the same effect on the rate of change for each of the four phases involved, and the upward demand shift to have the same effect on the product of the rate of change and time duration for each of the four phases involved. For W_g, that constraint reflects the hypothesis that there is a maximum per-year desirable rate of readjustment of deficient or excessive money balances; for S_g it reflects our earlier conclusion that the level figures could be approximated by two parallel constant elasticity functions. For each of the four rate of change observations affected by the shift, the size of the step is the same, but since the period over which the adjustment occurs depends on the interval between the preceding and following phases, that implies a constant product of time interval and change in the rate of change.

The effect of introducing these dummy variables is summarized in table 6.8. These results are much more satisfactory than those in table 6.7. Though the correlations between rates of change after allowing for shifts are understandably still much lower than between levels, they are all positive and most differ significantly from zero. More important, the ranges of the income elasticities are not inconsistent with the ranges derived from the levels. The postwar shift plays a much larger role, and the demand shift a smaller role, for rates of change than for levels, but again the size of the shifts is about the same for the United States and the United Kingdom. The one important difference from the levels is that there is no clear difference between United Kingdom and United States income elasticities—but that may reflect simply the much lower correlation for rates of change and hence the wider limits on the estimated elasticity.

As for levels, the inclusion of the shift variables seems to make the data homogeneous for the different periods. There is no significant difference between the computed regressions for the pre-World War I period, for which dummies are irrelevant, and the rest of the period, if dummies are included.[30] Neither is there a significant difference between the United States and the United Kingdom regressions once dummies are included.[31]

It is hard to compare directly the variability of the rates of change and the levels because the standard deviations are in different units. For the level of United States real per capita money for the period as a whole, for example, the standard deviation is 62.5 percent. This means that roughly

---

[30] The F values between the periods for regressions with a zero constant term of real per capita money on real per capita income and dummies are 1.3 for the United States (1 and 43 degrees of freedom); .01 for the United Kingdom (1 and 31 degrees of freedom); and .32 for the United States plus the United Kingdom (1 and 79 degrees of freedom). The results are similar for regressions including other variables, and for regressions with a constant term.

[31] The relevant analysis of variance is as shown in table 6.N.2 on p. 252.

Table 6.8 **Relation between Rates of Change of Real per Capita Money and Real per Capita Income: Effect of Allowing for Demand and Postwar Shifts, with Zero and Nonzero Constant Term**

	Zero Constant Term			Nonzero Constant Term		
	U.S.	U.K.	U.S. and U.K.	US.	U.K.	U.S. and U.K.
1a. Period affected by demand shift	1927–32	1920–24		1927–32	1920–24	
	1953–57	1952–58		1953–57	1952–58	
b. Period affected by postwar shift	1919–21	1919–21		1919–21	1919–21	
	1948–53	1951–55		1948–53	1951–55	
2a. Size of demand shift (annual percentage)	0.51(0.68)	0.45(0.48)	0.48(0.65)	0.28(0.81)	0.48(0.31)	0.38(0.84)
b. Size of postwar shift (annual percentage)	4.59(5.59)	4.18(6.33)	4.32(5.39)	4.62(5.66)	3.76(10.72)	4.40(5.45)
3a. Income elasticity, lower limit	0.91	0.85	0.89	0.64	0.47	0.68
b. Income elasticity, upper limit	1.27	1.84	1.42	1.42	5.23	1.83
4a. Simple correlation, money and income	0.75	0.49	0.66	0.55	0.08	0.41
b. Partial correlation, allowing for shifts	0.85	0.68	0.79	0.67	0.30	0.61
5. Multiple correlation, including shifts	0.85(0.87)	0.74(0.72)	0.81(0.82)	0.73(0.78)	0.64(0.43)	0.70(0.68)

	Standard Error of Estimate (%)					
Real per capita money						
6. Total	3.31	2.08	2.61	2.25	1.75	2.04
Maximum residual after allowing for:						
7. Constant velocity	2.32	1.91	2.07	2.31	1.93	2.08
8. Constant elasticity	2.21	1.84	1.98	1.90	1.78	1.88
9. Constant velocity plus shifts	2.20	1.78	2.00	2.20	1.78	2.00
10. Constant elasticity plus shifts	1.80	1.46	1.57	1.59	1.42	1.49
Real per capita income						
11. Total	3.25	1.60	2.36	2.57	0.96	1.78
Maximum residual after allowing for:						
12. Constant velocity	2.32	1.91	2.07	2.31	1.93	2.08
13. Constant elasticity	2.17	1.42	1.79	2.17	0.97	1.63
14. Constant velocity plus shifts	2.20	1.78	2.00	2.20	1.78	2.00
15. Constant elasticity plus shifts	1.67	1.17	1.40	1.68	0.90	1.33

Note: Figures in parentheses from regressions of g_y on g_m and dummy or dummies solved for g_m; others from regressions of g_m on g_y and dummy (or dummies).

two-thirds of the observations are between 100/1.625, or 61.5 percent of the mean, and 162.5 percent of the mean. For rates of change there is little point in expressing the variation relative to the mean, since the mean may be zero or negative. The comparable rate of change standard deviation is 2.25 percentage points per year, which means that two-thirds of the annual rates of changes are within 2.25 percentage points of the mean annual rate of change. Perhaps one way to make the comparison is to note that the mean interval spanned by the three phases from which rates of change are calculated is four years for the United States, five and one-half years for the United Kingdom, so the standard deviation of the total percentage change in money over this interval is about 9 percent of the initial value for the United States, only a seventh of the 62.5 percent standard deviation of levels. For the United Kingdom, the corresponding figures are 32.6 and 1.75 percentage points per year, or about 10 percent for a five and one-half year span, less than a third of the standard deviation of levels. Clearly, eliminating the longer period movements by the use of rates of change has also eliminated most of the variability.

As a result, allowing for the effect of real per capita income on real per capita money (or of real per capita money on real per capita income) and of the shifts has a much less dramatic effect on the standard deviations, reducing them, if trend is allowed for by including a constant term, from 2.25 percentage points per year to 1.59 for the United States, and from

Table 6.N.2 **Analysis of Variance of Differences between Demand for Money Regressions for the United States and the United Kingdom Based on Rates of Change**

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Within United States	44	.01428	.00032(.00030)[a]	
Within United Kingdom	32	.00684	.00021(.00023)[a]	
Within United States and United Kingdom	76	.01937	.00025	
Difference between United States and United Kingdom				
In slope	1	.00005	.00005	0.20
In shifts	2	.00001	.000005	0.02
Subtotal	79	.01943		
Attributable to				
Demand shift	1	.00546	.00546	21.84
Postwar shift	1	.00691	.00691	27.64
Slope	1	.02415	.02415	96.60
Total	82	.05595		

[a] See note 24, above.

1.75 to 1.42 for the United Kingdom (and for real per capita income from 2.57 to 1.68 for the United States, and .96 to .90 for the United Kingdom). Comparison of these reduced standard errors, multiplied by the average span on which they are based, with the residual standard errors of the levels, after allowing for real per capita income and the shifts, is much more illuminating than comparing the original standard deviations. For the United States the residual standard error of the level of real per capita money is 6.5 percent; four times the residual standard error of the rate of change is 6.4. For the United Kingdom, the residual standard error of the level is 6.3 percent; five and one-half times the residual standard error of the rate of change is 7.8 percent. The final results are therefore roughly the same: real per capita income and the shifts leave about the same residual error as the more sensitive adjustment for trend in the form of rates of change.

What these results mean is that the relatively small short-term movements of real money balances reflect primarily variables other than the contemporaneous changes in real per capita income or the shifts we have isolated. We have reached this conclusion before for movements within cycle phases. It is the basis for our suggestion that the demand for real money balances depends on permanent rather than measured income.[32] The present results indicate that a similar, but weaker, conclusion holds for the longer periods to which our rates of change refer—averaging four years for the United States, five and one-half for the United Kingdom. Over such periods, the variation in both real balances and real income is small, and each is heavily affected by variables that do not affect the other in the same way. Over longer periods, the changes in real money balances and real per capita income cumulate and become more and more important relative to the short-term perturbations.

6.5 Effect of Population and Prices

In the preceding section we assumed that $\beta = \gamma = 1$ in equation (3), that is, that any change in prices or population implied an equal percentage change in the quantity of money demanded (mathematically, the demand for money in nominal terms is homogeneous of the first degree in population and prices). This section tests that assumption. Take logarithms of both sides of equation (3), add terms to allow for the demand and postwar shifts, and rearrange to:

$$\log m = \log k + \alpha \log y + (\beta - 1) \log N + (\gamma - 1) \log P + \lambda_1 W + \lambda_2 S. \tag{20}$$

---

[32] M. Friedman, *The Demand for Money: Some Theoretical and Empirical Results*, Occasional Paper 68 (New York: NBER, 1959); reprinted in Friedman, *Optimum Quantity of Money;* Friedman and Schwartz, "Money and Business Cycles," p. 44.

Alternately, solve equation (20) for log y, giving:

$$\text{(21)} \qquad \log y = -(1/\alpha)\log k + (1/\alpha)\log m - [(\beta - 1)/\alpha]\log N - [(\gamma - 1)/\alpha]\log P - [\lambda_1/\alpha]W - [\lambda_2/\alpha]S.$$

Clearly, if $\beta = \gamma = 1$, the coefficients of log N and log P are zero in both equations (20) and (21).

The values of k, α, β, and γ obtained by fitting equations (20) and (21) to the data for the United States and the United Kingdom are summarized in table 6.9.[33] For both countries, the calculated price elasticities as computed from the money on income and income on money regressions are consistent with the theoretical value of unity—one estimate exceeding, the other falling short of unity. For the United Kingdom, that is true also of the population elasticity, but it is not for the United States, for which both limits exceed unity. For the United States plus the United Kingdom, the range of both population and price elasticities excludes unity, but the limit for each is clearly not significantly different from unity. All in all, therefore, the only serious discrepancy from theoretical expectations is for the United States population elasticities.

Columns 5 and 6 provide an additional test by comparing errors of estimate of equations omitting population and prices (i.e., forcing $\beta = \gamma = 1$) and equations including them. For four of the six comparisons, the inclusion of population and prices explicitly has a sizable effect, reducing the standard error substantially. There is one exception for each country: the income on money regression for the United States, the money on income regression for the United Kingdom.

One possible explanation for these results is that population and prices are serving as proxies for other variables. For prices, the other variable may be interest rates, since there is the well-known Gibson paradox phenomenon of a positive correlation between prices and interest rates. For population, the obvious other variable is time trend, since we know that population moves very slowly over time.

The effect of including population and prices on the value of α is consistent with the interpretation of population as a proxy for trend. The lower and upper limits of α (from the regressions of money on income and income on money, respectively) are far wider apart in table 6.9 than in table 6.6, indicating that inclusion of population and prices has reduced the correlation between real money per capita and real income per capita, both of which display consistent upward trends. Some other variable, presumably population, is clearly serving as a trend term.

---

[33] Note that because P and N enter explicitly in these equations, the statistical estimates of α are not subject to the bias arising from measurement errors in N and P discussed above.

Table 6.9 **Relation between Levels of Real per Capita Money and Real per Capita Income: Estimates of the Effects of Population and Prices, Allowing for Demand and Postwar Shifts**

					Standard Error of Estimate (Percentage) of per Capita Magnitude from Regressions among Real per Capita Money, Real per Capita Income, Shift Dummies, and			
Country and Regression	Constant (1)	Income Elasticity (α) (2)	Population Elasticity (β) (3)	Price Elasticity (γ) (4)	No Other Variables (5)	Population, Prices (6)	Interest Rate, Trend (7)	Interest Rate, Lagged per Capita Magnitude (8)
United States								
Money on income	−0.96	0.68	1.49	1.07	6.5	4.6	4.5	3.9
Income on money	−2.33	1.13	1.17	0.94	5.4	5.2	5.3	5.2
United Kingdom								
Money on income	−0.16	0.65	1.29	1.00	6.3	6.4	5.4	5.3
Income on money	0.70	1.62	−0.10	0.87	7.8	6.2	5.3	4.1
United States and United Kingdom[a]								
Money on income	−1.37	0.60	1.69	0.96	10.1	6.2	8.4	5.3
Income on money	3.66	1.45	1.01	0.75	9.2	6.6	5.6	6.1

[a] For United States and United Kingdom combined, country dummy (U.S. = 0, U.K. = 1) was also added to equation.

The final two columns in table 6.9 test this explanation by substituting a short-term interest rate for the price variable[34] and by allowing for trend instead of population—in column 7 by including time as a variable, in column 8 by including the lagged value of the dependent variable. These substitutions give lower standard errors of estimate in ten out of twelve comparisons, and in some comparisons decidedly lower standard errors.[35]

The equations including the interest rate term and trend (corresponding to column 7 of table 6.9) go even further than the equations including population and prices in the direction of absorbing some or most of the relation between money and income. For the United States and the United Kingdom, the effect is to widen the range between the lower and upper limits of the computed income elasticity—from .63 and 1.13 to .61 and 1.17 for the United States, from .65 and 1.62 to .42 and 2.54 for the United Kingdom. For the United States plus the United Kingdom, the effect, surprisingly, is the reverse, to narrow the range, from .60 and 1.45 to 1.23 and 1.86.

Parallel calculations for rates of change give very similar results. Differentiating equations (20) and (21) with respect to time gives:

$$g_m = \alpha\, g_y + (\beta - 1)g_N + (\gamma - 1)g_P + \lambda_1 W_g + \lambda_2 S_g \tag{22}$$

$$\begin{aligned} g_y = {} & [1/\alpha]g_m - [(\beta - 1)/\alpha]g_N - [(\gamma - 1)/\alpha]g_P \\ & - [\lambda_1/\alpha]W_g - [\lambda_2/\alpha]S_g. \end{aligned} \tag{23}$$

Table 6.10, the counterpart for rates of change of table 6.9, gives almost identical results. The computed population and price elasticities are in general, though not uniformly, consistent with the hypothetical values of unity; the inclusion of population and prices explicitly has a trivial effect on the standard error of estimate for three comparisons, reduces it moderately for three; in the latter cases the same or an even greater reduction in the standard error is achieved by including an interest rate and a trend term instead of population and prices.[36]

As for levels, the inclusion of population and prices absorbs some of the effects of the correlation between money and income, widening the range of the estimated income elasticities (compare table 6.10 and table 6.8, zero-constant term). Replacing population and prices by interest rate

---

[34] Tests summarized in section 6.6.1 indicate that a short-term rate generally gives better results than a long-term rate. However, the interest rate used here is not the differential interest rate we finally settled on in section 6.6.

[35] We have also computed regressions including population and prices along with interest rates and a trend term. In general, these give standard errors of about the same size as those in columns 7 and 8. However, in a few cases the addition of population and prices does reduce the standard error appreciably, indicating that they are not simply proxies for interest rates and time.

[36] Trend is allowed for simply by introducing a constant term into the regression. Equations (22) and (23) do not have a constant term.

Table 6.10 **Relation between Rates of Change of Real per Capita Money and Real per Capita Income: Estimates of the Effects of Population and Prices, Allowing for Demand and Postwar Shifts**

	Estimates of			Standard Error of Estimates (Percentage) of per Capita Real Magnitude from Regressions among Real per Capita Money, Real per Capita Income and			
Country and Regression	Income Elasticity (α) (1)	Population Elasticity (β) (2)	Price Elasticity (γ) (3)	No Other Variables (4)	Population, Prices (5)	Interest Rate, Trend (6)	Interest Rate, Lagged per Capita Magnitude (7)
United States							
Money on income	0.64	1.64	1.11	1.80	1.59	1.61	1.00
Income on money	1.37	0.95	0.90	1.67	1.69	1.69	1.52
United Kingdom							
Money on income	0.69	1.36	1.06	1.46	1.47	1.35	1.33
Income on money	3.49	−3.16	1.03	1.17	0.95	0.92	0.96
United States and United Kingdom[a]							
Money on income	0.67	1.57	1.08	1.57	1.48	1.47	1.23
Income on money	1.75	0.29	0.95	1.40	1.36	1.33	1.41

[a] For United States and United Kingdom combined, country dummy (U.S. = 0, U. K. = 1) was multiplied by demand shift and postwar shift dummies and added to equations.

and trend widens them further, modestly for the United States and United States plus the United Kingdom, substantially for the United Kingdom.[37]

The results for both levels and rates of change are reasonably consistent with the assumption $\beta = \gamma = 1$. Any apparent influence of population and prices can readily be interpreted as a disguised reflection of the influence of trend and interest rates.

The most interesting substantive result from our explorations is the effect on the computed elasticities of allowing explicitly for trend. For both the United States and the United Kingdom, introducing trend absorbs much of the effect of income and widens the range of the estimates of income elasticity. This result admits of two very different interpretations. One is that there is a secular trend in the quantity of money demanded produced by some secularly changing variables other than income itself—for example, urbanization, or increasing specialization in production—for which income is serving as a proxy and for which a trend term is a better proxy. This explanation is certainly plausible.

An alternative interpretation is that the quantity of money demanded is related to a longer term income magnitude than average income over a phase, and that the trend term itself is serving as a proxy for such a longer term magnitude. Numerous demand studies for money have shown that some concept of permanent real income or wealth is more closely connected with the real quantity of money demanded than is current income. However, these studies have mostly been for quarterly or annual data. We conjectured initially that our phase data would be closely enough linked to permanent income to make the distinction of little empirical significance. But perhaps that is wrong.

Column 8 of table 6.9 and column 7 of table 6.10 were included as a simple way to cast some light on this interpretation. Allowing for trend in a regression of money on income by including the lagged value of the quantity of money is equivalent to estimating permanent income as a weighted average of current and prior values of income with the weights given by an exponential whose value declines the earlier the income value weighted—this is the adaptive expectations approach—and hence to regarding the quantity of money demanded as a function of permanent income. Similarly, in a regression of income on money, including the lagged value of income is equivalent to regarding income as a function of "permanent money" estimated by the same kind of exponentially weighted average of present and past values of money.

The results are reasonably consistent with this interpretation. For the money on income regressions and for both levels and rates of change, the

---

[37] For the United States, to 0.64 and 1.42; for the United Kingdom, to 0.40 and 5.40; for the United States and United Kingdom, to 0.68 and 1.79.

inclusion of lagged income consistently gives a lower standard error than the use of the trend term. For the income on money regression, this is true for only three out of six comparisons, and in general any effect is smaller in magnitude—which is certainly what would be expected if permanent income systematically affects desired money balances. It is difficult to give any persuasive interpretation to "permanent money" as determining real income, which is why the asymmetry of the results supports the alternative interpretation. Moreover, the regression coefficients are not unreasonable. They imply that the mean period of averaging in computing permanent income is roughly two or three phases, and that the elasticity of money demand with respect to permanent income is roughly the same as the lower limits computed more directly, allowing only for the shifts, as given in tables 6.6 and 6.8.[38]

6.6 Effect of Costs of Holding Money

The quantity of money that people wish to hold is affected by the costs and returns from holding money. The direct costs and returns—in the form of expected losses, storage charges, interest received—are generally neglected and treated as zero. However, Benjamin Klein has demonstrated that, at least for the United States, allowance for these direct costs and returns can appreciably improve the statistical estimates of demand

---

[38] The estimated weight on the current phase value of income in computing permanent income (b), average period of weighting, in phase units (1/b), and estimated permanent income elasticity (dm/dy_p) are as follows for the United States and the United Kingdom, for levels and rates of change:

	Levels			Rates of Change		
	b	$1/b$	dm/dy_p	b	$1/b$	dm/dy_p
United States	.55	1.8	1.12	.40	2.5	1.55
United Kingdom	.62	1.6	0.79	.60	1.7	0.78

These estimates were derived from the coefficients of the following regressions:

$$m_t = \log k + \alpha y_t + \delta R_S + \Theta m_{t-1} + \lambda_1 W + \lambda_2 S,$$

$$g_{m_t} = \log k + \alpha g_{y_t} + \delta\, g_{R_S} + \Theta g_{m_{t-1}} + \lambda_1 W_g + \lambda_2 S_g .$$

We assumed that permanent income for the United States is estimated by adding a trend element of .02 per year, which for phases averaging two years in length means .04 per phase, to a weighted average of current and past measured incomes. On this assumption

$$b = 1 - \Theta + .04$$

$$\frac{dm}{dy_p} = \frac{\alpha}{b} .$$

For the United Kingdom, we assumed a trend element of .0125 per year, which multiplied by the 2.75 year phase average length of the United Kingdom phase, gives .034 as the trend allowance.

curves.[39] Klein obtains a statistical counterpart to these direct costs and returns by neglecting the direct costs and returns associated with currency and with losses from bank failures but including an allowance for interest paid on deposits, whether interest is paid directly, as was legal for all deposits in the United States before 1933 and for time deposits thereafter (and throughout in the United Kingdom),[40] or indirectly through rendering services without charge or granting loans at lower than market interest rates. He allows for both direct and indirect payments by assuming that banks pass on to their customers the net income they receive from their earning assets and estimating this net income on the basis of market interest rates. This procedure is equivalent to including the ratio of high-powered to total money as well as a market interest rate in the demand for money function.[41] We have experimented with this method of allowing for the direct costs and returns from holding money and report the results below.

The major return from holding money is indirect: the nonpecuniary services rendered by money. The adjustment we have made for increasing financial sophistication of the United States before 1903 already allows for one element of these services. For the rest, we treat this return

---

[39] Benjamin Klein, "The Payment of Interest on Commercial Bank Deposits and the Price of Money: A Study of the Demand for Money" (Ph.D. diss., University of Chicago, 1970).

[40] However, cartel agreements among the banks, dating from 1877 and effective from World War II on, prohibited direct interest payments on United Kingdom current accounts. In 1877 London joint stock banks agreed to abolish all interest payments on current accounts, although there is evidence that many banks evaded the agreement before World War I. Certain provincial banks did not end the practice of paying interest on current accounts until World War II, when all banks agreed not to compete for deposits by increasing interest payments. Cartel agreements among the banks dating to 1886 determined the rate of interest to be paid on deposit accounts by reference to the margin between bank rate and the deposit rate, but again there were many exceptions. See C. A. E. Goodhart, *The Business of Banking, 1891–1914* (London: Weidenfeld and Nicolson, 1972), pp. 178–88; B. Griffiths, "The Development of Restrictive Practices in the U.K. Monetary System," *Manchester School of Economics and Social Studies* 41 (March 1973): 6–7.

[41] Let $H = C + B$ be high-powered money, with C = currency, and B = bank reserves defined as the banks' total holdings of high-powered money. Let R = a market interest rate on securities comparable to bank earning assets in period to maturity. Then Klein treats interest paid per dollar of deposits as equal to $R(1 - B/D)$, where D equals total bank deposits. This includes interest paid indirectly on demand deposits plus interest paid directly on time deposits. Income earned by a holder of money per dollar would then be a weighted average of the zero earned per dollar of currency and the above sum earned per dollar of deposits, or

$$\frac{C(0) + DR(1 - \frac{B}{D})}{C + D} = \frac{R(D - B)}{C + D} = \frac{R(M - H)}{M} = R(1 - \frac{H}{M}),$$

where $M = C + D$ = the money supply.

as embodied in the demand function, which means as primarily a function of the real quantity of money held. This procedure is incorrect insofar as technical changes have altered the productivity of money balances in yielding monetary services—as apparently they did to a marked extent in the United States before 1903 and must have to some extent in the United States since then and in the United Kingdom before and since.

The major indirect cost of holding money is the income foregone from the assets that could have been held instead of the money. The opportunity cost depends on the alternative considered and on the period of time for which the alternative is regarded as sacrificed. In addition, we must distinguish between the alternative cost ex post—the realized return on the alternative asset—and the alternative cost ex ante—the anticipated return. We can only measure the cost directly ex post, yet what is relevant to the demand function is the alternative cost ex ante.

As we noted in chapter 2, possible alternative assets are numerous for each holder of money and vary widely from holder to holder. For business enterprises, the key alternatives to holding money are larger holdings of financial and physical assets or reduced liabilities in the form of borrowing from banks, other short-term borrowing, long-term debt, and equity. For ultimate owners of wealth, and for our concept of money, the major alternatives include mutual savings deposits, savings and loan shares, government and private debt obligations, equities, and physical assets, or reduced liabilities in the form of mortgages on owned homes and debt on other consumer goods.

For each alternative, the ex post return in nominal terms that is sacrificed for any given period includes the explicit yield in nominal terms during that period plus the change in the nominal price of the asset during that period. For some alternatives, for example borrowing from banks (neglecting the possibility of default), mutual savings deposits, savings and loan shares, and series E United States government bonds, the change in price is zero because the obligation can be paid off or redeemed on demand at a fixed nominal price. For these alternatives also, the yield is generally known in advance so that there is no significant difference between the nominal return ex ante and ex post—though of course there may still be a difference between the real return ex ante and ex post. For marketable obligations, such as commercial paper, bonds, and particularly equities, a change in the nominal price of the asset may be an important component of the ex post return. For some obligations that have a stated maturity value such as commercial paper and bonds, the change in price will itself consist of two parts, one that results from the amortization of any premium over, or discount from, the value at maturity, the other, from the change in market yields. The first is embedded in calculated market yields and hence can be regarded as known ex ante.

The second cannot be known in advance and the ex ante equivalent to this part of the ex post return is an *anticipated* change in market yields.

For equities, which have no stated maturity value, and for physical assets such as fixed capital and inventories for business enterprises, or owned homes, automobiles and other consumer capital for ultimate wealth owners, the change in market price tends to be a still larger part of the nominal return and no part of the change is known ex ante. Almost the whole of the return on these assets must be based on anticipations.

The period of time for which the alternative is regarded as sacrificed and therefore for which the ex ante yield is relevant is an elusive concept. Money balances are held for a wide variety of possible contingencies, the timing of some of which, such as recurrent trips to market, is reasonably predictable, the timing of others, such as emergency needs for ready funds, is highly uncertain. In principle, the whole term structure of yields, for all possible holding periods, is relevant to the quantity of money demanded. For example, no one would hold money instead of an interest-bearing asset if he were *certain* that he would not have to draw on it for a very long time. Yet he may hold money even though there is a sizable possibility that he may not have to draw on it for a very long time. Whether he will do so depends on the return available from assets held for long periods.[42]

Given the multiplicity and complexity of the yields that are relevant, it is perhaps not surprising that statistical studies of the demand for money, which have had to use a small number of observable yields, have produced divergent and confusing results. Almost all such studies confirm the expected negative relation between yields on alternative assets and the quantity of money demanded. However, studies for different countries and periods, and even studies by different scholars for the same country and period, have yielded widely divergent conclusions about the particular yield that is most closely related to the quantity of money demanded. The studies have explored (1) short-term yields on nominal assets; (2) long-term yields on nominal assets; (3) yields on equities; (4) anticipated rates of price change, computed as weighted averages of past rates of price change and interpreted as the nominal yield on physical assets. And these four classes of yields have of necessity covered only a small part of the range of assets that are effective alternatives to the holding of money. For most such assets, there simply are no market yields to observe.

Item 4, anticipated price change, has generally been dominant whenever there has been substantial inflation, perhaps partly because governmental intervention has often made it impossible to observe the other

---

[42] See Milton Friedman, "Time Perspective in the Demand for Money," *Scandinavian Journal of Economics* 79, no. 4 (1977): 397–416.

yields. For other countries and times, the effect of anticipated price change has been hard to detect, either because our measures are defective or because the bulk of its effect is allowed for by item 1, the yield on nominal assets, which tends to be raised by expectations of inflation. Item 3, yield on equities, is hard to measure and hence has been explored in only a few studies.

For countries and periods not characterized by substantial inflation, items 1 and 2, short- and long-term yields on nominal assets, have been the yields generally explored. These studies show no uniformity. Sometimes a short-term yield gives better results, sometimes a long-term yield, and when both are included the relative coefficients are unstable. This result is not inconsistent with theoretical reasoning, which suggests that the relative weight of yields for different holding periods should itself depend on the level and term structure of yields.

The consistent finding that the yield displays the expected negative relation with the quantity of money demanded presumably reflects the existence of a structure of yields that tends to move more or less together. If this is so, then the yield that is included in a demand study, or the small number of yields that are included, enter as a proxy or proxies for the whole structure of yields rather than in their own right. On this view, the particular yield that shows the closest relation depends on the yield that happens, in the particular circumstances, to be the best proxy for the yield structure as a whole.

We conclude that it would be desirable for the focus of research to change, that we should give up the attempt to find an asset or a small set of assets that can be regarded as the closest substitutes for money and instead recognize that money is so pervasive and the range of substitutes so broad that we should seek rather to find a compact way to describe the whole structure of yields—the "general" level; the "tilt" of the yield structure to maturity; and the "difference" between real and nominal yields—to suggest what seem offhand the key characteristics of that structure.[43]

---

[43] H. Robert Heller and Mohsin S. Khan, "The Demand for Money and the Term Structure of Interest Rates," *Journal of Political Economy* 87 (February 1979): 109–29, have taken up this suggestion and applied it to United States quarterly data for 1960–1976. They approximate the term structure by a quadratic in maturity, then use the three parameters of the quadratic in a demand for money function, concluding that "this approximation performed favorably relative to standard specifications of the money-demand function that utilized only one interest rate as the opportunity-cost variable as well as the ones that introduce several interest rates. Furthermore, . . . the function using this particular approximation appeared to be stable during a period when standard functions using only one interest rate display significant shifts in parameters" (p. 127).

We have ourselves not been able to use this approach for the United States and the United Kingdom because of the large research expenditure that would have been required to get

6.6.1 Yield on Nominal Assets

Our own explorations of different nominal yields have simply confirmed the results of other studies.

In our calculations, we have generally used two interest rates for each country: one a rate on a security with a short maturity, the other a rate on a security with a long maturity.[44] For levels, we have fitted equations of the form:

(24) $\log m = \log k + \delta_S R_S + \delta_L R_L +$ terms in other variables,

where m, as earlier, is real per capita money balances (for the United States, adjusted for changing financial sophistication), all logarithms are natural logarithms, R_S is a short-term rate of interest, R_L a long-term rate, and k, δ_S, δ_L and similar coefficients of other terms are parameters to be estimated. Some equations include only R_S, others, only R_L; still others, both.

For rates of change, we have fitted equations of the form:

(25) $g_m = \kappa + \delta_S DR_S + \delta_L DR_L +$ terms in other variables,

where D is used as a symbol for a time derivative or absolute difference (as compared with the symbol g that we have used for a percentage rate of change or time derivative of a logarithm). Some equations include an intercept (κ)—which is equivalent to allowing for a constant time trend in

---

satisfactory term structure data for the United Kingdom. We report in an appendix to this chapter (see 6.9) some experiments we have made for the United States.

Perhaps the most interesting and exciting empirical implication of their and our calculations is the evidence they provide on a conclusion reached on purely theoretical grounds in Friedman, "Time Perspective in the Demand for Money" (see note 42), that a steepening in tilt of the term structure, that is, a rise in long-term rates accompanied by a decline in short-term rates sufficient to keep the average level of interest rates constant, will reduce the quantity of money demanded. Both bodies of data confirm that conclusion, though we note that the relevant evidence from the Heller-Khan data is not that which they themselves adduce.

[44] In principle, holding period, not maturity, is the time-duration concept relevant to the demand for money. However, maturity is the only readily available proxy for holding period. The interest rates used are:

United States short rate: sixty-to-ninety-day commercial paper rate through 1923; thereafter four-to-six-month commercial paper rate.

United States long rate: high-grade corporate bond yield.

United Kingdom short rate: three month rate on bankers bills.

United Kingdom long rate: yield on consols.

Tables 5.7 and 5.8 give phase averages and rates of change for an alternative United States short rate, the call money rate ($R_{S'}$), and an alternative United States long rate, the basic yield on corporate bonds ($R_{L'}$). Our decision to use the ones listed above instead was based on a large number of trial regressions. In general the rates we use gave higher correlations than the alternatives, but the difference in results was often small and sometimes in the opposite direction.

the level variables, that is, to using the time differential of equation (24) when one of the other variables is time; other equations do not include an intercept, that is, force $\kappa = 0$, which makes equation (25) equivalent to the time differential of equation (24), when time is not included among the other variables. As for levels, some equations include only R_S; others, only R_L; still others, both.

The inclusion of interest rates rather than their logarithms in equations (24) and (25) implicitly assumes that the absolute rather than percentage change in interest rates is what matters for the demand for money—that a one-percentage point higher interest rate will produce the same percentage reduction in the quantity of money demanded whether it is added to a base rate of 5 percent or to a base rate of 10 percent. On this assumption, a semilogarithmic slope, like δ_S or δ_L, is the relevant measure of the interest rate effect, even though this parameter is not free from units of measure but has the dimension of time.

An alternative would be to include the logarithms of interest rates rather than the interest rates themselves. The corresponding regression coefficients would be interest elasticities, free from units of measure. This would assume that elasticities rather than slopes are the same at all levels of interest rates.

Theoretical considerations in favor of a constant slope rather than elasticity are: (1) per dollar of money held, the cost of a change in the interest rate depends on the absolute, not the relative change in the interest rate; to put this point differently, a doubling of an interest rate of 1 percent is much less of a stimulus to reduce cash balances than a doubling of an interest rate of 10 percent; (2) an interest rate of zero does not imply infinite desired cash balances, yet a constant elasticity other than zero produces infinite cash balances at an interest rate of zero; (3) an expected rate of price change is a measure of cost logically similar to a nominal interest rate, and it seems desirable to treat the two cost measures in the same way, yet the expected rate of price change can be negative, ruling out logarithms.

One consideration against a constant slope is that, at high interest rates, desired cash balances will be small and will be used for high-priority purposes so that a rise of one percentage point is likely to produce a smaller percentage reduction in money balances than a similar rise at lower levels of interest rates. However, that is not an argument for a constant elasticity but rather a warning that the linear approximation used in equations (24) and (25) cannot be expected to hold over more than a limited range of values of the interest rate or the expected rate of change in prices.

Empirically, the two approaches give similar results for interest rates in the general range of those observed in the past century in the United States and the United Kingdom. We have not compared the two approaches for the phase bases used in this book. However, in earlier work with annual data we concluded that the semilogarithmic form gave better results than the logarithmic.[45]

Before we discovered the critical importance of the postwar readjustment and the upward demand shift, our regression results were very mixed. We found it necessary to supplement regressions for the period as a whole by regressions for subperiods, and we also calculated a large number of alternative regressions to try out different interest rate series. All in all, we calculated literally hundreds of regressions for different periods and different sets of independent variables. The introduction of shift variables for the postwar readjustment and the upward demand shift brought order into our results, particularly by rendering the various subperiods homogeneous. As a result we have been able to simplify our exposition greatly by restricting attention to regressions for the whole period and, indeed, only a subset of those. We note here only that the conclusions we derive from the regressions we do present are entirely consistent with those we had earlier derived from a much more numerous set of regressions—which we now interpret as a cumbrous and inefficient way to allow for the postwar readjustment and the upward demand shift.

Table 6.11 summarizes regressions for both levels and rates of change, comparing short- and long-term interest rates as variables in money demand equations. Each successive line refers to a single regression. The lines come in triplets: the first is for a regression that includes only the short rate; the second, for one that includes only the long rate; the third, for one that includes both the short rate and the long rate. Columns 1 to 4 give limits on the semilogarithmic slope of real per capita balances with respect to the relevant interest rate. The remaining columns provide evidence on the goodness of fit. For pairs of equations containing the same variables except that one contains R_S and the other R_L, a higher t statistic implies a lower standard error of estimate, so that column 7 simply duplicates columns 5 and 6. For the third equation in the triplicate, containing both, this is also true, though less obviously: the standard error will be less or greater for the third than the first according as the t value for R_L is greater or less than unity. Similarly, it will be less or

---

[45] See also a similar conclusion reached by C. A. E. Goodhart and A. D. Crockett, "The Importance of Money," *Quarterly Bulletin, Bank of England* 10 (June 1970): 192. M. J. Hamburger found that the semilogarithmic form gave better results than the logarithmic for quarterly German (1963–70) and United Kingdom (1963–71) demand functions ("The Demand for Money in an Open Economy: Germany and the United Kingdom," *Journal of Monetary Economics* 3 (January 1977): 29, 34). However, most authors of money demand functions have measured interest rates in logarithmic form. In particular, S. M. Goldfeld obtained good results from such a form for post–World War II United States data (see his "The Demand for Money Revisited," *Brookings Papers on Economic Activity*, no. 3 (1973) pp. 577–646).

greater for the third equation than for the second according as the *t* value for R_S is greater or less than unity.

The results are highly consistent: when only one interest rate is included in the regression, the interest rate slopes are negative, and, with one exception, the *t* values are higher for the short- than for the long-term rate; when both interest rates are included, the short rate dominates and, with one exception, the standard error is intermediate between those for the equations including only one interest rate. These are the results for level equations, for rate of change equations with a zero intercept, and for rate of change equations with a nonzero intercept. (The exception for *t* values is for the United States short-term rate, for rates of change, nonzero intercept; for standard errors, for the United Kingdom rates of change, zero intercept.) The conclusion is clear: the short-term interest rate is preferable to the long, and there is no justification for including both.

As to the magnitude of effect, the limits on the slopes are as usual farther apart for the rate of change equations than for the level equations, but generally the wider rate of change range includes the narrower level range, so the two sets of results from levels and rates of change confirm one another. There seems no appreciable difference in slope between the United States and the United Kingdom.

To judge from the results based on the greatest amount of evidence—for the United States and the United Kingdom combined and the level data—the slope for the short-term interest rate is between -2.8 and -11.8. Since the average value of the short-term interest rate was .037, this corresponds to an elasticity at that average value between $-.10$ and $-.44$.

6.6.2 Interest on Deposits

Nominal yields on assets other than money measure the return on substitutes for money—in terms of the usual demand function for a commodity, they are the counterpart of the price of a substitute, and the elasticities cited in section 6.6.1 are the counterpart of cross-elasticities.

The counterpart of own-price in the usual demand function is the yield on money itself in the form either of services rendered without charge or explicit interest paid on demand deposits or the time deposits we include in our concept of money. However, whereas the own-price is usually the first variable to be considered in demand studies for most commodities and services, we have left it next to last. It is ordinarily completely neglected in money-demand studies on the (implicit) ground that it can be treated as zero. As we noted earlier, Benjamin Klein's study is an exception, and he concluded that allowing for own-yield improves demand functions estimated for the United States from annual data.

Table 6.11 Comparison of Short-Term (R_S) versus Long-Term (R_L) Interest Rate as Variables in Money-Demand Equation

	Interest Rate Slope[c]						
	Short-Term Rate		Long-Term Rate		*t* Statistic[d]		
Country	Lower Limit (1)	Upper Limit (2)	Lower Limit (3)	Upper Limit (4)	Short-Term Rate (5)	Long-Term Rate (6)	Standard Error of Estimate (%) (7)
			Levels				
United States	−2.85	−14.69			3.18		5.54
			−1.95	−25.60		1.86	5.93
	−3.70	−26.48	1.19	85.42	2.58	0.76	5.57
United Kingdom	−3.05	−9.45			3.78		5.37
			−3.06	−13.07		3.03	5.72
	−3.19	−27.28	0.18	584.64	1.96	0.10	5.47
United States and United Kingdom[a]	−2.75	−11.84			4.80		5.51
			−2.29	−18.53		3.28	5.88
	−3.55	26.59	1.10	98.45	3.40	0.92	5.51
			Rates of Change (Zero Intercept)				
United States	−1.92	−27.32			1.78		1.57
			−2.22	−47.36		1.44	1.59
	−1.70	−67.51	−0.40	−554.60	1.03	0.17	1.59

United Kingdom	−1.73	−24.46			1.51		1.42
			−0.62	−93.15		0.45	1.46
	−3.40	−31.94	2.48	55.47	1.86	1.17	1.41
United States and United Kingdom[b]	−1.60	−27.95			2.15		1.45
			−0.94	−74.68		0.99	1.48
	−2.52	−43.23	1.49	109.21	2.15	1.01	1.45
			Rates of Change (Nonzero Intercept)				
United States	−1.37	−33.86			1.32		1.49
			−2.22	−40.00		1.55	1.47
	−0.33	−323.56	−1.87	−112.22	0.20	0.82	1.49
United Kingdom	−2.12	−19.63			1.88		1.37
			−1.13	−51.14		0.81	1.44
	−3.42	−29.49	1.98	64.43	1.92	0.94	1.37
United States and United Kingdom[b]	−1.59	−25.78			2.21		1.40
			−1.29	−51.06		1.39	1.42
	−2.02	−50.65	0.70	224.91	1.75	0.48	1.40

Note: All regressions include real per capita income, postwar readjustment dummy, demand shift dummy, and rate of change of nominal income.

[a] Includes also dummy for country and product of country dummy and real per capita income to allow for different income elasticity in United States and United Kingdom.

[b] Includes also product of country dummy and real per capita income.

[c] "Upper" and "lower" limits refer to absolute value, not algebraic value.

[d] Absolute value.

Following Klein, we allow for own-yield by assuming that currency yields a zero nominal yield and that banks are forced to pass on to their depositors the bulk of the interest they receive on their assets. However, assets held in the form of reserves, that is, high-powered money, earn zero return; hence this approach is equivalent to assuming that the return on money is a weighted average of zero (the return on currency and the high-powered reserves that are the counterpart to some deposits) and the market interest rate (the return on the rest of deposits), the weights being H/M (the ratio of high-powered money to total money) and $1 - H/M$. This requires, in effect, including $R_S(1 - H/M)$, where R_S is the market interest rate, in regressions as an own-price, along with R_S as a price of a substitute.

Table 6.12 shows the effect of the inclusion of own-yield, namely $(1 - H/M)R_S$, in the demand for money regressions. The inclusion of own-yield, in addition to the short-term rate (R_S), consistently lowers the standard error of estimate (compare lines 1 and 2 in each triplet).

Theory suggests that the difference between the yield on close substitute assets and on money should be the relevant variable, since this difference measures the marginal cost of holding an extra dollar as money rather than as an alternative nominal asset. The closeness in absolute value of the coefficients of the short rate and the own rate (columns 1 and 2 for line 2 of each triplet) is consistent with this theoretical expectation. Accordingly, in line 3 of each triplet we include the difference in yield—that is, $R_S - (1 - H/M)R_S = R_S H/M$, which we designate R_N, the N referring to nominal. With two exceptions (levels for United Kingdom and for United States and United Kingdom) this equation yields a lower standard error than either of the others. Moreover, the resulting coefficients are very similar for the United States and the United Kingdom and are consistent for levels and rates of change.[46]

[46] In an excellent article examining the statistical validity of Klein's results, John A. Carlson and James R. Frew ("Money Demand Regressions with the Rate of Return on Money: A Methodological Critique," *Journal of Political Economy* 88 [June 1980]: 598–607) correctly point out that the improvement in Klein's fit obtained by approximating R_M (yield on money) by $(1 - H/M)\ R_S$ and including R_M in a demand equation along with R_S may be spurious, reflecting errors of measurement common to m and H/M. If both numerator and denominator of H/M are divided by NP, the denominator becomes m. Errors of measurement in m that arise from errors in NP will not be common (they cancel out in H/M), but errors of measurement in M will be. They will tend to produce a spurious negative correlation between m and H/M, or a spurious positive correlation between m and $(1 - H/M)$. Carlson and Frew point out further that such common errors will also introduce spurious elements into the computed coefficients of R_S and R_M. (They note quite properly that the fact that R_S is common to R_S and R_M does not of itself introduce any bias in the estimates because both are independent variables. This may however affect the variances and covariances of the estimates.) The spurious element contributes to the coefficients

These results confirm Klein's and suggest including in the demand function the differential yield on nominal assets, rather than the short rate alone or the short rate plus own-yield. At the same time they also show that the effect of substituting the differential yield for the short rate alone, while statistically significant, is quantitatively small: for levels, the substitution reduces the standard error by 8 percent for the United States but raises it by 3 percent for the United Kingdom; for rates of change, the largest improvement is a reduction of the standard error by 2 percent. Hence, for some purposes in later chapters, where including the differential yield greatly complicates the analysis, we shall omit it and include instead the short rate.

---

having opposite signs and, if the coefficient of R_M is positive and that of R_S negative, as is the case in both Klein's and our calculations, to a higher absolute value of the coefficient of R_S.

Some sample calculations by Carlson and Frew for Klein's results suggest that the spurious element could explain his results, though they are careful to note that the case is not proved, since the economic forces Klein emphasizes would work in the same direction as the spurious correlation, and their sample calculations do not discriminate between the spurious and the real effects.

One implication of the Carlson-Frew analysis of spurious correlation, pointed out but not exploited by them, does differ from an implication of economic reasoning and hence provides a basis for discriminating between the spurious and the real effects. As noted in the text, economic forces suggest that the differential rate, $R_S - R_M$, is the relevant alternative cost of holding money. It follows that on economic grounds the semilog *slopes* of R_S and R_M should be numerically equal. On the other hand, Carlson and Frew point out that the spurious statistical effects would produce numerically equal *elasticities* (p. 601, especially footnote 1). For their replication of Klein's result, the slope of R_S is 12 percent less in numerical value than on the slope of R_M; the elasticity of R_S is 20 percent greater (our calculation for the elasticity at the mean values of R_S and R_M). So far as this calculation goes, it argues somewhat in favor of real forces rather than spurious correlation accounting for Klein's results.

The purely spurious statistical element that Carlson and Frew emphasize affects our results in the same direction as Klein's. However, the effect is very likely decidedly smaller quantitatively. Our basic data are phase averages, Klein's annual, which tends to reduce the variance of errors of measurement compared with "true" fluctuations. For rates of change, our estimating them from triplets of phases reduces still further the relative magnitude of pure errors of measurement. This conclusion is supported by a number of pieces of empirical evidence. (1) The correlation between M and H/M or their rates of change is uniformly small. For levels it is negative, as both the spurious element and the secular decline in H/M would imply, but only $-.135$ for the United States, $-.184$ for the United Kingdom. For rates of change it is positive, .045 for the United States, .095 for the United Kingdom. (2) Adding of R_M reduces the standard error of estimate far more in Klein's equation, as recomputed by Carlson and Frew, than in ours (in Klein's by 37 percent; in ours by a maximum of 12 percent in the three comparisons for levels; by much less, in the comparisons for rates of change). (3) Klein's equation yields statistically significant coefficients for both R_S and R_L in an equation including R_M, for R_L but not R_S in an equation excluding R_M. In our case R_S has a statistically significant coefficient in level equations (including and excluding R_L, including and excluding R_M) (see tables 6.11 and 6.12). (4) Including R_M raises the t-statistic (in absolute value) for R_S for Klein's equation from 1.4 to 11.9; for our calculations for United States levels, from 3.2 to 4.1; for United Kingdom

Table 6.12 **Effect of Own Yield on Money in Money-Demand Function**

	Slope			*t* Statistic[c]			
Country	Short-Term Rate (1)	Own Rate (2)	Differential Rate (3)	Short-Term Rate (4)	Own Rate (5)	Differential Rate (6)	Standard Error of Estimate (%) (7)
			Levels				
United States	−2.85			3.18			5.54
	−8.38	7.34		4.13	2.98		5.09
			−8.82			4.45	5.09
United Kingdom	−3.05			3.78			5.37
	−4.84	3.19		5.27	3.08		4.74
			−11.16			3.42	5.54
United States and United Kingdom[a]	−2.75			4.80			5.51
	−5.43	3.98		6.79	4.38		4.95
			−9.28			5.48	5.32
			Rates of Change (Zero Intercept)				
United States	−1.92			1.78			1.57
	−10.52	10.88		1.59	1.32		1.56
			−9.65			2.25	1.54

United Kingdom	−1.73			1.51			1.42
	−18.67	21.43		1.42	1.29		1.40
			−9.48			1.86	1.39
United States and United Kingdom[b]	−1.60			2.15			1.45
	−11.94	13.03		2.04	1.78		1.43
			−8.74			2.74	1.42
		Rates of Change (Nonzero Intercept)					
United States	−1.37			1.32			1.49
	−6.11	5.93		0.91	0.72		1.49
			−6.60			1.51	1.48
United Kingdom	−2.12			1.88			1.37
	−19.12	21.50		1.51	1.34		1.35
			−11.19			2.24	1.34
United States and United Kingdom[b]	−1.59			2.21			1.40
	−11.13	12.04		1.97	1.70		1.38
			−8.48			2.76	1.37

Note: All regressions include real per capita income, postwar readjustment dummy, demand shift dummy, and rate of change of nominal income.

[a] Includes also dummy for country and product of country dummy and real per capita income to allow for different income elasticity in United States and United Kingdom.

[b] Includes also product of country dummy and real per capita income.

[c] Absolute value.

6.6.3 Yield on Physical Assets

Alternatives to holding money include not only the nominal-value assets whose yields were considered in section 6.6.2 but also financial assets, such as equities, that have no stated nominal value and physical assets held directly, such as land, buildings, machinery, and consumer capital. The financial assets generally represent indirect titles to physical assets, so we shall refer to this class of assets as physical assets, in effect treating all nonhuman wealth as falling into one of three classes: money, other nominal-value assets, and physical assets.

The nominal yield on physical assets, like the yield on financial assets, consists of two parts: the direct yield—rent on literal physical assets, dividends on equities—and the change in the nominal price of the assets. And again, the relevant yields for determining the attractiveness of these assets relative to money are the *anticipated* yields.

Most money-demand studies that have included the yield on physical assets have allowed for only the second part of the yield and have approximated it by the rate of change of a price level, often consumer

---

levels, from 3.8 to 5.3; for the two countries together, the t-statistic is reduced a trifle, and that is also true for all but one of the rate-of-change comparisons.

Including R_M does raise the absolute value of the coefficient on R_S for all of our results as well as for the one result of Klein recomputed by Carlson and Frew. But that result is also the one predicted by economic theory, as also is the positive sign of the own rate and the numerically larger coefficient of R_N than of either R_S or R_M. Hence we do not regard these findings as having much bearing on the quantitative importance of the spurious correlation.

In "Competitive Interest Payments and the Demand for Money: Economic Forces or Spurious Correlation?" Michael Melvin recalculates Klein's demand for money function, 1919–70, substituting r_d, the return on deposits, for Klein's r_m variable, and tests the regression results for evidence of real rather than spurious effects. For the shorter period Melvin investigates, he finds that the slopes differ by more than the elasticities, suggestive of spurious effects.

With respect to slopes versus elasticities, our results for levels support real forces versus spurious effects much more strongly than Klein's. The slope of R_S is numerically greater than the slope of R_M for all three comparisons: by 14 percent for the United States, 52 percent for the United Kingdom, and 88 percent for the United States plus United Kingdom (the opposite direction from Klein's result). Since the elasticity is the product of the slope and the interest rate, and since R_S as calculated is necessarily greater than R_M, it follows that elasticities will differ in the same direction and by even more than the slopes. Evaluated at the mean values of R_S and R_M, the numerical elasticity is 58, 94, and 144 percent greater for R_S than R_M for the United States, United Kingdom, and United States plus United Kingdom, respectively.

For rates of change, the slope of R_S is numerically lower than of R_M for five out of six comparisons: by 3, 13, 9, 11, and 9 percent, and higher by 3 percent for the remaining comparison (United States nonzero intercept). The elasticity of R_S is numerically higher for all comparisons, the percentage excess corresponding to the above differentials for slopes, being 32, 11, 21, 12, 20, and, for the one in the same direction, 41 percent. In all but one case, the percentage difference is larger for the elasticity, and generally much larger.

All in all, we conclude that we can have considerable confidence that our results reflect primarily real effects rather than spurious statistical effects.

prices. This procedure has been dictated by paucity of data. There are no satisfactory measures of the direct yield, except for dividends on equities, or of the change in the nominal price of physical assets, except again for equity prices on organized markets. But such equities are only a small part of the total class of assets under consideration.

The use of the rate of change of prices—generally a weighted average of past rates of change—as a proxy for the anticipated nominal yield on physical assets raises the obvious problem that it may be a poor proxy. But it also raises a more subtle problem. Arbitrage on capital markets tends to equalize the attractiveness of the anticipated yields on different assets—which means not that the yields will be equal but that differences will reflect the value placed by investors on such features of assets as risk and liquidity. Anything that changes the yield on physical assets will be transmitted in whole or part also to nominal assets. As Irving Fisher taught many years ago, the anticipated nominal yield on any asset, expressed as a percentage of its value, can be regarded as the sum of an anticipated real yield—that is, a yield after allowance for the effect of inflation on both the income stream and the capital value—and an anticipated rate of change of prices.

Suppose we compute a multiple regression including as variables the nominal yield on nominal assets and the anticipated rate of inflation (rather than nominal yield) on physical assets. We are then measuring the effect of a changed anticipated rate of inflation for a *given* nominal yield on nominal assets. How can these coexist if there is arbitrage between the assets? Only if the implicit real yield changes in the opposite direction so that, for example, a higher inflation rate plus a lower real yield gives a nominal rate on physical assets roughly equivalent to the assumed unchanged rate on nominal assets. But in that case there are two opposite effects on desired balances; by itself, the rise in the rate of price change would lower desired balances; but, by itself, the decline in the real rate would raise desired balances. Hence, in principle, even the sign of the coefficient is ambiguous.[47] In practice there is not full adjustment; the nominal yield serves, as we have stressed, as a proxy for the structure of yields; the rate of price change is likely to enter as another proxy, and a rise in it with a particular nominal yield such as R_S held constant is probably to be interpreted as reflecting a rise in other yields at least as much as a fall in the real yield; hence there is some presumption that its coefficient will be negative. Nonetheless, the effect of the rate of price change is diluted by the failure to include the real rate explicitly.

One way to allow for the real rate is to use the rate of change of nominal income rather than of prices as a proxy for the nominal yield on physical

[47] This point is developed fully by Norman Lefton, "The Demand for Real Cash Balances and the Expected Permanent and Contemporaneous Rates of Change of Prices" (Ph.D. diss., University of Chicago, 1972).

assets—a device employed by Maurice Allais on somewhat different grounds.[48] The use of the rate of change of nominal income has the purely statistical advantage that it is likely to be more accurately measured than the rate of change of prices, particularly during periods of price control (see secs. 4.1.3 and 4.2.3). But its main attraction is economic. The rate of change of nominal income is the sum of the rate of change of prices and the rate of change of output, and the rate of change of output is an estimate, though a downward biased estimate, of the real yield.[49] Hence the rate of change of nominal income can be regarded as a better proxy than the rate of change of prices alone for the *total nominal yield* on physical assets.

On these grounds we have experimented with both the percentage rate of change of prices (g_p) and the percentage rate of change of nominal income (g_Y) as proxies for the nominal return on physical assets in equations for levels of money holdings. In equations for rate-of-change of money holdings, we have used the time derivatives of these rates of change, designated D_{g_P} D_{g_Y} respectively. One problem with both is that we need a measure of anticipated rather than actual rate of change. However, for our initial explorations we have assumed that for our phase data we can regard actual and anticipated rates of change as identical.

Table 6.13 summarizes our regressions for both levels and rates of change: g_Y clearly performs better than g_P; every t value for g_Y is higher than for g_P in absolute value. Hence we shall use it as our proxy for the nominal return on physical assets. The coefficients of g_Y uniformly have the expected negative sign. For the United States and the United States and United Kingdom combined, the coefficients deviate from zero by an amount that is statistically significant. That is not true for the United Kingdom alone. Even when statistically significant, however, the slope coefficients are not very accurately determined. The range between lower and upper limits is wide, so that while the ranges overlap for the United States and for the United Kingdom, and for levels and rates of change, that is no great comfort. The overlap may merely reflect the limited accuracy with which we can estimate the coefficients.

The slope for g_Y tends to be decidedly smaller than that for R_S. It is arithmetic that a one percentage point increase in the nominal yield, whether on physical assets or on nominal assets, raises the cost of holding money instead of the corresponding asset by the same amount. As a

---

[48] See Maurice Allais, "Réformulation de la Théorie Quantitative de la Monnaie," *Bulletin Sedeis*, no. 928, suppl., 10 September 1965; "A Restatement of the Quantity Theory of Money," *American Economic Review* 56 (December 1966), 1123–57; "Growth and Inflation," *Journal of Money, Credit and Banking* 1 (August 1969), 355–426, in which the velocity function depends upon past rates of growth of nominal income.

[49] See R. J. Gordon, ed. *Milton Friedman's Monetary Framework* (Chicago: University of Chicago Press, 1974), p. 37.

matter of economics, it does not follow that the two increases would have the same effect on cash balances. That depends on the composition of the asset portfolio that would be held instead of cash. If that portfolio consisted exclusively of nominal assets, a change in the yield on physical assets would have no effect on cash balances for a given yield on nominal assets, and conversely if the portfolio consisted exclusively of physical assets. We have no direct evidence on what the proportions are, though we do know that the bulk of all material wealth consists of physical assets, if all accounts are consolidated.[50] It does not follow that, at the margin, holders of money could not treat other nominal assets as the chief substitute for money, though it does render any such result rather implausible. If our computed slopes were accurate measures of the effect of a change in the corresponding yield, their relative values would provide a measure of the composition of the substitute portfolio.

The slope coefficients from table 6.11 for R_S are in general a substantial multiple of the corresponding slope coefficients from table 6.13 for g_Y.[51] The implication, if the observed multiples were accurate estimates of the

---

[50] In such a consolidation the only nominal value assets are high-powered money plus government interest-bearing liabilities.

[51] See table 6.N.3.

Table 6.N.3 **Comparison of Effect on Demand for Money of Yields on Nominal and Physical Assets**

	Slope Coefficient(R_S)		Slope Coefficient(g_Y)		Ratio of Coefficients(R_S/g_Y)	
	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit
			Levels			
United States	−2.85	−14.69	−0.71	−2.82	4.0	5.2
United Kingdom	−3.05	−9.45	−0.28	−8.83	10.9	1.1
United States and United Kingdom	−2.75	−11.84	−0.57	−3.97	4.8	3.0
			Rates of Change (Zero Intercept)			
United States	−1.92	−27.32	−0.60	−2.20	3.2	12.4
United Kingdom	−1.73	−24.46	−0.28	−5.57	6.2	4.4
United States and United Kingdom	−1.60	−27.95	−0.45	−3.02	3.6	9.3
			Rates of Change (Nonzero Intercept)			
United States	−1.37	−33.86	−0.45	−2.65	3.0	12.8
United Kingdom	−2.12	−19.63	−0.11	−14.95	19.3	1.3
United States and United Kingdom	−1.59	−25.78	−0.33	−4.03	4.8	6.4

Table 6.13 **Comparison of Rate of Change of Nominal Income (g_Y) and of Prices (g_P) as Variables Representing Nominal Return on Physical Assets in Money-Demand Equation**

Country	Slope[c]				t Statistic[d]		Standard Error of Estimate (%) (7)
	Rate of Change of Income		Rate of Change of Prices				
	Lower Limit (1)	Upper Limit (2)	Lower Limit (3)	Upper Limit (4)	g_Y (5)	g_P (6)	
			Levels				
United States	−0.71	−2.82			3.77		5.54
			−0.64	−5.89		2.27	6.05
	−1.26	−5.86	0.86	13.53	3.35	1.67	5.43
United Kingdom	−0.28	−8.83			1.00		5.37
			0.04	45.18		0.16	5.46
	−1.84	−7.69	1.42	6.61	3.02	2.81	4.84
United States and United Kingdom[a]	−0.57	−3.97			3.54		5.51
			−0.26	−10.23		1.42	5.87
	−1.34	−6.04	0.98	8.58	4.26	3.12	5.22
			Rates of Change (Zero Intercept)				
United States	−0.60	−2.20			3.96		1.57
			−0.30	−9.51		1.17	1.81
	−0.91	−2.81	−0.65	6.42	4.44	2.15	1.51

United Kingdom	−0.28	−5.57			1.25		1.42
			−0.12	−6.95		0.73	1.44
	−0.61	−11.54	0.28	13.34	1.28	0.79	1.43
United States and United Kingdom[b]	−0.45	−3.02			3.66		1.45
			−0.17	−7.61		1.30	1.56
	−0.80	−4.19	0.43	6.26	4.20	2.35	1.41
			Rates of Change (Nonzero Intercept)				
United States	−0.45	−2.65			2.91		1.49
			−0.17	−13.02		0.74	1.62
	−0.73	−3.26	0.54	6.99	3.40	1.83	1.44
United Kingdom	−0.11	−14.95			0.46		1.37
			−0.05	−15.48		0.31	1.37
	−0.22	−36.79	0.09	44.43	0.41	0.23	1.39
United States and United Kingdom[b]	−0.33	−4.03			2.58		1.39
			−0.09	−11.84		0.77	1.45
	−0.63	−5.25	0.35	7.21	3.19	1.96	1.37

Note: All regressions include real per capita income, postwar readjustment dummy, demand shift dummy, and short-term interest rate.

[a] Includes also dummy for country and product of country dummy and real per capita income to allow for different income elasticity in United States and United Kingdom.

[b] Includes also product of country dummy and real per capita income.

[c] "Upper" and "lower" limits refer to absolute value, not algebraic value.

[d] Absolute value.

"true" multiples, would be that, at the margin, the bulk of any change in money balances is as a substitute for nominal value assets. In light of the composition of total wealth, that implication seems extremely implausible.[52]

An alternative interpretation is that the market rate of interest is a better measure of the relevant nominal yield on nominal value assets than the rate of change of nominal income is of the relevant nominal yield on physical assets. The one is a direct measure of yield, the other an indirect measure; the one is quoted contemporaneously on a marketplace, the other must be computed from inexact measures of a hypothetical total after the event.

If nominal yields on nominal and physical assets tended to move together, either would serve as a proxy for the other. That is the case for short and long rates, which tend to move together. It is not, however, the situation for nominal and physical assets. The correlation between R_S and g_Y is $-.189$ for the United States, $-.153$ for the United Kingdom; between their rates of change, $-.102$ for the United States, $-.132$ for the United Kingdom, so that inclusion of g_Y introduces a variable largely independent of others. It follows that it is desirable to include yields on both nominal and physical assets in a demand equation for real balances.

6.7 Effect of All Variables Combined

The preceding sections suggest that the following variables should be included in functions for the United States and the United Kingdom to represent the demand for money, expressed in real terms and per capita (m):

1. Changing financial sophistication for the United States—for which we allow by replacing the raw monetary totals before 1903 by adjusted totals that allow for a 2.5 percent per year increase in the quantity of money demanded arising from this source alone.
2. Real per capita income (y).
3. The difference between the nominal yield on short-term securities and the hypothetical yield on money, as a proxy for the differential yield on nominal value assets sacrificed by holding money (R_N).
4. The rate of change of nominal income, as a proxy for the yield on physical assets (g_Y).

---

[52] However, we should note that the standard Keynesian liquidity preference approach takes bonds as the only substitute for money and so assumes that the whole of any change in money balances is as a substitute for nominal value assets.

Even if physical assets are a decidedly poorer substitute for money per dollar than nominal assets, they are so much more plentiful that the fraction of an additional dollar added to or subtracted from money balances that is matched by an offsetting change in physical assets might be expected to be of at least the same order of magnitude as the fraction matched by an offsetting change in nominal assets.

5. A postwar readjustment, allowed for by including a dummy variable ($W = -T_i$ after World War I, and $+T_i$ after World War II)
6. An upward demand shift, produced by economic depression and war, allowed for by including a dummy variable ($S = 1$ for the phases affected, 0 for the remaining phases).

We found that these six variables enable a single demand function to describe fairly accurately the demand for money in each country for the whole of the century our data cover. In addition, variables 3, 4, 5, and 6 have had about the same quantitative effect in the two countries, provided allowance is made for the different time periods for which items 5 and 6 are relevant.

Item 1 refers only to the United States. The one significant difference of response to any of these variables that we have detected is with respect to item 2: real per capita income. The income elasticity of demand for money is apparently above unity for the United States, below unity for the United Kingdom.

Table 6.14 presents the final equations for the United States and the United Kingdom as derived in three ways: from the levels, from the rates of change with a zero constant term, and from rates of change allowing for a trend via a nonzero constant term. All the coefficients are of the correct sign, and most differ from zero by a statistically significant amount. Three features of these results are most encouraging:

1. The consistency of results from levels and rates of change for both the United States and the United Kingdom, though for reasons discussed earlier the postwar readjustment has a larger effect on rates of change than on levels and the upward demand shift has a larger effect on levels than on rates of change. As usual, the correlation is higher for levels than for rates of change.

Consistency does not mean that the coefficients are identical. The coefficients of the two shift variables aside, because these have a different meaning for levels and rates of change, all but one of the other coefficients is less in absolute value for the rate of change equations with a nonzero intercept than for the level equations. This is the relation to be expected. The coefficients are lower limits (in absolute value) because of the regression effect (the coefficient of income is also biased toward unity because of statistical errors in prices and population). The lower correlation for the rate of change equations means that the regression effect is larger for them.

The standard errors of estimate for level and rate-of-change regressions are not readily comparable. However, as earlier, we can make a rough comparison by multiplying the percentage per year standard errors from the rate of change equations by the average interval between the first and third of the triplet of phases from which each rate of change is calculated. The product should be comparable to the standard error of

Table 6.14 Demand Equations for Money: Final Estimates for Regressions of Real per Capita Money on Other Variables

	Coefficients and t Statistic[a] (in Parentheses)							
Country	Constant	Real per Capita Income	Differential Yield on Money	Proxy Yield on Physical Assets	Postwar Readjustment	Demand Shift	R^2	Standard Error of Estimate (%)
				Levels				
United States	1.53	1.15	−8.82	−0.59	0.025	0.17	.994	5.09
	(9.42)	(50.72)	(4.45)	(3.54)	(3.84)	(6.90)		
United Kingdom	0.16	0.88	−11.16	−0.22	0.014	0.21	.970	5.54
	(0.08)	(18.13)	(3.42)	(0.74)	(2.38)	(7.56)		
			Rates of Change (Zero Constant)					
United States		1.09	−9.65	−0.54	0.033	0.0045	.807	1.54
		(12.03)	(2.25)	(3.84)	(3.00)	(3.45)		
United Kingdom		0.92	−9.48	−0.24	0.030	0.0030	.616	1.39
		(5.41)	(1.86)	(1.08)	(2.26)	(1.21)		
			Rates of Change (Nonzero Constant)					
United States	0.008	0.88	−6.59	−0.42	0.036	0.0033	.618	1.48
	(2.14)	(6.82)	(1.51)	(2.87)	(3.43)	(2.40)		
United Kingdom	0.008	0.44	−11.19	−0.06	0.026	0.0027	.502	1.34
	(1.81)	(1.44)	(2.24)	(0.26)	(2.03)	(1.13)		

[a] Absolute value.

the level, which is a value at a point of time. For the United States, the result is modestly higher than the level standard error (6.16 and 5.92, for equations with a zero and a nonzero constant, vs. 5.09); for the United Kingdom, considerably higher (7.64 and 7.37 vs. 5.54).

These results are in line with the theoretical expectation that estimates from the level equations can be expected to be the most reliable of the three sets of estimates.

2. The closeness of the coefficients for the United States and the United Kingdom. Only the coefficients of the income term from the level equations differ significantly between the two countries. Even the standard errors of estimate are not far apart for the two countries.

For the income term from levels, the estimated elasticity for the United States is 1.15, for the United Kingdom, 0.88; the difference, 0.27, is roughly five times the standard error of the difference—highly significant. For the income term from rates of change, the difference is in the same direction as for levels, but is not statistically significant, being roughly equal to the standard error of the difference for both forms of rate-of-change equations.

The only other difference is in overall level of demand, on which we cannot use the results in table 6.14, since the United States data are in dollars and the United Kingdom data in pounds. We shall return to this point later.

Allowing for the yields has reduced the size of the difference between the United States and the United Kingdom in income elasticity, but unfortunately it has not eliminated the difference. The reason for this difference is the major mystery for which we have been unable to find an explanation.

The implication is that, income aside, the same basic forces affect money demand in the two countries, have the same quantitative impact, and leave the same residual to be explained by statistical error or omitted economic variables.

We shall exploit this implication by constructing a single equation for the two countries combined.

3. How far we have been able to go in accounting for fluctuations over a century in real money holdings on the basis of six variables only. The implication is that money demand is stable over time—in the sense of a demand function—as well as between countries.

This excellent overall result is misleading in one respect. Although the income elasticities can be regarded as having been estimated fairly precisely, the slopes of the two yield variables are much less precisely estimated. This is brought out by table 6.15, which gives upper and lower limits based on regressions run both ways for the coefficients of the three quantitative economic variables.

The generally far wider limits for the rate-of-change regressions than for the level regressions simply reinforce the earlier conclusion about the greater reliability of the level regressions.

But even for the level regressions the coefficients of the yield terms are not specified very precisely. Taking the results from the regressions for two countries, the range is from -9 to -40 for the slope of the differential yield on money and from -0.2 to -12 for the proxy yield on physical assets.

Note that these are slopes, not elasticities. To convert them into elasticities requires multiplying them by the value of the relevant yield. Doing so at the mean value of the yields gives an elasticity of -0.10 to -0.32 for the differential yield and -0.01 to -0.49 for the proxy yield.

As a final summary of our results, we present two single equations for the two countries combined, one computed from levels, one from rates of change. For the level equation, we add a country dummy ($Z = 1$ for the United Kingdom, 0 for the United States) to allow for the difference in level (and perhaps also for a deviation between the market and relevant exchange rate), and we add a term equal to the product of the country dummy and the logarithm of real per capita income (Z times $\log y$) to allow for the difference in income elasticity between the United States and the United Kingdom. The result is

$$(26)\qquad \begin{array}{l} \log m = -\underset{(9.6)}{1.47} + \underset{(52.1)}{71.14} \log y - \underset{(5.5)}{9.3}\, R_N - \underset{(3.1)}{0.47}\, g_Y \\[1em] \qquad + \underset{(4.7)}{.019} W + \underset{(10.8)}{.193} S + \underset{(6.2)}{1.64} Z - \underset{(6.0)}{0.25}\, Z \log y \\[1em] \qquad\qquad R^2 = .9889 \\ \qquad\qquad SEE = 5.32 \text{ percent.} \end{array}$$

For income elasticities, this gives 1.14 for the United States and 0.89 (1.14 − 0.25) for the United Kingdom as lower limits—or very much the same as for the separate country equations (table 6.15).

The -9.3 slope for the differential yield on money corresponds to an elasticity of -0.19, the -0.47 slope for the proxy yield, to an elasticity of $-.02$.

These results simply repeat the earlier ones.[53] The new result is the coefficient of Z, which indicates the difference in level for the United States and the United Kingdom. For this equation, the United Kingdom figures were converted to dollars at the 1929 exchange rate ($4.862 to the £). At that exchange rate table 6.16 gives the differences in level of money

---

[53] We do not give upper limits because of difficulties in determining them produced by the $Z \log y$ term.

Table 6.15 **Upper and Lower Limits[a] of Income Elasticity of Money Demand and Semilog Slopes of Differential Yield on Money, and Proxy Yield on Physical Assets**

			Semilog Slope			
	Income Elasticity		Differential Yield on Money		Proxy Yield on Physical Assets	
Country	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit
			Levels			
United States	1.15	1.17	−8.82	−27.52	−0.59	−2.59
United Kingdom	0.88	0.96	−11.16	−39.76	−0.22	−11.92
			Rates of Change (Zero Constant)			
United States	1.09	1.40	−9.65	−89.74	−0.54	−2.08
United Kingdom	0.92	1.86	−9.48	−91.49	−0.24	−6.29
			Rates of Change (Nonzero Constant)			
United States	0.88	1.66	−6.59	−124.65	−0.42	−2.50
United Kingdom	0.44	6.62	−11.19	−75.94	−0.06	−26.71

[a] Upper and limits refer to absolute, not algebraic values.

Table 6.16 **Differences in Level of Money Demand in the United States and the United Kingdom at Three Income Levels**

Income Levels	Real per Capita Money Holdings (in Dollars)		Percentage Excess of United Kingdom Money Holdings over United States Money Holdings (%)
	United States	United Kingdom	
United States, 1877	109	136	+25
United Kingdom, 1975	454	423	−7
United States, 1975	832	679	−18

demand at the three income levels[54] isolated in chart 6.3, and at the values of the other variables for the United States at the corresponding dates.

Though the difference between the two countries is substantial, it is appreciably less than the difference when no allowance is made for differences between them in yields on alternative assets. The corresponding differences in chart 6.3 are +40 percent for 1877 income levels; −13 percent for the 1975 United Kingdom income level; −30 percent for the 1975 United States income level. Allowing for yields has cut these differentials almost in half. Nonetheless, the mystery of why there is a difference in income elasticity remains.

The rate-of-change equation for the two countries combined, calculated with a zero constant term, to correspond with the absence of a trend term in the level equation, is as follows:

$$
\begin{aligned}
(27) \qquad g_m = {} & \underset{(10.9)}{1.03\, g_y} - \underset{(2.7)}{8.74\, DR_N} - \underset{(3.6)}{0.41\, Dg_Y} + \underset{(3.7)}{.030\, W_g} \\
& + \underset{(3.3)}{.0040 S_g} - \underset{(0.3)}{.04\, Zgy} \\
& R^2 = .72 \\
& SEE = 1.54 \text{ percent.}
\end{aligned}
$$

The results are very similar to those for the level equation, except that the difference between the income elasticities for the United States and the United Kingdom is not statistically significant, as it was not for the separate country equations. As is to be expected, the regression bias is more serious, so that the coefficients of the income and rate-of-return

---

[54] In 1877, United States and United Kingdom real per capita incomes (in 1929 dollars) were equal. In 1975, United States real per capita income was higher than that of the United Kingdom.

variables are lower in absolute value (except for the United Kingdom income elasticity).

6.8 Appendix A: Issues in Using the Exchange Rate to Pool United States and United Kingdom Data

This appendix explores the problem of the exchange rate to use in pooling the data for the United States and the United Kingdom.

The market exchange rate may not be the relevant exchange rate for our regressions for two different reasons. One has been discussed in chapter 5: the relation between the exchange rate and price levels in the two countries can be affected by the composition of the balance of payments, in particular the role of capital exports from or imports to each country and of income from abroad from sources other than the current sale of goods and services.[55] The second reason is more immediately pertinent. The domestic prices in the two countries of goods traded internationally for which there are no tariffs or export subsidies and for which transportation costs can be neglected are necessarily in the same ratio as the exchange rate. For example, if the exchange rate for the pound sterling is $4.86, then such a good which sells for £100 in the United Kingdom will sell for $486 in the United States. (This proposition is sometimes known in the literature of international economics as "the law of one price.") Of course, even for international goods, tariffs and transportation costs may be sufficiently important to introduce substantial deviations between the relative prices in different countries and the exchange rate. But for other goods and services that are not traded the deviation may be far greater. For each such good or service, an implicit exchange rate can be calculated from domestic prices that may differ widely from the market rate. For example, consider the price of an hour of roughly comparable domestic service in the United States and the United Kingdom. The exchange rate required to make these two prices—one in pounds and the other in dollars—equal when converted to a common currency might be, say, $10 to the pound rather than $4.86 if domestic service is cheaper relative to internationally traded goods in the United Kingdom than in the United States (as seems indeed to have been the case). There are as many implicit exchange rates as there are identifiable goods and services available in both countries. There is still a link between these implicit exchange rates and the market rate—via substitution of trade for production, substitution among factors of production, and migration of labor and capital—but these links may be very loose indeed.[56] The tighter they are, the closer relative prices for different

---

[55] See notes 4 and 14 in chapter 5.

[56] For an excellent summary discussion of many of these issues and references to the literature on them, see Irving B. Kravis, Alan W. Heston, and Robert Summers, "Real

goods in different countries will be to the market exchange rate; the looser they are, the wider will be the variation of relative prices for different goods from the market exchange rate.

Consider a money-holding unit (an individual, a business enterprise, or a government or other unit) deciding on the amount of money to hold. By holding an extra dollar (or pound) of money, it saves, say c cents (or pence) in pecuniary costs plus receiving, say, nonpecuniary returns it values at n cents (or pence). Neither c nor n, of course, is a constant; the size of both depends on the real amount of money held (m) and the level of real income (y) or wealth. By holding an extra dollar (or pound) of money, the unit incurs a cost of, say, R cents (or pence) that could have been earned by holding alternative assets (including, of course, borrowing less). And R too, of course, need not be a constant but may depend on m, y, and other variables. The unit will tend to hold an amount of money so that

$$c + n = R$$

for that amount of money and its level of income. In general, of course, the higher c and n, and the lower R, for given values of m and y, the higher will be desired m.

Each of the magnitudes introduced in the preceding paragraph (c, n, R, m, y) raises a problem in using the market exchange rate for international comparisons.

For c, the services in question are those that the use of money economizes, such as bookkeeping services. Casual observation suggests that the relevant services involve mostly personal labor-intensive white-collar services, and that such services have been cheaper relative to internationally traded goods in the United Kingdom than in the United States. If these casual impressions are correct, they establish a presumption that the relevant implicit exchange rate for such services (for 1929) is greater than \$4.86.

For n it is difficult to be equally specific because the alternative ways of obtaining nonpecuniary services substituting for those rendered by money are harder to visualize specifically (presumably, alternatives include insurance, credit lines, credit cards, larger holdings of financial or physical assets, including such things as jewelry, and even different employment).

For R there is clearly an international capital market. Nonetheless, yields on domestic assets, for a country that has net foreign assets, are, as Ricardo pointed out long ago, likely to be lower than on actuarially

---

GDP *per Capita* for More Than One Hundred Countries," *Economic Journal* 88 (June 1978): 215–42.

equivalent foreign assets because of additional transactions costs associated with holding foreign assets, and because of the general preference for investing at home. In principle, both these factors are taken into account through *c* and *n*—the transactions costs through *c*, the preference for domestic assets through *n*—provided that the interest rate used is the higher yield on foreign assets. In practice, however, we use the domestic interest rate, so some effect may remain. However, a more important factor is probably the capital exporting or importing status of the country, which will affect not only the market exchange rate but also interest rates.

For *M* and *Y*, nominal money and income, the market exchange rate is clearly the relevant one—nominal dollars and pounds are internationally traded goods par excellence. But what of *m* and *y*, real money and real income per capita? For each country separately, we have followed the practice of deflating nominal money and nominal income by the same price index, the index implicit in estimates of real net national product. The same practice for intercountry as for intertemporal comparisons would mean using the same exchange rate for real money and real income. The common exchange rate comparable to the internal price index we use would be the purchasing power parity exchange rate, that is, the average obtained by weighting all implicit exchange rates for goods and services by the contribution of each good and service to national product. As noted earlier (sec. 6.4.1), an estimate of the purchasing-power-parity (PPP) exchange rate in 1929 can be made on the basis of a recent direct estimate for 1970. The estimated rate in $5.50.

Kravis, Heston, and Summers (1978) estimate that, for 1970, the purchasing-power-parity exchange rate of the United Kingdom pound in terms of the United States dollar was 117 percent of the market rate for traded goods, 173 percent of the market rate for nontraded goods, and 139 percent for both combined.[57] The large difference for traded goods may reflect primarily the definition of traded goods used in making this calculation rather than the effect of tariffs and transportation costs. Traded goods were defined as including all commodities as finally sold, so even the final prices for most commodities "contain large service elements attributable to trade and transport margins."[58]

The 1970 estimate of the ratio of the purchasing-power-parity exchange rate to the market rate can be extrapolated to a different year by multiplying it by the ratio of the change in internal prices between 1970 and that year in the United States to the corresponding change in the United Kingdom. If *PPP*(*t*) is the estimate of the purchasing-power-parity exchange rate in year *t*, then

---

[57] Ibid., p. 216.

[58] Ibid., p. 224.

$$PPP(t_1) = \frac{P_{us}(t_1)/P_{us}(t_2)}{P_{uk}(t_1)/P_{uk}(t_2)} PPP_d(t_2),$$

where PPP_d is a direct estimate for time t_2 and P is the implicit price index for the country designated by the subscript and the indicated date. This is essentially the method used in footnote 5 of chapter 5 to get an estimate for 1873–78 based on 1929 data of the ratio of incomes in the United States and the United Kingdom. For 1929, this method gives a purchasing-power-parity exchange rate of $5.50 in 1929, or a ratio of 1.13, compared with the 1970 ratio of 1.39. Similar estimates for the ratio of the purchasing-power-parity to the United States–United Kingdom market exchange rate are plotted in chart 6.5, annually, 1868–1975.

Chart 6.5 shows that "the law of one price" came much closer to being satisfied before 1932 than from that year on. This is the first and most important conclusion. The first sharp isolated peak is for the year 1932. That point provides a dividing line for the periods before and after 1932. For the period before 1932, the average ratio was about 1.12; that is, on the average, the purchasing-power-parity exchange rate was 12 percent above the dollar exchange rate. The dollar exchange rate during the fixed exchange-rate period was $4.86, so the purchasing-power-parity change rate was 12 percent above that, or about $5.40, or very close to the estimated value of $5.50 for 1929, the year we use as the base for our real income and money series. What is even more interesting is the fluctuation about the mean. Before 1932, the highest ratio is only 10 percent above the average; the lowest ratio, only about 10 percent below. So the purchasing-power-parity exchange rate fluctuated within a narrow range of plus and minus 10 percent, which seems consistent with a reasonably unified market, given the statistical error in estimates extrapolated back by a period of 38 to 102 years (from 1970 to 1868 to 1932), plus the time it takes for adjustment. The main movements of the purchasing-power-parity rate seem to reflect inflation and deflation in the United States. The purchasing-power-parity exchange rate went down from the 1880s to the 1890s when prices were falling in the United States relative to prices in the United Kingdom and went up thereafter when prices were rising.

The situation after 1931 is very different. In September 1931 the United Kingdom went off the gold standard. Thereafter, the United Kingdom had an exchange rate that was sometimes floating but mostly pegged by the Bank of England, with occasional large devaluations—temporarily fixed but subject to change. The United States stayed on the gold standard for two more years and thereafter also had a mixture of floating and pegged rates. The source of the sharp spike in 1932 is that 1932 was the one year when the United Kingdom was off the gold standard and the United States stayed on it. The depreciation of sterling made United Kingdom goods cheap relative to United States goods;

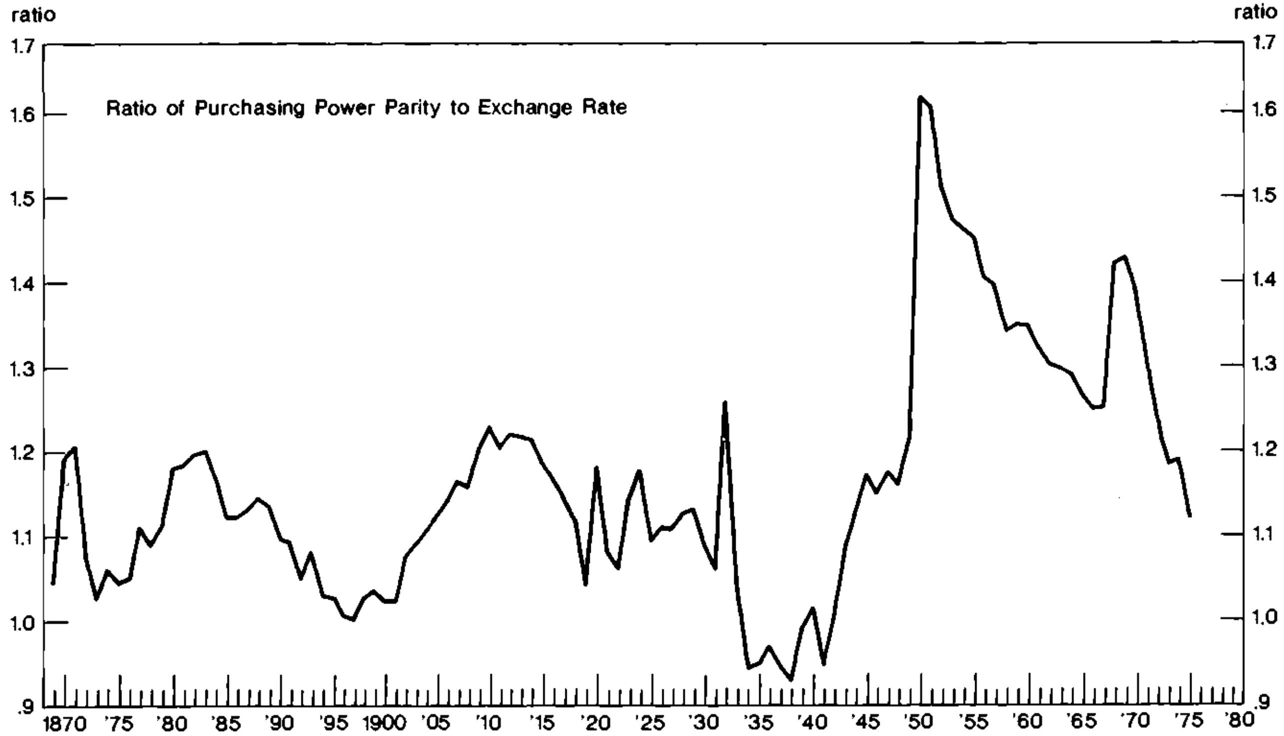

Chart 6.5 Ratio of the purchasing-power-parity exchange rate to the market exchange rate between the United States dollar and the United Kingdom pound, annually, 1868–1975.

equivalently, it meant that one pound could buy a larger quantity of goods than could the number of dollars corresponding to that pound according to the market exchange rate. In 1933 the United States went off the gold standard. The dollar depreciated even more than the pound had. As a result the ratio drops suddenly and drastically. From then on there were successive ups and downs. Each prominent peak in the series thereafter corresponds to a United Kingdom devaluation.

The picture is quite clear. Every now and then, for whatever reason, the United Kingdom government has stepped in and either devalued the currency deliberately or permitted it to depreciate. That has forced the purchasing-power-parity exchange rate out of line with the market exchange rate. Then market forces set in to correct it. The highest peak is for 1949–50, and that corresponds to the immediate post–World War II devaluation of the pound in 1949. The market then gradually starts to bring the ratio back down toward about 1.2. Before it gets there, another devaluation occurs and the ratio shoots up again. The market again brings it down, the United States devaluation in 1971 speeding up the process. Then, to go beyond the period covered by the chart, the depreciation of the pound in 1975 and 1976 pushed it up again. Once again, it has started coming down toward the range that market forces dictate.

The average for the period after 1931 is not much different from the average for the period before 1931—1.21 compared with 1.12. The real difference is in the fluctuations—plus and minus 10 percent before 1931, from minus one-quarter to plus one-third thereafter.

Since 1931 there have been tremendous improvements in communications and in transportation. The jet aircraft now spans the ocean in a few hours. Satellite transmission and television and radio communications link countries instantaneously and at relatively low cost. The cliché is it has become one world. In the economic world, the reality is clearly the reverse. The law of one price was far closer to being satisfied before 1931 than after. The technological improvements, which might have been expected to unify the world, have been more than offset by governmental intervention, which has fragmented the world into separate, isolated markets. Chart 6.5 demonstrates vividly how powerful and effective government intervention has been in rendering the law of one price far less applicable after 1931 than it was before.[59]

[59] For an interesting approach to the same issue from a different perspective, see Sven Grassman, "Long-Term Trends in Openness of National Economies," *Oxford Economic Papers*, n.s., 32 (March 1980): 123–33. His broad conclusion is that "several European and North-American economies have roughly the same degrees of openness today as a century ago" (p. 123). His detailed tables for decades (tables A1 and A3) confirm our own finding about the difference between the periods before and after 1930. For decades before 1925, both foreign trade and foreign capital movements are more important for the United Kingdom relative to national income than for decades after 1944. That is on the average also true for the United States though less consistently.

It is worth repeating that the deviations of the purchasing-power-parity exchange rate from the market rate do not affect the slope of our money-income relations so long as the same adjustment is made for both money and income, but they do affect the United Kingdom income in pounds that is regarded as equivalent to the United States income in dollars. Hence they will affect our estimates of the difference in level of real money balances for a given real income—though even here, as we have seen, any effect is likely to be quantitatively minor.

The effects would be more far-reaching if a different adjustment were made for money than for income. But that issue is not basically one of how to make international comparisons; it arises for each country separately. For example, consider the use of time-series data to derive the relation between real food expenditure per capita and real income per capita. For such a study it might well be fruitful to deflate nominal food expenditure by a food price index and nominal income by a general price index. Using the latter index for both is equivalent to examining the effect of real per capita income not on real food expenditure but on the fraction of income spent on food—the counterpart to our relations that can be regarded as relating velocity to income. Put differently, our use of the same index for money and income means that we are in effect defining real money as the number of weeks' income to which it is equivalent.

The use of different index numbers for money and income could in principle contribute to explaining our puzzle of the different elasticities in the United States and the United Kingdom. For example, suppose that the ratio of the relevant price index for money to the relevant index for income fell in the United Kingdom over the century but remained constant in the United States. That would make the computed income elasticity of separately deflated money higher than the elasticity we computed. However, we know no student of the demand for money who has explored this route, we have not ourselves done so, and we have not thought of any feasible way to establish a presumption about the direction of effect. Hence we simply mention but do not further explore this possibility.

The same issue arises of course with respect to the unobserved variables c and n. If, for given real income, and other relevant variables, the ratio of the price index relevant to these variables to the general price index had fallen over time in the United Kingdom relative to the corresponding ratio for the United States, that would have reduced over time the ratio of desired m in the United Kingdom to desired m in the United States for the equivalent real incomes per capita—which would, in our calculations, have reduced computed United Kingdom elasticity relative to United States elasticity. Offhand, it seems more nearly feasible to pursue this possibility than the preceding one, by estimating purchasing-power-parity exchange rates for nontraded goods, on the assumption that *c*, at least, and perhaps also *n*, depend disproportionately on the price of nontraded goods. However, that would be a far more laborious task than we have thought it justified to undertake for our purpose.

6.9 Appendix B: Incorporating the Term Structure of Interest Rates in the United States Demand for Money Equation

This appendix describes the experiment, referred to in footnote 43 above, to incorporate the term structure of interest rates in the United States demand for money equation. We first discuss the term structure data set, then the generation of the parameters to describe the term structure of yields, and finally the equation replacing the short-term interest rate with the term structure parameters.

6.9.1 United States Term Structure Data

Annual estimates of yield curves for the best grade corporate bonds are available since 1900.[60] For 1900–1941, the annual yield curves were derived from averages of high and low sale prices of the best-grade corporate bonds during each month of the first quarter of the year. Yields for 1941 were based on January and February prices, and yields for later years on February prices.

To match the period covered by our phase data, we needed term structure yields for 1873–99 for all five maturities available for 1900–1970 and for 1971–75 for the one-year-to-maturity yields.

Our basic source for the period before 1900 was Macaulay's tables of railroad bond yields.[61] He gives the yield of high-grade railroad bonds if held to maturity for each month of each year. Maturity date is given as part of the name for each bond. Following Durand, we used the February yields for each year, excluding any bond whose maturity exceeded fifty years—in some cases bonds with three hundred years to maturity were listed. For each year *i* we estimated a yield curve, using the quadratic function:

$$R_i(\tau) = a_{0i} + a_{1i}\,\tau + a_{2i}\,\tau^2\,, \tag{A1}$$

---

[60] The estimating procedure is described in David Durand, *Basic Yields of Corporate Bonds, 1900–1942*, Technical Paper 3 (New York: NBER, 1942), p. 4. The data for five maturities (1, 5, 10, 20, and 30 years) are given in *Historical Statistics of the United States, Colonial Times to 1970*, Bicentennial Edition, part 2, (Washington, D.C.: Bureau of the Census, 1975), ser. X487-491, p. 1004. Data for 1979–75 for four maturities (5, 10, 20, and 30 years) are given in *Statistical Abstract of the United States, 1976* (Washington, D.C.: Bureau of the Census, 1976), p. 495. The Durand estimates were for the following years to maturity (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 20, 25, 30, 40, 50, 60).

[61] F. R. Macaulay, *The Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1856* (New York: NBER, 1938), pp. A45–A78.

where R_i is the yield per bond in year i, τ is years to maturity, and a_{0i}, a_{1i}, and a_{2i} are parameters to be estimated for year i.[62] From this equation we calculated the estimated bond yield for each year from 1873 to 1902 for the five maturities (1, 5, 10, 20, and 30 years) that are available from 1943 to 1970. We linked these estimates to the corresponding Durand estimates by lowering the yield for each maturity by the average difference between each pair of series in the overlap period, 1900–1902.

To supplement the four longer maturities for 1971–75, we estimated the one-year yield to maturity, using the yields for the other four maturities (the only ones available for these years) for each year in the quadratic function.

Table 6.17 lists the anual yield to maturity estimates for 1873–1975. We then calculated the phase averages for each maturity to conform to the rest of our observations.

6.9.2 Estimating the Parameters of the Term Structure Yields

The estimation process described by Heller and Khan (see footnote 43 above) has two stages. In the first stage, estimates of a_0, a_1, and a_2 were obtained from the quadratic function ($A1$) for individual phases rather than years. In the second stage, these parameters replace the phase single interest rate included in the standard version of the demand for money equation. As Heller and Khan note, a_0 "can be viewed as a shift parameter for the entire term structure of interest rates" (p. 114). The slope of the term structure equals $a_1 + 2a_2\,\tau$, and the curvature $2a_2$. If the term structure slopes positively, at least initially, and is concave downward, generally taken to be normal, a_1 will be positive and a_2 negative. With a "U-shaped" yield curve, a_1 will be negative, and a_2 positive. Table 6.18 lists the parameters of the yield curve for each phase observation.

For pre–World War I phases, the yield curve alternated between a positive slope and negative curvature and negative slope and positive curvature. The latter shape dominated the interwar years until 1927–29. Thereafter, with only one exception through 1961–66, and again in the 1970s, the former shape predominated. Both parameters positive (1882–85, 1957–58) or negative (1914–18, 1967–69, 1969–70) was rare. Heller and Khan describe the case of negative slope and curvature as "fairly unrealistic and does not appear to have occurred during our sample period." The basic corporate yield data do, however, conform to such a yield curve in 1967–70.

---

[62] Heller and Khan (see note 43 of this chapter) used the same quadratic function but fitted it to log R, instead of R, as we do.

Table 6.17 **Estimated Yields of United States Corporate Bonds by Term to Maturity, 1873–1975 (Annual Percentage): Number of Years to Maturity**

Date	1	5	10	20	30
1873	6.39	6.43	6.50	6.60	6.61
1874	6.15	6.24	6.35	6.51	6.53
1875	5.17	5.41	5.67	6.08	6.26
1876	4.42	4.79	5.19	5.72	5.84
1877	5.23	5.35	5.49	5.68	5.69
1878	4.41	4.79	5.18	5.66	5.67
1879	4.32	4.58	4.86	5.19	5.18
1880	4.41	4.55	4.72	4.98	5.08
1881	4.58	4.52	4.48	4.51	4.61
1882	4.36	4.35	4.38	4.51	4.69
1883	4.28	4.32	4.39	4.56	4.68
1884	4.40	4.38	4.38	4.44	4.51
1885	4.22	4.25	4.30	4.40	4.44
1886	3.79	3.82	3.87	3.97	4.03
1887	3.87	3.88	3.92	4.01	4.04
1888	4.12	4.07	4.05	4.05	4.07
1889	3.61	3.69	3.79	3.94	3.95
1890	4.10	4.05	4.01	3.97	3.93
1891	4.26	4.01	4.18	4.14	4.08
1892	4.30	4.01	4.14	4.06	3.99
1893	4.66	4.47	4.28	4.02	3.88
1894	4.88	4.61	4.33	3.98	3.84
1895	5.66	4.39	4.13	3.80	3.70
1896	4.59	4.32	4.05	3.73	3.62
1897	4.39	4.13	4.88	3.60	3.54
1898	4.02	3.79	3.59	3.38	3.38
1899	3.77	3.59	3.43	3.26	3.23
1900	3.97	3.36	3.30	3.30	3.30
1901	3.25	3.25	3.25	3.25	3.25
1902	3.30	3.30	3.30	3.30	3.30
1903	3.45	3.45	3.45	3.45	3.45
1904	3.60	3.60	3.60	3.60	3.60
1905	3.50	3.50	3.50	3.50	3.50
1906	4.75	3.67	3.55	3.55	3.55
1907	4.87	3.87	3.80	3.80	3.80
1908	5.10	4.30	4.02	3.95	3.95
1909	4.03	3.97	3.91	3.82	3.77
1910	4.25	4.10	3.99	3.87	3.80
1911	4.09	4.05	4.01	3.94	3.90
1912	4.04	4.00	3.96	3.91	3.90
1913	4.74	4.31	4.12	4.02	4.00
1914	4.64	4.45	4.32	4.16	4.10
1915	4.47	4.39	4.31	4.20	4.15
1916	3.48	4.03	4.05	4.05	4.05
1917	4.05	4.05	4.05	4.05	4.05
1918	5.48	5.25	5.05	4.82	4.75
1919	5.58	5.16	4.97	4.81	4.75

Table 6.17 (Continued)

Date	1	5	10	20	30
1920	6.11	5.72	5.43	5.17	5.10
1921	6.94	6.21	5.73	5.31	5.17
1922	5.31	5.19	5.06	4.85	4.71
1923	5.01	4.90	4.80	4.68	4.61
1924	5.02	4.90	4.80	4.69	4.66
1925	3.85	4.46	4.50	4.50	4.50
1926	4.40	4.40	4.40	4.40	4.40
1927	4.30	4.30	4.30	4.30	4.30
1928	4.05	4.05	4.05	4.05	4.05
1929	5.27	4.72	4.57	4.45	4.42
1930	4.40	4.40	4.40	4.40	4.40
1931	3.05	3.90	4.03	4.10	4.10
1932	3.99	4.58	4.70	4.70	4.70
1933	2.60	3.68	4.00	4.11	4.15
1934	2.62	3.48	3.70	3.91	3.99
1935	1.05	2.37	3.00	3.37	3.50
1936	0.61	1.86	2.64	3.04	3.20
1937	0.69	1.68	2.38	2.90	3.08
1938	0.85	1.97	2.60	2.91	3.00
1939	0.57	1.55	2.18	2.65	2.75
1940	0.41	1.28	1.95	2.55	2.70
1941	0.41	1.21	1.88	2.50	2.65
1942	0.81	1.50	2.16	2.61	2.65
1943	1.17	1.71	2.16	2.61	2.65
1944	1.08	1.58	2.20	2.60	2.60
1945	1.02	1.53	2.14	2.55	2.55
1946	0.86	1.32	1.88	2.35	2.43
1947	1.05	1.65	2.08	2.40	2.50
1948	1.60	2.03	2.53	2.73	2.80
1949	1.60	1.92	2.32	2.62	2.74
1950	1.42	1.90	2.30	2.48	2.58
1951	2.05	2.22	2.39	2.59	2.67
1952	2.73	2.73	2.73	2.88	3.00
1953	2.62	2.75	2.88	3.05	3.15
1954	2.40	2.52	2.66	2.88	3.00
1955	2.55	2.70	2.80	2.95	3.04
1956	2.70	2.78	2.86	2.99	3.09
1957	3.50	3.50	3.50	3.50	3.68
1958	3.21	3.25	3.33	3.47	3.61
1959	3.67	3.80	4.03	4.10	4.10
1960	4.95	4.73	4.60	4.55	4.55
1961	3.10	3.75	4.00	4.12	4.22
1962	3.50	3.97	4.28	4.40	4.42
1963	3.25	3.77	3.98	4.10	4.16
1964	4.00	4.15	4.25	4.33	4.33
1965	4.15	4.29	4.33	4.35	4.35
1966	5.00	4.97	4.91	4.80	4.75
1967	5.29	5.28	5.23	5.00	4.95
1968	6.24	6.24	6.20	6.00	5.93

Table 6.17 (Continued)

Date	1	5	10	20	30
1969	7.05	7.05	7.05	6.77	6.54
1970	8.15	8.10	8.00	7.60	7.60
1971	5.32	5.85	7.05	7.12	7.12
1972	6.26	6.50	7.05	7.05	7.01
1973	6.69	6.85	7.05	7.20	7.20
1974	7.32	7.47	7.67	7.80	7.80
1975	7.39	7.70	8.00	8.35	8.35

6.9.3 The Demand for Money, with the Term Structure of Interest Rates

We estimated the United States money demand equation, replacing R_S with the three parameters of the term structure shown in table 6.18. The results[63] for the alternate versions are as follows:

(A2)
$$\log m = \underset{(14.9)}{-1.91} + \underset{(64.4)}{1.21 \log y} - \underset{(3.2)}{2.85\, R_S}$$
$$\underset{(3.8)}{-0.71\, g_Y} + \underset{(4.3)}{0.15S} + \underset{(2.6)}{.019\, W}$$
$$R^2 = .9919$$
$$SEE = 5.54 \text{ percent.}$$

(A3)
$$\log m = \underset{(15.1)}{-1.93} + \underset{(60.2)}{1.21 \log y} \underset{(2.6)}{-2.78\, a_0} \underset{(2.7)}{-298 a_1}$$
$$\underset{(2.8)}{-13823 a_2} \underset{(3.6)}{-0.71 g_Y} + \underset{(4.8)}{0.185S} + \underset{(2.9)}{0.021\, W}$$
$$R^2 = .9916$$
$$SEE = 5.64 \text{ percent.}$$

In terms of goodness of fit, equation (A3) with term structure parameters does about as well as equation (A2) using the short-term interest rate, and it yields statistically significant coefficients for all parameters. It provides more information than equation (A2) by indicating the effect on the demand for money of changes in the term structure of interest rates.

The intercept parameter has a clear meaning: an increase in a_0 by .01 implies a one percentage point rise in the whole term structure. Hence, the slope of the intercept parameter, a_0, indicates that a uniform upward

---

[63] Note that in these equations, the values for R_S and for a_0, a_1, and a_2 are entered as decimals, not as percentage points, so that the values entered are 1/100 of values such as given in tables 6.17 and 6.18. That is true also, of course, of g_Y.

Table 6.18 **Parameters of the Yield Curve for Each Phase Observation (Interest Rates Expressed as Percentages)**

Phase Number	$100a_0$	$100a_1$	$100a_2$
4	5.2055	0.061700	-0.001068
5	4.3764	0.040500	-0.000638
6	4.3085	0.005876	0.000116
7	3.8986	0.010700	-0.000090
8	3.9847	-0.000119	0.000087
9	3.8424	0.008997	-0.000148
10	4.1429	-0.007933	0.000122
11	4.2070	-0.011000	0.000195
12	4.6457	-0.046800	0.000734
13	5.2581	-0.127000	0.002630
14	5.1112	-0.126400	0.002654
15	4.2360	-0.038400	0.000363
16	3.8472	-0.059800	0.001379
17	3.4065	-0.012400	0.000313
18	3.5250	0.000000	0.000000
19	4.0921	-0.064000	0.001619
20	4.8914	-0.121300	0.003001
21	4.3506	-0.046800	0.000996
22	4.1823	-0.021100	0.000340
23	4.2330	-0.026600	0.000553
24	4.6965	-0.057400	0.001217
25	4.2981	-0.001875	-0.000092
26	5.5507	-0.065100	0.001299
27	5.8662	-0.080000	0.001647
28	6.5733	-0.118600	0.002388
29	5.6752	-0.060100	0.001042
30	5.0341	-0.027600	0.000480
31	4.3279	0.028500	-0.000772
32	4.3500	0.000000	0.000000
33	4.4089	-0.020700	0.000473
34	4.0621	0.036700	-0.000930
35	1.8658	0.180700	-0.004025
36	0.7257	0.208500	-0.004462
37	0.6257	0.179100	-0.003708
38	0.8618	0.145700	-0.003022
39	1.0643	0.126200	-0.002581
40	1.5238	0.102300	-0.002053
41	2.0426	0.053000	-0.000937
42	2.4779	0.033600	-0.000457
43	2.7208	0.020900	-0.000217
44	3.3578	0.002016	0.000247
45	3.8492	0.015400	-0.000247
46	4.0451	0.029800	-0.000638
47	3.7954	0.055200	-0.001254
48	5.1789	-0.014400	0.000103
49	6.2397	-0.009713	-0.000141
50	7.6547	-0.019200	-0.000037
51	6.2427	0.102100	-0.002409
52	7.1241	0.057300	-0.001177

rise in the entire term structure by one percentage point will result in a decrease of the logarithm of money balances by .0278, or in money balances by 2.74 percent. This effect on money balances is almost identical, as of course it should be, with the effect obtained by using the short-term interest rate, namely, a reduction in the logarithm of money balances by .0285, or in money balances by 2.81 percent for each percentage point rise in R_S.

It is more difficult to interpret the coefficients of a_1 and a_2. If either a_1 or a_2 alone rises while a_0 and the other one are held constant, that will raise the value of the interest rate for every maturity given by equation (A1), and that in turn can be expected to reduce log m, as the negative coefficients in equation (A3) indicate. However, that simply repeats the finding for R_S from equation (A2) and for a_0 from equation (A3). The more interesting question is the effect of an increase in the slope of the term structure while at the same time an appropriate *average* rate is constant, namely, an average rate obtained by weighting different maturities (really holding periods) by their importance in determining desired money balances.[64] The calculated coefficients of the parameters a_0, a_1, a_2 have imbedded in them the appropriate weighting function, but we do not see any way to extract it.

This question has particular interest because a theoretical analysis of money demand implies that an increase in the slope in this sense, which decreases short rates while increasing long rates by enough to keep the appropriate average rate constant, will *decrease* the quantity of money demanded. This is counterintuitive. Short-term assets are a closer substitute for money than long-term assets, and hence our intuitive expectation is that a decline in short rates would tend to raise the quantity of money demanded by more than the associated rise in long-term rates would decrease it. The counterintuitive result reflects the countervailing influence of the weights. In general, closer substitutability of short-term than of long-term assets for money will mean that they get a higher weight in the appropriate substitute portfolio, which means that, to keep the average yield constant, long-term rates will have to rise by more than short-term rates fall, which offsets the closer substitutability of short-term assets. Moreover, the greater the difference in substitutability, the greater the difference in weights. Nonetheless, it is far from intuitively clear that the *net* result must be in the direction of decreasing the demand for money, which is why an empirical test is of great interest.[65]

---

[64] For a fuller discussion, see M. Friedman, "Time Perspective in the Demand for Money," *Scandinavian Journal of Economics* 79, no. 4 (1977), equation (10) p. 408.

[65] Ibid., pp. 410–11.

Aside from the unavailability of the appropriate weighting function, a further complication is that the slope is itself a function of maturity unless $a_2 = 0$.

Let $\bar{\tau}$ be a weighted average maturity, using the (unknown) appropriate weights. Then the average yield for the appropriately weighted portfolio is

(A4) $$\bar{R} = a_0 + a_1\bar{\tau} + a_2\overline{\tau^2} = a_0 + a_1\bar{\tau} + a_2\,(\bar{\tau}^2 + \sigma_\tau^2).$$

Suppose we concentrate on the slope of equation (A1) at $\bar{\tau}$, which is

(A5) $$DR\,(\bar{\tau}) = a_1 + 2a_2\bar{\tau}.$$

Consider a change of ΔS in the slope of equation (A1) at $\bar{\tau}$ produced by a change in a_1, from a_1 to $a_1 + \Delta a_1$, and in a_2 to $a_2 + \Delta a_2$, so that

(A6) $$\Delta a_1 + 2(\Delta a_2)\,\bar{\tau} = \Delta S.$$

The condition that $\bar{R}$ be unchanged requires that

(A7) $$(\Delta a_1)\,\bar{\tau} + (\Delta a_2)\,(\bar{\tau}^2 + \sigma_\tau^2) = 0.$$

Solving equations (A6) and (A7) simultaneously gives

(A8) $$\Delta a_1 = -\frac{\bar{\tau}^2 + \sigma_\tau^2}{\bar{\tau}^2 - \sigma_\tau^2}\,\Delta S,$$

(A9) $$\Delta a_2 = \frac{\bar{\tau}}{\bar{\tau}^2 - \sigma_\tau^2}\,\Delta S.$$

Let b_1 be the coefficient of a_1 in equation (A3); b_2, the coefficient of a_2. Then the change in log m produced by a change of Δa_1 and Δa_2 in a_1 and a_2 is

(A10) $$\Delta \log m = b_1\Delta a_1 + b_2\Delta a_2\,.$$

To go further requires some assumption about the portfolio for which $\bar{\tau}$ and σ_τ^2 are to be calculated. We have considered two weighting patterns which are simple to handle mathematically because both are characterized by only a single parameter. One weighting pattern corresponds to a portfolio equally divided among all maturities from 0 to τ_o, or a weighting pattern of

(A11) $$w_1(\tau) = \frac{1}{\tau_o},$$

which has

(A12) $$\bar{\tau} = \frac{\tau_o}{2},$$

(A13) $$\sigma_\tau^2 = \frac{1}{12}\tau_o^2,$$

(A14) $$\Delta a_1 = -2\Delta S,$$

$$\text{(A15)} \qquad \Delta a_2 = \frac{3}{\tau_o}\Delta S.$$

The other weighting pattern involves triangular weights declining from 0 to τ_o, or a weighting pattern of

$$\text{(A16)} \qquad w_2(\tau) = \frac{2}{\tau_o}\left(1 - \frac{\tau}{\tau_o}\right),$$

which has

$$\text{(A17)} \qquad \bar{\tau} = \frac{\tau_o}{3}$$

$$\text{(A18)} \qquad \sigma_\tau^2 = \frac{\tau_o^2}{18}$$

and

$$\text{(A19)} \qquad \Delta a_1 = -3\Delta S_1$$

$$\text{(A20)} \qquad \Delta a_2 = \frac{6}{\tau_o}\Delta S.$$

The second pattern corresponds to the intuition that short-term securities should have a heavier weight than long-term securities.

It turns out that for both weighting patterns, and for the calculated values of b_1 and b_2, the effect on log m is the theoretically predicted negative effect for small τ_o. As τ_o rises, the negative effect becomes smaller in absolute value and ultimately turns positive. The value of τ_o at which it turns positive is 69.6 years, or a mean maturity of 34.8 years, for weighting pattern w_1; it is 92.8 years, or a mean maturity of 30.9 years, for weighting pattern w_2. These are far longer mean maturities than seem at all plausible for relevant actual portfolios—indeed, the mean maturity is longer than the longest maturities that we have used in calculating a_0, a_1, a_2. Hence, the theoretical implication about direction is amply confirmed.

What about magnitude of effect? Consider a rise in the slope at τ of .033 percentage points (.00033). That is equivalent to increasing the differential between maturities thirty years apart by one percentage point. And consider a value of $\bar{\tau}$ of 10 years, that is, a maximum maturity in the portfolio of twenty years for weighting pattern w_1 and of thirty years for weighting pattern w_2. That rise in slope, for fixed $\bar{R}$, would decrease desired m by

39 percent for weighting pattern w_1
46 percent for weighting pattern w_2.

These seem extremely large effects, hard to reconcile with the value of the coefficient of a_0 in equation (A3), or of R_S in equation (A2). However, the coefficients of a_1 and a_2 in equation (A3) are subject to large sampling errors, hence these estimates are subject to a large margin of error. To check this effect, we raised the absolute value of b_1 by one standard error and reduced the absolute value of b_2 by one standard error—both changes tending to raise $\Delta \log m$ algebraically. The result was still a reduction in desired m but by much smaller, and more plausible, amounts: by

16 percent for weighting pattern w_1
2 percent for weighting pattern w_2.

The inaccuracy of the estimate reflects the inaccuracy in the computed values of a_0, a_1, a_2 based as they are on only five observations, and these frequently not differing much. All in all, considering the element of chance error in our estimates, we regard these results as a strong confirmation of theoretical expectations.

A direct comparison of our results with those of Heller and Khan is not easy because of the different periods covered (1873 to 1975 versus 1960 to 1976); the different time unit (a cycle phase versus a quarter); the different form of the quadratic term structure equation (dependent variable R versus $\log R$); the different range of interest rates entering into the calculations (our one to thirty years versus their three months to twenty years); and the different demand equation (per capita real money and income versus aggregate real money and income, our inclusion and their exclusion of g_Y versus our exclusion and their inclusion of a lagged dependent variable).

In terms of direction, the coefficients of their counterparts of our a_0, a_1, a_2, are, like ours, all negative, though decidedly larger in absolute value when the difference between the use of R and the use of $\log R$ is allowed for, yet with roughly similar t values. In terms of the effect of an increase in slope holding $\overline{R}$ constant (in this case, $\overline{\log R}$, that is, the geometric mean), their results show a decrease in demand for τ_o less than 9.3 years for weighting pattern w_1, less than 12.35 years for weighting pattern w_2, or mean maturities of 4.6 or 4.1 years. It is perhaps not surprising that, given their short time unit and narrower range of maturities, these numbers are lower than those we found. In addition, the higher and more erratic inflation after 1960 than during most of our period could be expected to lead to considerably shorter-term portfolios, so, all in all, we regard their results as not inconsistent with ours.

Two recent studies by Bilson and Hale and by Allen and Hafer point out that Heller and Khan entered maturity incorrectly in computing the quadratic term structure. Bilson and Hale correct the error and also find that expressing the logarithm of the interest rate as a quadratic in the logarithm of maturity fits the term-structure data better than a quadratic in maturity. Allen and Hafer, using quarterly data for 1960–79, recompute both the Heller-Khan and the Bilson-Hale equations and also find that a cubic in the logarithm of maturity gives still better results.

The Bilson-Hale recomputation of the Heller-Khan equation strengthens its conformity with our results. The value of τ_o at which the effect on log m turns positive is 35.9 years for weighting pattern 1 and 48.1 years for weighting pattern 2, still less than the corresponding values for our equation but much closer to them.

The Bilson-Hale and Allen-Hafer use of the logarithms of interest rate and maturity rather than the interest rate and maturity complicates the comparison of their results with our own. However, it is straightforward to estimate the effect of a change in the slope of the term structure at the geometric mean maturity while keeping the geometric average yield constant. The results are that the effect of an increase in slope is in the theoretically predicted negative direction for weighting pattern 1 (weighting pattern 2) for τ_o less than 3.2 (5.1) months and between 7.4 (10.7) months and roughly 100 (300) years; it is in the opposite direction for τ_o between 3.2 (5.1) months and 7.4 (10.7) months, and greater than roughly 100 (300) years. These results, therefore, are in line with our own as well as with Heller and Khan's.

We have not attempted to analyze in the same way the Allen-Hafer equation using a cubic to approximate the term structure. However, the coefficients of the linear and quadratic terms (the counterparts of $b_1$and b_2 in equation A11) are almost identical with those for the Bilson-Hale equation. Hence, the results cannot differ much.