WHAT CLASSICAL AND NEOCLASSICAL MONETARY THEORY REALLY WAS*

PAUL A. SAMUELSON *Massachusetts Institute of Technology*

Qu'était au juste la théorie monétaire classique et néo-classique ? L'article présente le point de vue d'un de ceux qui ont contribué à la théorie monétaire classique et néo-classique. Comme l'auteur à cru pendant les années 1932 à 1937, cette théorie (jamais formulée d'une façon formelle sous forme d'un système d'équations) supposait, sans en contester le bien-fondé, qu'en longue période le volume monétaire n'avait aucune importance une fois que l'économie considérée était devenu une économie monétaire. Toutefois, la théorie n'allait pas jusqu'à prétendre qu'une économie monétaire et une économie de troc fussent identiques même si l'on supposait des goûts, des connaissances techniques et des quantités de facteurs de production identiques.

Certains tenants ont aussi postulé que ces facteurs réels affectaient les prix et les niveaux de production relatifs, alors que le volume monétaire affectait le niveau absolu des prix. Il y avait ainsi deux dichotomies au lieu d'une, mais la seconde dichotomie n'a jamais été prise tellement au sérieux par qui que ce soit. Elle constituait plûtot une simplification provisoire. L'essentiel de la formulation présente est d'inclure la monnaie dans la fonction d'utilité, puis de considérer la fonction comme jouissant de la propriété d'homogénéité suivant laquelle un doublement de tous les prix et de la monnaie n'avantage personne. Par conséquent, lorsque chacun pèse la commodité de détenir de la monnaie en comparaison de son coût en intérêt, sa fonction de démande pour les biens est indépendante du niveau absolu des prix, mais sa démande pour la monnaie est proportionnelle aux augmentations balancées de tous les prix.

Quoi qu'il en soit, les deux dichotomies sont légitimes pourvu que les modèles sousjacents soient définis en conséquence. L'auteur présente ensuite un modèle qui démontre le bien-fondé d'une dichotomie entre « les éléments réels » et « l'élément monétaire qui ne détermine que le niveau absolu des prix ». L'auteur prétend que les meilleurs auteurs néo-classiques avaient intuitivement ce modèle en tête, même s'ils ne l'ont jamais explicité ou publié. L'auteur termine son article par une discussion des contributions de Lange, Patinkin, et Archibald et Lipsey.

To know your own country you must have travelled abroad. To understand modern economics it is good to have lived long enough to have escaped competent instruction in its mysteries. When Archibald and Lipsey try to draw for Patinkin a picture of what a "classical" monetary theorist believed in, they are pretty much in the position of a man who, looking for a jackass, must say to himself, "If I were a jackass, where would I go?"

Mine is the great advantage of having once been a jackass. From 2 January 1932 until an indeterminate date in 1937, I was a classical monetary theorist. I do not have to look for the tracks of the jackass embalmed in old journals and monographs. I merely have to lie down on the couch and recall in tranquillity, upon that inward eye which is the bliss of solitude, what it was that I believed between the ages of 17 and 22. This puts me in the same advantageous position that Pio Nono enjoyed at the time when the infallibility of the Pope was being enunciated. He could say, incontrovertibly, "Before I was Pope, I believed he was infallible. Now that I am Pope, I can *feel* it."

Essentially, we believed that in the longest run and in ideal models the

*I owe thanks to the National Science Foundation.

Canadian Journal of Economics/Revue canadienne d'Economique, I, no. 1
February/février 1968

amount of money did not matter. Money could be "neutral" and in many conditions the hypothesis that it was could provide a good first or last approximation to the facts. To be sure, Hume, Fisher, and Hawtrey had taught us that, under dynamic conditions, an increase in money might lead to "money illusion" and might cause substantive changes—e.g., a shift to debtor-entrepreneurs and away from creditor-rentiers, a forced-saving shift to investment and away from consumption, a lessening of unemployment, a rise in wholesale prices relative to sticky retail prices and wage rates, *et cetera.*

But all this was at a second level of approximation, representing relatively transient aberrations. Moreover, this tended to be taught in applied courses on business cycles, money and finance, and economic history rather than in courses on pure theory. In a real sense there *was* a dichotomy in our minds; we were schizophrenics. From 9 to 9:50 a.m. we presented a simple quantity theory of neutral money. There were then barely ten minutes to clear our palates for the 10 to 10:50 discussion of how an engineered increase in M would help the economy. In mid-America in the mid-1930s, we neoclassical economists tended to be mild inflationists, jackasses crying in the wilderness and resting our case essentially on sticky prices and costs, and on expectations.

Returning to the 9 o'clock hour, we thought that *real* outputs and inputs and price ratios depended essentially in the longest run on real factors, such as tastes, technology, and endowments. The stock of money we called M (or, to take account of chequable bank deposits, we worked in effect with a velocity-weighted average of M and M'; however, a banking system with fixed reserve and other ratios would yield M' proportional to M, so M alone would usually suffice). An increase in M—usually we called it a doubling on the ground that after God created unity he created the second integer—would cause a proportional increase in *all* prices (tea, salt, female labour, land rent, share or bond prices) and values (expenditure on tea or land, share dividends, interest income, taxes). You will hardly believe it, but few economists in those days tried to write down formal equations for what they were thinking. Had we been asked to choose which kinds of equation system epitomized our thinking, I believe at first blush we would have specified:

A. Write down a system of real equations involving *real* outputs and inputs, and *ratios* of prices (values), and depending essentially on real tastes, technologies, market structures, and endowments. Its properties are invariant to change in the stock of money M.

B. Then append a fixed-supply-of-M equation that pins down (or up) the absolute price level, determining the scale factor that was essentially indeterminate in set A. This could be a quantity equation of exchange—$MV = PQ$—or some other non-homogeneous equation. More accurately, while A involves homogeneity of degree zero in *all Ps*, B involves homogeneity of degree 1 of *Ps* in terms of M.

I have purposely left the above paragraphs vague. For I doubt that the typical good classical monetary theorist had more definite notions about the *mathematics* of his system.

Moreover, I must leave room for an essential strand in our thinking. Our expositions always began with barter and worked our fundamental pricing in barter models. But then we, sensibly, pointed out the *real* inconvenience of barter and the real convenience of an abstract unit of money. Here we made explicit and tacit reference to the real facts of brokerage or transaction charges, of uncertainties of income and outgo, and so on. In short, we did have a primitive inventory theory of money holding, but we were careful to note that true money—unlike pearls, paintings, wine, and coffee—is held only for the *ultimate exchange* work it can do, which depends upon the scale of *all Ps* in a special homogeneous way.

So there was another dichotomy in our minds, a very legitimate one. We had, so to speak, *qualitative* and *quantitative* theories of money. According to our qualitative theory, money was not neutral; it made a big difference. Pity the country that was still dependent upon barter, for it would have an inefficient economic system. But once this qualitative advantage had been realized by the adoption of market structures using M, the *quantitative* level of M was of no particular significance (except for indicated transient states and uninteresting resource problems involved in gold mining or mint printing). We liked the image of John Stuart Mill that money is the *lubricant* of industry and commerce. As even women drivers know, lubrication is important. But M is quantitatively a special lubricant: a drop will do as well as a poolful. So an even better image was the post-Mill one: money is like a catalyst in a chemical reaction, which makes the reaction go faster and better, but which, like the oil in the widow's cruse, is never used up. To push the analogy beyond endurance, only an iota of catalyst is needed for the process.

What I have just said makes it unmistakably clear that a classical monetary theorist would not go the stake for the belief that the real set of equations A are independent of M, depending essentially only on price ratios as in barter. If time were short on a quiz, I might carelessly write down such an approximation. But if asked specifically the question "Is Set A really independent of M?" I and my classmates would certainly answer "No" and we would cite the qualitative aspects mentioned earlier.

In a moment we shall see that this considered qualitative view requires that *M enter quantitatively in Set A in certain specified homogeneous ways.* But first let us investigate how those of us who were mathematically inclined would have handled the Set A and Set B problem. The economists interested in mathematics tended to be specialists in value theory. They had a big job just to describe the real relations of A, whether under barter or otherwise. They wanted to simplify their expositions, to sidestep extraneous complication. Hence, many would have followed the practice (which I seem to connect with Cassel's name, at least) of writing Set A purely in barter terms, and essentially giving enough equations to determine real quantities and price ratios—as follows:

$$\text{A}' \quad f_i(Q_1, \ldots, Q_n, P_1, \ldots, P_n) = 0 \ (i = 1, 2, \ldots, 2\,n)$$

where there are n inputs or outputs, with n prices. However, the f_i functions are made to be homogeneous of degree zero in all the Ps, and, luckily, the $2n$ functions f_i are required to involve one of them as being dependent on the other, thus avoiding an overdetermination of the $2n$ functions. This homogeneity and dependence postulate enables us to write A′ in the equivalent form:

$$\text{A}' \quad f_i(Q_1, \ldots, Q_n, \lambda P_1, \ldots, \lambda P_n)_\lambda \equiv 0 \ (i = 1, 2, \ldots, 2n).$$

This formulation does not contain price ratios explicitly. But since λ is arbitrary, it can be set equal to $1/P_1$ to give us price ratios, $P_i/_1$. Or if you have an interest in some kind of average of prices, say $\pi(P_1, \ldots, P_n) = \pi(P)$, where π is a homogeneous function of degree one, you can rewrite A′ in terms of ratios $P_i/\pi(P)$ alone, by suitable choice of λ. Hence, Set A′ involves $2n-1$ independent functions which hopefully determine a unique (or multiple) solution to the $2n - 1$ real variables $(Q_1, \ldots, Q_n, P_2/P_1, \ldots, P_n/P_1)$. With the special structure of A′, we are now free to add any non-homogeneous B′ we like, of the following types:

$P_1 = 1$, good 1 being taken as numéraire, or
B′ $P_1 + P_2 = 3.1416$, or
$P_1 + P_2 + \ldots + P_n = 1$, or
$P_1[(Q_1^* + (P_2^*/P_1)Q_2^* + \ldots + (P_n^*/P_1)Q_n\text{o}] = \overline{M}$, Fisher's Constant,
where Q_i^*, $(P_i/P_1)^*$ are solutions of A′.

Of course, the last of these looks like the Fisher-Marshall formulation of the "quantity equation of exchange." But, since some Q_i are inputs, my way of writing it recognizes the realistic fact that money is needed to pay factors as well as to move goods.[1]

I do not defend this special A′, B′ formulation. I am sure it was often used. And even today, if I am behind in my lectures, I resort to it in courses on pure theory. But we should admit that it is imperfect. And we should insist that the classical writers, when they did full justice to their own views, did not believe that this formulation was more than a provisional simplification.

What is a minimal formulation of (A, B) that does do full justice? I am sure that I personally, from 1937 on at least, had a correct vision of the proper version. It is as if to understand Gary, Indiana, I had to travel to Paris. I began to understand neoclassical economics only after Keynes' *General Theory* shook me up. But I am sure that I was only learning to articulate what was intuitively felt by such ancients as Ricardo, Mill, Marshall, Wicksell, and Cannan. I regret that I did not then write down a formal set of equations. I did discuss the present issue at the Econometric Society meetings of 1940, of which only an incomplete abstract appeared, and also at its 1949 meetings, where W. B. Hickman, Leontief and others spoke; and there are fragmentary similar remarks in half a dozen of my writings of twenty years ago. The nub of the matter is contained in my 1947 specification[2] that the utility function contain in it, along with physical quantities of good consumed, the stock of M and all money Ps, being homogeneous of degree zero in $(M, P_1, \ldots, P_n)$ in recognition of money's peculiar "neutral" quantitative properties.

Frankly, I was repelled by the abstract level at which Oskar Lange, Hicks, and others carried on their discussion of Say's Law, staying at the level of equation counting and homogeneity reckoning, without entering into the

[1]An equation like the last one could be split into two equations without altering the meaning:
B'_1 $1/FC\ \Sigma^n{}_{j=1}\ (P_j/M)Q_j^* = 1$
B'_2 $M = M$, a prescribed total. The important thing to note is that B'_1, even if it looks a little like some A′ equations, is completely decomposable from the set A′.
[2]P. A. Samuelson, *Foundations of Economic Analysis* (Cambridge, Mass., 1947), 119.

concrete character of the models. And this was one of the few continuing controversies of economics from which I steadfastly abstained.

For the rest of this discussion, what I propose to do is to get off the couch and go to the blackboard and write down an organized picture of what we jackasses implicitly believed back in the bad old days.

The way things are

I abstract heroically. We are all exactly alike. We live forever, We are perfect competitors and all-but-perfect soothsayers. Our inelastic labour supply is fully employed, working with inelastically supplied Ricardian land and (possibly heterogeneous) capital goods. We have built-in Pigou-Böhm rates of subjective time preference, discounting each next-year's independent utility by the constant factor $1/(l + \rho)$, $\rho > 0$. We are in long-run equilibrium without technical change or population growth: the stock of capital goods has been depressed to the point where all own-interest-rates yielded by production are equal to r, the market rate of interest; in turn, r is equal to the subjective interest rate ρ, this being the condition for our propensity to consume being 100 per cent of income, with zero net capital formation.

We equally own land, and such capital goods as machinery and material stocks. We own, but legally cannot sell, our future stream of labour earnings. We hold cash balances, because we are *not* perfect soothsayers when it comes to the uncertainty of the timing of our in-and-out-payments, which can be assumed to follow certain probability laws in the background; this lack of synchronization of payments plus the indivisible costs of transactions (brokerage charges, need for journal entries, spread between bid and ask when earning assets are converted into or out of cash, etc.) requires us to hold money. To keep down inessential complications, while not omitting Hamlet from the scenario, I am neglecting the need for cash balances for corporations; it is as if consumer families alone need cash balances for their final consumption purchases, whereas in real life cash is needed at every vertical stage of the production process. Later we can allow our holdings of earning assets—titles to land and machines—to economize on our need for M balances, just as does the prospect of getting wage increases.

Our system is assumed to come into long-run equilibrium. This equilibrium can be deduced to be unique if we add to our extreme symmetry assumptions the conventional strong convexity assumptions of neoclassical theorizing—constant returns to scale with smooth diminishing returns to proportions, quasi-concave ordinal utility functions that guarantee diminishing marginal rates of substitution, and so on.

We should be able to *prove rigorously* what is probably intuitively obvious—doubling all M will exactly double *all* long-run prices and values, and this change in the absolute price level will have absolutely no effect on real output-inputs, on price ratios or terms of trade, on interest rate and factor shares generally.

For this system, it is not merely the case that tautological quantity equations of exchange can be written down. Less trivially, a simple "quantity theory of prices and money" holds exactly for the long-run equilibrium model. Although Patinkin has doubts about the propriety of the concept, I think our meaning was unambiguous—and unobjectionable—when we used to say that the "demand curve for money" (traced out by shifts in the vertical supply curve of M) plotted in a diagram containing, on the x axis, M and, on the Y axis, the "value of money," (as measured by the reciprocal of *any* absolute money price $1/P_i$ or any average price level) would be a rectangular hyperbola with a geometrical Marshallian elasticity of exactly minus one.

To prove this I write down the simplest possible set of equations. These do split up into two parts, showing that there is a legitimate "dichotomy" between "real elements" and "monetary elements which determine only the absolute level of prices." Call these two parts A and B. Now this legitimate dichotomy will not be identical with the over-simple dichotomy of A′ and B′ mentioned earlier. If Patinkin insists upon the difference, I am in complete agreement with him. If he should prefer not to call the (A, B) split a dichotomy, that semantic issue is not worth arguing about so long as enough words are used to describe exactly what the (A, B) split is, and how it differs from the (A′, B′) split. If Patinkin insists on saying that my A equations do have in them a "real balance effect," I see no harm in that—even though, as will be seen, my formulation of A need involve no use of an average price index, and hence no need to work with a "deflated M" that might be called a real balance. Peculiarly in the abstract neoclassical model with its long-run strong homogeneity properties, all Ps move together in strict proportion when M alone changes and hence no index-number approximations are needed. By the same token, they do absolutely no harm: Patinkin is entitled to use any number of average price concepts and real-balance concepts he wishes. If Patinkin wishes to say that the principal neoclassical writers (other than Walras) had failed to *publish* a clear and unambiguous account of the (A, B) equation such as I am doing here, I would agree, and would adduce the worth and novelty of Patinkin's own book and contributions. On the other hand, the present report on my recollections claims that the best neoclassical writers did *perceive* at the intuitive level the intrinsic content of the (A, B) dichotomy which I am about to present. All the more we should regret that no one fully set down these intuitions thirty years ago!

Now what about Archibald and Lipsey?[3] I want to avoid semantic questions as to what is meant by real-balance effects being operative. If they claim that the (A′, B′) dichotomy does justice to the tacit neoclassical models of 1930, I think they are wrong. If they think an (A′, B′) dichotomy does justice to a reasonably realistic long-run model of a monetary economy, I think they are also wrong. Whether, as a *tour de force*, some special, flukey (A′, B′) model might be found to give a representation of some monetary economy is a possibility that I should hate to deny in the abstract; but I should be surprised if this issue turned out to be an interesting one to linger on or to

[3]Don Patinkin, in his *Money, Interest, and Prices* (New York, 1966), summarizes his path-breaking writings on money over the last twenty years. For a critique of aspects of its first (1954) edition, see Archibald and Lipsey (*Review of Economic Studies*, XX) and articles in subsequent numbers of that journal.

debate. For what a casual opinion is worth, it is my impression that Patinkin's general position—which I interpret to be essentially identical to my (A, B) dichotomy *and* to the tacit neoclassical theory of my youth—is left impregnable to recent attacks on it. There is one, and only one, legitimate dichotomy in neoclassical monetary theory.

Abjuring further doctrinal discussion, I proceed now to the equations of my simplest system.

Structure of the model

1. PRODUCTION RELATIONS

To keep down inessentials, let land, T, real capital, K (assumed homogeneous merely as a preliminary to letting K stand for a vector of heterogeneous capital goods), and labour, L, produce real output which, because of similarity of production factors in all sectors, can be split up into the linear sum of different physical consumption goods $\pi_1 q_1 + \ldots + \pi_n q_n$ and net capital formation $\dot{K}(=dK/dt)$, namely: $\dot{K} + \pi_1 q_1 + \pi_2 q_2 + \ldots + \pi_m q_m = f(K, \bar{L}, \bar{T})$ where F is a production function of the Ramsey-Solow type, homogeneous of first degree, and where the π_i are constants, representing marginal costs of the ith goods relative to machines. From this function, we can deduce all factor prices and commodity prices relative to the price of the capital good P_K, namely:

$$A_{I,1} \quad \frac{P_i}{P_K} = \pi_i \qquad (i = 1, 2, \ldots, n)$$

$$A_{I,2} \quad \frac{W}{P_K} = \frac{\partial F(K, \bar{L}, \bar{T})}{\partial L}, \quad \text{the marginal productivity wage,}$$

$$\frac{R}{P_K} = \frac{\partial F(K, \bar{L}, \bar{T})}{\partial T}, \quad \text{the marginal productivity rent,}$$

$$r = \frac{\partial F(K, \bar{L}, \bar{T})}{\partial K}, \quad \text{the marginal productivity interest rate.}$$

Bars are put over L and T because their supplies are assumed to be fixed. To determine the unknown stock of capital K we need:

$r = \rho$, the subjective time preference parameter;[4]

A_{II} $\dot{K} = 0$, the implied steady-state long-run equilibrium condition;

$r = R/P_T$, the implicit capitalization equation for the price of land.

Hence, $\rho = \partial F(\bar{K}, \bar{L}, T)/\partial K$ henceforth gives us our fixed $\bar{K}$.

The above relationships determine for the representative man the wage and interest income (inclusive of land rentals expressed as interest on land values) which he can spend on the $(q_1, q_2, \ldots, q_n)$ goods and on holding of M cash balances which bear no interest and thus cost their opportunity costs in terms of interest forgone (or, to a net borrower, the interest on borrowings). What

[4]In unpublished memos and lectures, using a Ramsey maximum analysis I have shown how the long-run steady-state condition where $r = \rho$ is approached so that $\dot{K}$ (or $K^{t+1} - K^t$) is zero. The steady-state analysis of U $(q$: $M, \ldots)$ here is shorthand for the perpetual stream $\Sigma_0^\infty U\,(q^t;\, M^{t+1}, \ldots)/(1+p)^t$, etc. My colleague, Professor Miguel Sidrouski, has independently arrived at such dynamic formulations.

motive is there for holding any M? As I pointed out in *Foundations*, one can put M into the utility function, along with other things, as a real convenience in a world of stochastic uncertainty and indivisible transaction charges.[5]

If, however, one does put M directly into U, one must remember the crucial fact that M differs from every other good (such as tea) in that it is not really wanted for its own sake but only for the ultimate exchanges it will make possible. So along with M, we must always put all Ps into U, so that U is homogeneous of degree zero in the set of monetary variables $(M, P_1, \ldots, P_m)$, with the result that $(\lambda M, \lambda P_1, \ldots, \lambda P_m)$ leads to the same U for all λ.

In *Foundations*, I wrote such a U function:

$$U(q_1, q_2, \ldots, q_n; M, P_1, P_2, \ldots, P_n)_\lambda \equiv U(q_1, \ldots, q_m; \lambda M, \lambda P_1, \ldots, \lambda P_n),$$

where Ps are prices in terms of money. Here I want merely to add a little further cheap generality. The convenience of a given M depends not only on Ps, but also upon the earning assets you hold and on your wage prospects. It is not that we will add to M the earning-asset total EA, which equals $P_T\bar{T} + P_K\bar{K}$. Nor shall we add EA after giving the latter some fractional weight to take account of brokerage and other costs of liquidating assets into cash in an uncertain world. Rather, we include such new variables in U to the right of the semicolon to get:

$$U(q_1, \ldots, q_n; M, EA, W\bar{L}, P_1, \ldots, P_n) = U(q; x) = U(q; \lambda x).$$

That is, increasing all Ps, including those of each acre of land and machine and of hourly work along with M, will not make one better off. Thus U ends up homogeneous of degree zero in M and *all* prices $(M, P_K, P_T, W, P_1, \ldots P_n)$ by postulate.

Now, subject to the long-run budget equation indicated below, the representative man maximizes his utility:

$$\max_{\{q_1, \ldots, q_n, M\}} \quad U(q_1, \ldots, q_n; M, P_K\bar{K} + P_T\bar{T}, W\bar{L}, P_1, \ldots, P_n)$$

subject to

$$P_1 q_1 + \ldots + P_n q_n = W\bar{L} + r\,(\text{Total Wealth} - M)$$

or

$$P_1 q_1 + \ldots + P_n q_n + rM = W\bar{L} + r\,(\text{TW}) = W\bar{L} + r\,(P_K\bar{K} + P_T\bar{T} + M^*),$$

where each representative man has Total Wealth defined as:

$$\text{Total Wealth (in money value)} = EA + \text{Money Endowment} = P_K\bar{K} + P_T\bar{T} + M^*,$$

where M^* is the money created in the past by gold mining or by government.

[5]This is not the only way of introducing the real convenience of cash balances. An even better may would be to let U depend only on the time stream of qs, and then to show that holding an inventory of M does contribute to a more stable and greatly preferable stream of consumptions. The present oversimplified version suffices to give the correct general picture.

The maximizing optimality conditions give the demand for all q_1 and for M in terms of the variables prescribed for the individual, namely:

$$(P_1, \ldots, P_n, W, P_K, P_T; r, \bar{K}, \bar{L}, \bar{T}).$$

The optimality equations can be cast in the form:

$$\frac{\partial U/\partial q_1}{P_1} = \cdots = \frac{\partial U/\partial q_n}{P_n} = \frac{\partial U/\partial M}{r}$$

or

$$(A_{III,1}) \qquad \frac{\partial U/\partial M}{\sum_1^n q_j \frac{\partial U}{\partial q_i} + M \frac{\partial U}{\partial M}} = \frac{r}{W\bar{L} + r(P_K\bar{K} + P_T\bar{T} + M^*)}$$

$$(A_{III,2}) \qquad \frac{\partial U/\partial q_i}{\sum_1^n q_j \frac{\partial U}{\partial q_j} + M \frac{\partial U}{\partial M}} = \frac{P_i}{W\bar{L} + r(P_K\bar{K} + P_T\bar{T} + M^*)}$$

$$(i = 1, 2, \ldots, n).$$

But for society as a whole (and hence for the representative man who, even if he does not know it, represents $1/N$th of the total in our symmetrical situation) total money demanded, M, must end up equalling total money endowment, M^*:

$$(A_{III,3}) \qquad M = M^*.$$

An important comment is in order.[6] Although $A_{III,3}$ holds for society as a whole, being essentially a definition of demand-for-money equilibrium, each representative man (one of thousands of such men) can*not* act in the belief that his budget equation has the form:

$$P_1q_1 + \ldots + P_nq_n + rM = W\bar{L} + r(P_K\bar{K}_tP_T\bar{T} + M),$$

even though substituting $A_{III,3}$ into the earlier budget equation would yield this result. What is true for all is not true for each. Each man thinks of his cash balance as costing him forgone interest and as buying himself convenience. But for the community as a whole, the total M^* is there and is quite costless to use. Forgetting gold mining and the historical expenditure of resources for the creating of M^*, the existing M^* is, so to speak, a free good from society's viewpoint. Moreover, its *effective* amount can, from the community's viewpoint, be indefinitely augmented by the simple device of having a lower absolute level of *all* money prices. To see this in still another way, with fixed labour L and land T and capital K big enough to give the interest rates equal to the psychological rate ρ, the community can consume on the production possibility equation:

$$P_1q_1 + \ldots P_nq_n = F(\bar{K}, \bar{L}, T) = W\bar{L} + r(P_KK + P_T\bar{T})$$

and to *each* side of this could be added rM of any size without affecting this true physical menu.

[6]The next few paragraphs can be skipped without harm.

Evidently we have here an instance of a lack of optimality of laissez-faire: there is a kind of fictitious internal diseconomy from holding more cash balances, as things look to the individual. Yet if all were made to hold larger cash balances, which they turned over more slowly, the resulting lowering of absolute price would end up making everybody better off. Better off in what sense? In the sense of having a higher U, which comes from having to make fewer trips to the bank, fewer trips to the brokers, smaller printing and other costs of transactions whose only purpose is to provide cash when you have been holding too little cash.

From society's viewpoint, the optimum occurs when people are satiated with cash and have:

$$\partial U/\partial M = 0 \text{ instead of } r \times (\text{positive constant}) > 0.$$

But this will not come about under laissez-faire, with stable prices.[7]

Now let us return from this digression on social cost to our equations of equilibrium. Set A consists of the A_I equations relating to production and implied pricing relations, and of the A_{II} equations relating to long-run equilibrium of zero saving and investment, where technological and subjective interest rates are equal and provide capitalized values for land and other assets. Finally, A_{III} are the demand conditions for the consumer, but generalized beyond the barter world to include explicitly the qualitative convenience of money *and to take into account the peculiar homogeneity properties of money resulting from the fact that its usefulness is in proportion to the scale of prices.* Though the exact form of A_{III} is novel, its logic is that implied by intuitive classical theories of money.

All of equations A have been cast in the form of involving ratios of prices, values, and M^* only (to put A_{III} in this form, multiply M into the numerators on each side). That means they are homogeneous functions of degree zero in all Ps, and M^* or M, being capable of being written in the general form:

$$\text{A} \qquad G_i\left(q_1, \ldots, q_n, K, \bar{L}, \bar{T}, r; \frac{P_K}{M}, \frac{W}{M}, \frac{R}{M}, \frac{P_1}{M}, \ldots, \frac{P_n}{M}, \frac{TW}{M}\right) = 0$$

where all the magnitudes to the left of the semicolon are "real" and all those to the right are *ratios* of a price or a value to the quantity of money. If a price ratio like P_i/P_j appears in an equation and no M, we can rewrite the ratio as $(P_i/M)/(P_j/M)$.

To the set A, we now append a decomposable single equation to fix the supply of money:

$$\text{B} \qquad M \text{ or } M^* = \bar{M}, \text{ an exogenous supply.}$$

This single equation is not homogeneous of degree zero in Ps and M and therefore it does pin down the absolute scale of all Ps and values in direct proportion to the quantity of M. Why? Because Set A consists of as many indepen-

[7]See P. Samuelson, "D. H. Robertson, "*Quarterly Journal of Economics*, LXXVII, 4 (Nov. 1963), 517–36, esp. 535 where reference is made to earlier discussions by E. Phelps, H. G. Johnson, and R. A. Mundell. This article is reproduced in Joseph E. Stiglitz, ed., *The Collected Scientific Papers of Paul A. Samuelson* (Cambridge, Mass., 1966).

dent equations as there are unknown real quantities and ratios. Let us check this. Omitting fixed $(\bar{L},\bar{T})$, we count $n + 2 + n + 5$ unknowns in G_i when we ignore both $\dot{K}$ and the $\dot{K} = 0$ equation. We count $n + 3$ equations in A_I, 2 equations in A_{II}, and $n + 2$ *equations in* A_{III}. Thus $2n + 7 = 2n + 7$. Another way of looking at the matter is this: A_I and A_{II} determine all Ps as proportional to P_K. Then for fixed P_K and M^*, A_{III} determines all qs and M, the latter doubling when P_K and M^* double.

Summarizing, Set A determines all real quantities and all prices and values in ratio to the stock of M^*. Then equation B determines $M^* = M$ and hence the absolute level of all prices in proportion to $\bar{M}$.

Where in A or B is the quantity theory's "equation of exchange" to be found? Certainly not in B. If anywhere, an $MV = PQ$ equation must be found in A. Where? Certainly not in A_I or A_{II}. In A_{III}, equation $A_{III,1}$ deals with the relative marginal utility of the cash balance. By itself, it is not an $M = PQ/V$ equation. Only after all the A_{III} equations are solved, can we express M in a function that is proportional to any (and all) P_i:

$$M = P_i \psi_i (\ldots)$$

where the ψ functions depend on a great variety of real magnitudes.

This suggests to me that the late Arthur Marget was wrong in considering it a fault of Walras that, after the second edition of his *Elements,* he dropped a simple $MV = PQ$ equation. Classical and neoclassical monetary theory is much better than a crude quantity theory, although it can report similar results from special ideal experiments. In particular, correct neoclassical theory does not lead to the narrow anti-Keynesian view of those Chicago economists who allege that velocity of circulation is not a function of interest rates.

How M gets allocated

Symmetry plays an important role in the model given here. With every man exactly alike, it does not matter where or how we introduce new money into the system; for it gets divided among people in exactly the same proportions as previous M. We classical writers were aware that the strict (A,B) dichotomy held only when every unit's M (say M^1, M^2, ...) stayed proportional to total $\bar{M} = \Sigma M^k$. But being careless fellows, we often forgot to warn that this was only a first approximation to more complicated incidents of gold inflations and business cycle expansions.

Can this rock-bottom simplicity be retained if we relax this extreme symmetry assumption (which renders the problem almost a Robinson Crusoe one)? Providing all income elasticities, including that for M, are (near) unity, it never matters (much) how things are divided among people. Collective indifference curves of the Robinson Crusoe type then work for all society. The simple structure of A_{III} is preserved and the uniqueness of equilibrium is assured. Again, it matters not how the new M is introduced into the system.

Finally, there was an even more interesting third assumption implicit and explicit in the classical mind. It was a belief in unique long-run equilibrium independent of initial conditions. I shall call it the "ergodic hypothesis" by analogy to the use of this term in statistical mechanics. Remember that the classical economists were fatalists (a synonym for "believers in equilibrium"!). Harriet Martineau, who made fairy tales out of economics (unlike modern economists who make economics out of fairy tales), believed that if the state redivided income each morning, by night the rich would again be sleeping in their comfortable beds and the poor under the bridges. (I think she thought this a cogent argument against egalitarian taxes.)

Now, Paul Samuelson, aged 20 a hundred years later, was not Harriet Martineau or even David Ricardo; but as an equilibrium theorist he naturally tended to think of models in which things settle down to a unique position independently of initial conditions. Technically speaking, we theorists hoped not to introduce *hysteresis* phenomena into our model, as the Bible does when it says "We pass this way only once" and, in so saying, takes the subject out of the realm of science into the realm of genuine history. Specifically, we did not build into the Walrasian system the Christian names of particular individuals, because we thought that the general distribution of income between social classes, not being critically sensitive to initial conditions, would emerge in a determinate way from our equilibrium analysis.

Like Martineau, we envisaged an oversimplified model with the following ergodic property: no matter how we start the distribution of money among persons—M^1, M^2, . . .—after a sufficiently long time it will become distributed among them in a unique ergodic state (rich men presumably having more and poor men less). I shall not spell out here a realistic dynamic model but content myself with a simple example.

Half the people are men, half women. Each has a probability propensity to spend three-quarters of its today's money on its own products and one-quarter on the other sex's. We thus have a Markov transitional probability matrix of the form

$$A = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}+\frac{a}{2} & \frac{1}{2}-\frac{a}{2} \\ \frac{1}{2}-\frac{a}{2} & \frac{1}{2}+\frac{a}{2} \end{bmatrix}$$

with $a = \frac{1}{2}$ and

$$A^t = \begin{bmatrix} \frac{1}{2}+\frac{a^t}{2} & \frac{1}{2}-\frac{a^t}{2} \\ \frac{1}{2}-\frac{a^t}{2} & \frac{1}{2}+\frac{a^t}{2} \end{bmatrix}$$

$$\lim_{t\to\infty} A^t = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \text{ the ergodic state.}$$

Suppose we start out with men and women each having M of ($100, $100). Now introduce a new $100 to women only. Our transitional sequent in dollars will then be ($200, $100), ($175, $125), ($162½, $137½), ($156½, $143¾),

($151 9/16, $148 7/16), . . . with the obvious limiting ergodic state $150, $150) since the divergence from this state is being halved at each step. Such an ergodic system will have the special homogeneity properties needed for the (A,B) dichotomy.[8]

None of this denies the fact that the leading neoclassical economists often recognized cases and models in which it does make a difference, both in the short and the long run, how the new money is introduced and distributed throughout the system. One of the weaknesses of a crude quantity theory is that it treats M created by open-market purchases by the central bank as if this were the same as M left over from last century's (or last minute's) mining. A change in M, accompanied by an opposite change in a near-M substitute like government short-term bonds, is *not* shown in my Set A.

Indeed, when all men are alike and live for ever, we have too simple a model to take account of the interesting effect upon the system of permanent interest-bearing public debt which we as taxpayers know we will not have to pay off or service beyond our lifetimes.[9]

Epilogue

With the positive content of traditional monetary theory now written down concretely for us to see, kick, and kick at, a few comments on some controversies of the last twenty years may be in order.

Oskar Lange began one line of reasoning on price flexibility in 1939 which culminated in his 1944 Cowles book, *Price Flexibility and Employment.*[10] Hicks' *Value and Capital,*[11] with its attempt to treat bonds and money just as some extra $n + 1$ and $n + 2$ goods along with n goods like tea and salt, had, I fear, a bad influence on Lange. It led to his suppressing possible differences between stocks and flows, to attempts to identify or contrast Say's Law with various formalisms of Walrasian analysis (such as the budget equation), and to discussion in the abstract of functions of many variables possessing or not

[8]Let me warn that this discussion in terms of a Markov probability matrix is meant to be only indicative. The temporal sequence of decisions to exchange money for goods and services and goods for money, with all that is implied for the distribution among units of the stock of M at any time, is more complicated than this. In our most idealized models, we assumed that, whatever the complexity of the process, after enough time had elapsed the M would get distributed in a unique ergodic way. This does not beg the question, since there are models in which this is a theorem. In our more realistic moods, we tacitly used models involving *hysteresis*: Spain would never be the same after Columbus; Scarlett O'Hara would be permanently affected by the Confederate inflation, just as Hugo Stinnis was by the 1920–23 German inflation. Obviously, in such models all real variables do not end up unchanged as a result of certain unbalanced introductions of new M into the system. In that sense realistic equations do not seem to have the homogeneity properties in (M, P, . . .) of my Set A; but if we were to write in A the variables (M^1, M^2, . . .) and not merely their sum ΣM^k, it is still possible that homogeneity properties would hold—so that doubling *all* M^k together would be consistent with doubling all Ps. But this is too delicate a question to attempt in brief compass here.

[9]My *Economics* (6th ed., New York, 1964), 342, shows that (M, public debt) and (λM, λ public debt) play the role in more complicated systems that (M) and (λM) play in the simple classical system given here. Crude quantity theorists should take note of this distinction, which Franco Modigliani has also insisted on.

[10](New York, 1944).

[11](Oxford, 1939).

possessing certain abstract homogeneity properties. There are many interesting points raised in Lange's book, and several analytical contributions to non-monetary economic theory. But only about a dozen pages grapple with the key problem of money (e.g., pp. 5–19), and these stay at a formalistic level that never deals with the peculiar properties and problems of cash balances. I do not say that this approach of Lange's cannot be used to arrive at valid results, but in fact it remained rather sterile for twenty years.

I had thought that Don Patinkin's work from 1947 on, culminating in his classic *Money, Interest, and Prices* was much influenced by the Lange approach, and I thought this a pity. But, on rereading the book, I am not sure. What Patinkin and Lange have in common is a considerable dependence upon the *Value and Capital* device of lumping money in as an extra good. This approach has not kept Patinkin from arriving at a synthesis consistent with what I believe was the best of neoclassical theory, or from going beyond anything previously appearing in the literature. But it may help to account for his attributing error to earlier thinkers when a more sympathetic reading might absolve them from error. When we become accustomed to approaching a problem in a certain way and using a certain nomenclature, we must not confuse the failure to use this same language and approach with substantive error. Still, beyond that, Patinkin scores many legitimate points: monetary economists had better intuitions than they were able to articulate. Thus I suspect that my (A,B) dichotomy is really very similar to what Cassel had in mind, but the only form in which he could render it mathematically was (A′, B′), which is inadequate (as Patinkin insists, though perhaps not for all the reasons he insists on). In what sense can one say that a man believes one thing when he says something else? In this non-operational sense: if one could subpoena Cassel, show him the two systems and the defects in one, and then ask him which fits in best with his over-all intuitions. I believe he would pick (A,B) and not his own (A′, B′).[12] I might add that Cassel is not Walras; and it seems to me that Walras comes off better on Patinkin's own account than he is given credit for.

Some will interpret Archibald and Lipsey as defending an (A′, B′) dichotomy against Patinkin's rejection of that dichotomy. If that is their primary intention —and I am not sure that it is—I fear I must side with Patinkin. Logically, one can set up (A′, B′), as I did here and as Cassel did. But I think it is bad economics to believe in such a model. *All* its good features are in the (A,B) dichotomy and none of its bad ones.

On the other hand, there is certainly much more in Archibald and Lipsey than a defence of (A′, B′) and this important part of their paper seems to me to be quite within the spirit of Patinkin's analysis and my own. Here, however, I shall comment on the two different dichotomies.

I begin with (A′, B′).

[12]Needless to say, the test is not whether Aristotle, apprised of Newton's improvements over Aristotle, would afterwards acquiesce in them; the test is whether in Aristotle's writings there are non-integrated Newtonian elements. If so, we credit him only with non-integrated intuitions.

A′ $F_i(q, P)_\lambda \equiv F_i(q, \lambda P)$ $(i = 2, \ldots 2n)$

B′ $P_1 = 1$ or $\sum_{j=1}^{n} q_j^k P_j = \bar{V}M,\ M = \bar{M}.$

Suppose that we can solve n of the A′ equation to eliminate the qs, ending up with the independent homogeneous functions

A′ $f_i(P)_\lambda \equiv f_i(\lambda P) \equiv f_i(1, P_2/P_1, \ldots, P_2/P_1)$ $(i = 2, \ldots, n - 1)$

B′ $\sum q_j^*(P_j/M) = \bar{V},\ M = \bar{M}.$

Although f_i involve actually money Ps, it is not logically or empirically mandatory to interpret them as "excess-demand" functions which drive up (or down) the *money* Ps. Some students of Hicks, Lange, and Patinkin fall into this presupposition. Logically, there *could* be dynamic adjustments of price ratios—as e.g. P_i/P_1 or P_i/P_j, either of which could be written as $(P_i/M)/(P_j/M)$—of the type

a′ $[d\,(P_i/P_1)]/dt = k_i f_i(1, P_2/P_1, \ldots, P_n/P_1)\ (i = 2, \ldots, n)$

b′ $[d\,(P_1/M)]/dt = k_M[\bar{M} - \sum (P_j/P_1)(P_1/M)Q_j^*(1/V)]k_j, k_m > 0,$

where the ks are positive speed constants of adjustment and where the q^* and V may be functions of relative Ms. Such a system could dynamically determine *relative* prices within a decomposable real set A′ and then determine the absolute price level in Set B. Note that no version of Walras' Law relates B′ to A′ or b′ to a′. Walras' Law in the form that merely reflects the Budget Equation of each consumer is expressed in the functional dependence of the $f_1(1, P_2/P_1, \ldots)$ function (which we can ignore) on the rest—namely

$$f_1(1, P_2/P_1, \ldots) \equiv -\sum_2^n (P_j/P_1) f_j(1, P_2/P_1, \ldots).$$

If (a′,b′) is dynamically stable, $P_i/M \rightarrow$ constant is in agreement with the long-run quantity theory.[13]

[13]A short-run quantity theory need not hold. Doubling M this minute or this week need not double this week's prices. But there is a sense in which homogeneity holds in *every* run. Suppose as a *fait accompli* we are all made to wake up with every dollar of M *exactly* doubled and every P (present *and* future) exactly doubled. If nought else has changed, we recognize this to be indeed a new equilibrium. And if the time-profile of equilibrium is unique, how can we have any other time-profile of prices? At the root of this paradox is the assumption of perfectly balanced changes in M, perfect foresight, and the postulate of uniqueness of equilibrium. All this is a far cry from interpreting the stream of contemporary history.