WESTERN CLASSICS

EDITED BY
[illegible] FRIEDMAN

STUDIES IN THE QUANTITY
THEORY OF [illegible]

AMARTYA K. SEN

COLLECTIVE CHOICE AND
SOCIAL WELFARE

STUDIES IN THE QUANTITY THEORY OF MONEY

EDITED BY MILTON FRIEDMAN
WITH ESSAYS BY
MILTON FRIEDMAN
PHILLIP CAGAN
JOHN J. KLEIN
EUGENE M. LERNER
RICHARD T. SELDEN

AMARTYA K. SEN

COLLECTIVE CHOICE AND SOCIAL WELFARE

CHINA SOCIAL SCIENCES PUBLISHING HOUSE
CHENGCHENG BOOKS LTD.

图书在版编目(CIP)数据

货币数量理论研究:英文/(美)弗里德曼著. -影印本.
-北京:中国社会科学出版社,1999.12
(西学基本经典·经济学类/西学基本经典工作委员会 编)
ISBN 7-5004-2644-5

Ⅰ.货… Ⅱ.弗… Ⅲ.弗里德曼,M. -经济思想-英文 Ⅳ.F091.353

中国版本图书馆 CIP 数据核字(1999)第 68421 号

版权总代理:TAO MEDIA INTERNATIONAL
(2790 19th Avenue,Ste.20,San Francisco,CA 94132 U.S.A)

西学基本经典·经济学类
西学基本经典工作委员会 编
中国社会科学出版社 出版发行
(北京鼓楼西大街甲 158 号 邮编 100720)
E-mail:5004@Peoplespace.net
诚成图书有限公司制作
中国建筑工业出版社密云印刷厂印刷
新华书店经销
1999 年 12 月第 1 版 1999 年 12 月第 1 次印刷
开本 880×1230 1/32 印张 184.5
经济学类全 12 册 定价:600.00 元

总策划 严 平 野 夫
项目策划 张自文 任建成

西学基本经典工作委员会

※ ※ ※ ※ ※

西学基本经典专家委员会

STUDIES IN THE QUANTITY THEORY OF MONEY

COLLECTIVE CHOICE AND SOCIAL WELFARE

STUDIES IN THE QUANTITY THEORY OF MONEY

EDITED BY MILTON FRIEDMAN
WITH ESSAYS BY
MILTON FRIEDMAN
PHILLIP CAGAN
JOHN J. KLEIN
EUGENE M. LERNER
RICHARD T. SELDEN

Table of Contents

I

The Quantity Theory of Money—A Restatement

MILTON FRIEDMAN

I

The Quantity Theory of Money—A Restatement

MILTON FRIEDMAN

The Quantity Theory of Money—A Restatement

THE quantity theory of money is a term evocative of a general approach rather than a label for a well-defined theory. The exact content of the approach varies from a truism defining the term "velocity" to an allegedly rigid and unchanging ratio between the quantity of money—defined in one way or another—and the price level—also defined in one way or another. Whatever its precise meaning, it is clear that the general approach fell into disrepute after the crash of 1929 and the subsequent Great Depression and only recently has been slowly re-emerging into professional respectability.

The present volume is partly a symptom of this re-emergence and partly a continuance of an aberrant tradition. Chicago was one of the few academic centers at which the quantity theory continued to be a central and vigorous part of the oral tradition throughout the 1930's and 1940's, where students continued to study monetary theory and to write theses on monetary problems. The quantity theory that retained this role differed sharply from the atrophied and rigid caricature that is so frequently described by the proponents of the new income-expenditure approach—and with some justice, to judge by much of the literature on policy that was spawned by quantity theorists. At Chicago, Henry Simons and Lloyd Mints directly, Frank Knight and Jacob Viner at one remove, taught and developed a more subtle and relevant version, one in which the quantity theory was connected and integrated with general price theory and became a flexible and sensitive tool for interpreting movements in aggregate economic activity and for developing relevant policy prescriptions.

To the best of my knowledge, no systematic statement of this theory as developed at Chicago exists, though much can be read between the lines of Simons' and Mints's writings. And this is as it should be, for the Chicago tradition was not a rigid system, an unchangeable orthodoxy, but a way of looking at things. It was a theoretical approach that insisted that money does matter—that any interpretation of short-term movements in economic activity is likely to be seriously at fault if it neglects monetary changes and repercussions and if it leaves unexplained why people are willing to hold the particular nominal quantity of money in existence.

The purpose of this introduction is not to enshrine—or, should I say, inter—a definitive version of the Chicago tradition. To suppose that one

could do so would be inconsistent with that tradition itself. The purpose is rather to set down a particular "model" of a quantity theory in an attempt to convey the flavor of the oral tradition which nurtured the remaining essays in this volume. In consonance with this purpose, I shall not attempt to be exhaustive or to give a full justification for every assertion.

1. The quantity theory is in the first instance a theory of the *demand* for money. It is not a theory of output, or of money income, or of the price level. Any statement about these variables requires combining the quantity theory with some specifications about the conditions of supply of money and perhaps about other variables as well.

2. To the ultimate wealth-owning units in the economy, money is one kind of asset, one way of holding wealth. To the productive enterprise, money is a capital good, a source of productive services that are combined with other productive services to yield the products that the enterprise sells. Thus the theory of the demand for money is a special topic in the theory of capital; as such, it has the rather unusual feature of combining a piece from each side of the capital market, the supply of capital (points 3 through 8 that follow), and the demand for capital (points 9 through 12).

3. The analysis of the demand for money on the part of the ultimate wealth-owning units in the society can be made formally identical with that of the demand for a consumption service. As in the usual theory of consumer choice, the demand for money (or any other particular asset) depends on three major sets of factors: (*a*) the total wealth to be held in various forms—the analogue of the budget restraint; (*b*) the price of and return on this form of wealth and alternative forms; and (*c*) the tastes and preferences of the wealth-owning units. The substantive differences from the analysis of the demand for a consumption service are the necessity of taking account of intertemporal rates of substitution in (*b*) and (*c*) and of casting the budget restraint in terms of wealth.

4. From the broadest and most general point of view, total wealth includes all sources of "income" or consumable services. One such source is the productive capacity of human beings, and accordingly this is one form in which wealth can be held. From this point of view, "the" rate of interest expresses the relation between the stock which is wealth and the flow which is income, so if Y be the total flow of income, and r, "the" interest rate, total wealth is

$$W = \frac{Y}{r}. \tag{1}$$

Income in this broadest sense should not be identified with income as it is ordinarily measured. The latter is generally a "gross" stream with respect

to human beings, since no deduction is made for the expense of maintaining human productive capacity intact; in addition, it is affected by transitory elements that make it depart more or less widely from the theoretical concept of the stable level of consumption of services that could be maintained indefinitely.

5. Wealth can be held in numerous forms, and the ultimate wealth-owning unit is to be regarded as dividing his wealth among them (point [*a*] of 3), so as to maximize "utility" (point [*c*] of 3), subject to whatever restrictions affect the possibility of converting one form of wealth into another (point [*b*] of 3). As usual, this implies that he will seek an apportionment of his wealth such that the rate at which he *can* substitute one form of wealth for another is equal to the rate at which he is just willing to do so. But this general proposition has some special features in the present instance because of the necessity of considering flows as well as stocks. We can suppose all wealth (except wealth in the form of the productive capacity of human beings) to be expressed in terms of monetary units at the prices of the point of time in question. The rate at which one form can be substituted for another is then simply $1.00 worth for $1.00 worth, regardless of the forms involved. But this is clearly not a complete description, because the holding of one form of wealth instead of another involves a difference in the composition of the income stream, and it is essentially these differences that are fundamental to the "utility" of a particular structure of wealth. In consequence, to describe fully the alternative combinations of forms of wealth that are available to an individual, we must take account not only of their market prices—which except for human wealth can be done simply by expressing them in units worth $1.00—but also of the form and size of the income streams they yield.

It will suffice to bring out the major issues that these considerations raise to consider five different forms in which wealth can be held: (i) money (M), interpreted as claims or commodity units that are generally accepted in payment of debts at a fixed nominal value; (ii) bonds (B), interpreted as claims to time streams of payments that are fixed in nominal units; (iii) equities (E), interpreted as claims to stated pro-rata shares of the returns of enterprises; (iv) physical non-human goods (G); and (v) human capital (H). Consider now the yield of each.

(i) Money may yield a return in the form of money, for example, interest on demand deposits. It will simplify matters, however, and entail no essential loss of generality, to suppose that money yields its return solely in kind, in the usual form of convenience, security, etc. The magnitude of this return in "real" terms per nominal unit of money clearly

depends on the volume of goods that unit corresponds to, or on the general price level, which we may designate by P. Since we have decided to take \$1.00 worth as the unit for each form of wealth, this will be equally true for other forms of wealth as well, so P is a variable affecting the "real" yield of each.

(ii) If we take the "standard" bond to be a claim to a perpetual income stream of constant nominal amount, then the return to a holder of the bond can take two forms: one, the annual sum he receives—the "coupon"; the other, any change in the price of the bond over time, a return which may of course be positive or negative. If the price is expected to remain constant, then \$1.00 worth of a bond yields r_b per year, where r_b is simply the "coupon" sum divided by the market price of the bond, so $1/r_b$ is the price of a bond promising to pay \$1.00 per year. We shall call r_b the market bond interest rate. If the price is expected to change, then the yield cannot be calculated so simply, since it must take account of the return in the form of expected appreciation or depreciation of the bond, and it cannot, like r_b, be calculated directly from market prices (so long, at least, as the "standard" bond is the only one traded in).

The nominal income stream purchased for \$1.00 at time zero then consists of

$$r_b(0) + r_b(0)\, d\frac{\left(\frac{1}{r_b(t)}\right)}{dt} = r_b(0) - \frac{r_b(0)}{r_b^2(t)} \cdot \frac{d\, r_b(t)}{dt}, \qquad (2)$$

where t stands for time. For simplicity, we can approximate this functional by its value at time zero, which is

$$r_b - \frac{1}{r_b}\frac{d\, r_b}{dt}. \qquad (3)$$

This sum, together with P already introduced, defines the real return from holding \$1.00 of wealth in the form of bonds.

(iii) Analogously to our treatment of bonds, we may take the "standard" unit of equity to be a claim to a perpetual income stream of constant "real" amount; that is, to be a standard bond with a purchasing-power escalator clause, so that it promises a perpetual income stream equal in nominal units to a constant number times a price index, which we may, for convenience, take to be the same price index P introduced in (i).[1] The nominal return to the holder of the equity can then be regarded as taking three forms: the constant nominal amount he would receive per year in

1. This is an oversimplification, because it neglects "leverage" and therefore supposes that any monetary liabilities of an enterprise are balanced by monetary assets.

the absence of any change in P; the increment or decrement to this nominal amount to adjust for changes in P; and any change in the nominal price of the equity over time, which may of course arise from changes either in interest rates or in price levels. Let r_e be the market interest rate on equities defined analogously to r_b, namely, as the ratio of the "coupon" sum at any time (the first two items above) to the price of the equity, so $1/r_e$ is the price of an equity promising to pay \$1.00 per year if the price level does not change, or to pay

$$\frac{P(t)}{P(0)} \cdot 1$$

if the price level varies according to $P(t)$. If $r_e(t)$ is defined analogously, the price of the bond selling for $1/r_e(0)$ at time 0 will be

$$\frac{P(t)}{P(0)\, r_e(t)}$$

at time t, where the ratio of prices is required to adjust for any change in the price level. The nominal stream purchased for \$1.00 at time zero then consists of

$$r_e(0) \cdot \frac{P(t)}{P(0)} + \frac{r_e(0)}{P(0)} \cdot d\frac{\left[\frac{P(t)}{r_e(t)}\right]}{dt} = r_e(0) \cdot \frac{P(t)}{P(0)} + \frac{r_e(0)}{r_e(t)} \cdot \frac{1}{P(0)} \cdot \frac{dP(t)}{dt} - \frac{P(t)}{P(0)} \cdot \frac{r_e(0)}{r_e^2(t)} \cdot \frac{d\,r_e(t)}{dt}. \tag{4}$$

Once again we can approximate this functional by its value at time zero, which is

$$r_e + \frac{1}{P}\frac{dP}{dt} - \frac{1}{r_e}\frac{d\,r_e}{dt}. \tag{5}$$

This sum, together with P already introduced, defines the "real" return from holding \$1.00 of wealth in the form of equities.

(iv) Physical goods held by ultimate wealth-owning units are similar to equities except that the annual stream they yield is in kind rather than in money. In terms of nominal units, this return, like that from equities, depends on the behavior of prices. In addition, like equities, physical goods must be regarded as yielding a nominal return in the form of appreciation or depreciation in money value. If we suppose the price level P, introduced earlier, to apply equally to the value of these physical goods, then, at time zero,

$$\frac{1}{P}\frac{dP}{dT} \tag{6}$$

is the size of this nominal return per $1.00 of physical goods.[2] Together with P, it defines the "real" return from holding $1.00 in the form of physical goods.

(v) Since there is only a limited market in human capital, at least in modern non-slave societies, we cannot very well define in market prices the terms of substitution of human capital for other forms of capital and so cannot define at any time the physical unit of capital corresponding to $1.00 of human capital. There are some possibilities of substituting non-human capital for human capital in an individual's wealth holdings, as, for example, when he enters into a contract to render personal services for a specified period in return for a definitely specified number of periodic payments, the number not depending on his being physically capable of rendering the services. But, in the main, shifts between human capital and other forms must take place through direct investment and disinvestment in the human agent, and we may as well treat this as if it were the only way. With respect to this form of capital, therefore, the restriction or obstacles affecting the alternative compositions of wealth available to the individual cannot be expressed in terms of market prices or rates of return. At any one point in time there is some division between human and non-human wealth in his portfolio of assets; he may be able to change this over time, but we shall treat it as given at a point in time. Let w be the ratio of non-human to human wealth or, equivalently, of income from non-human wealth to income from human wealth, which means that it is closely allied to what is usually defined as the ratio of wealth to income. This is, then, the variable that needs to be taken into account so far as human wealth is concerned.

6. The tastes and preferences of wealth-owning units for the service streams arising from different forms of wealth must in general simply be taken for granted as determining the form of the demand function. In order to give the theory empirical content, it will generally have to be supposed that tastes are constant over significant stretches of space and time. However, explicit allowance can be made for some changes in tastes in so far as such changes are linked with objective circumstances. For example, it seems reasonable that, other things the same, individuals want

2. In principle, it might be better to let P refer solely to the value of the services of physical goods, which is essentially what it refers to in the preceding cases, and to allow for the fact that the prices of the capital goods themselves must vary also with the rate of capitalization, so that the prices of services and their sources vary at the same rate only if the relevant interest rate is constant. I have neglected this refinement for simplicity; the neglect can perhaps be justified by the rapid depreciation of many of the physical goods held by final wealth-owning units.

to hold a larger fraction of their wealth in the form of money when they are moving around geographically or are subject to unusual uncertainty than otherwise. This is probably one of the major factors explaining a frequent tendency for money holdings to rise relative to income during wartime. But the extent of geographic movement, and perhaps of other kinds of uncertainty, can be represented by objective indexes, such as indexes of migration, miles of railroad travel, and the like. Let u stand for any such variables that can be expected to affect tastes and preferences (for "utility" determining variables).

7. Combining 4, 5, and 6 along the lines suggested by 3 yields the following demand function for money:

$$M = f\left(P,\ r_b - \frac{1}{r_b}\frac{dr_b}{dt},\ r_e + \frac{1}{P}\frac{dP}{dt} - \frac{1}{r_e}\frac{dr_e}{dt},\ \frac{1}{P}\frac{dP}{dt};\ w;\ \frac{Y}{r};\ u\right). \quad (7)$$

A number of observations are in order about this function.

(i) Even if we suppose prices and rates of interest unchanged, the function contains three rates of interest: two for specific types of assets, r_b and r_e, and one intended to apply to all types of assets, r. This general rate, r, is to be interpreted as something of a weighted average of the two special rates plus the rates applicable to human wealth and to physical goods. Since the latter two cannot be observed directly, it is perhaps best to regard them as varying in some systematic way with r_b and r_e. On this assumption, we can drop r as an additional explicit variable, treating its influence as fully taken into account by the inclusion of r_b and r_e.

(ii) If there were no differences of opinion about price movements and interest-rate movements, and bonds and equities were equivalent except that the former are expressed in nominal units, arbitrage would of course make

$$r_b - \frac{1}{r_b}\frac{dr_b}{dt} = r_e + \frac{1}{P}\frac{dP}{dt} - \frac{1}{r_e}\frac{dr_e}{dt}, \quad (8)$$

or, if we suppose rates of interest either stable or changing at the same percentage rate,

$$r_b = r_e + \frac{1}{P}\frac{dP}{dt}, \quad (9)$$

that is, the "money" interest rate equal to the "real" rate plus the percentage rate of change of prices. In application the rate of change of prices must be interpreted as an "expected" rate of change and differences of opinion cannot be neglected, so we cannot suppose (9) to hold; indeed,

one of the most consistent features of inflation seems to be that it does not.[3]

(iii) If the range of assets were to be widened to include promises to pay specified sums for a finite number of time units—"short-term" securities as well as "consols"—the rates of change of r_b and r_e would be reflected in the difference between long and short rates of interest. Since at some stage it will doubtless be desirable to introduce securities of different time duration (see point 23 below), we may simplify the present exposition by restricting it to the case in which r_b and r_e are taken to be stable over time. Since the rate of change in prices is required separately in any event, this means that we can replace the cumbrous variables introduced to designate the nominal return on bonds and equities simply by r_b and r_e.

(iv) Y can be interpreted as including the return to all forms of wealth, including money and physical capital goods owned and held directly by ultimate wealth-owning units, and so Y/r can be interpreted as an estimate of total wealth, only if Y is regarded as including some imputed income from the stock of money and directly owned physical capital goods. For monetary analysis the simplest procedure is perhaps to regard Y as referring to the return to all forms of wealth other than the money held directly by ultimate wealth-owning units, and so Y/r as referring to total remaining wealth.

8. A more fundamental point is that, as in all demand analyses resting on maximization of a utility function defined in terms of "real" magnitudes, this demand equation must be considered independent in any essential way of the nominal units used to measure money variables. If the unit in which prices and money income are expressed is changed, the amount of money demanded should change proportionately. More technically, equation (7) must be regarded as homogeneous of the first degree in P and Y, so that

$$f\left(\lambda P,\ r_b,\ r_e,\ \frac{1}{P}\frac{dP}{dt};\ w;\ \lambda Y;\ u\right) = \lambda f\left(P,\ r_b,\ r_e,\ \frac{1}{P}\frac{dP}{dt};\ w;\ Y;\ u\right). \tag{10}$$

where the variables within the parentheses have been rewritten in simpler form in accordance with comments 7 (i) and 7 (iii).

3. See Reuben Kessel, "Inflation: Theory of Wealth Distribution and Application in Private Investment Policy" (unpublished doctoral dissertation, University of Chicago).