"Reply to Ando and Modigliani and to Deprano and Mayer"

by Milton Friedman with David Meiselman

American Economic Review 55, September 1965, pp. 753-785

Because our article ventured into almost virgin territory, it was necessarily tentative, probing, and imperfect.[1] Because it questioned the new orthodoxy, we expected it to provoke controversy. That it has done so impresses us less than the large area of agreement between our critics and ourselves. Professors Ando and Modigliani (referred to hereafter as AM), and DePrano and Mayer (referred to hereafter as DM), all agree with us that changes in the quantity of money are connected with changes in nominal income over short-run periods and that this connection is at least to some extent independent of concurrent changes in autonomous expenditures. So far as we can judge, they also all agree with at least the broad outlines of our analysis of the channels through which the changes in the quantity of money exert their influence, though no doubt there are differences about details.

One other important element of agreement is implicit, and extends also to Professor Hester, who has published a criticism of our work in another journal.[2] Neither we nor our critics had much problem selecting a useful empirical counterpart for the theoretical concept of "money." There is some leeway, but the alternatives are fairly clear-cut and differences in results are relatively minor. No elaborate process of trial and error, of combination and recombination of components, is required to get measures that are closely correlated with nominal income and the nominal value of consumption. We and our critics all used the same measures without much ado.[3] The contrast with "autonomous expenditures" could hardly be sharper. Among us, we have produced more measures than there are critics. We settled on one; AM on a different one, which is the sum of two separable components; DM, after running "basic tests . . . on 20 different, but not unreasonable, definitions of autonomous expenditures" settle on two, but also carry two others along for the ride; Hester came up with four measures that only partly overlap the others. And all of us harbor serious doubts about the measures we settled on. However useful "autonomous expenditures" may be as a theoretical construct, it is still far from having any generally accepted empirical counterpart.

This contrast between the stock of money and autonomous expenditures is relevant in judging the import of the gauntlet of tables run by the reader who has arrived at this point. The single set of relations between money and the dependent variables comes from the first test runs. The multiple sets of relations between autonomous expenditures and the dependent variables are the end-product of a whole series of test runs. We could, of course, have made further test runs for alternative measures of money. No doubt we could have found measures that were more highly correlated with income and consumption than the measure we used, especially if we had restricted ourselves to recent years. We have refrained from doing so in order not to confuse the issue, which is not how high a correlation can be wrung from a computer but rather how well two alternative theories correspond to experience when the central concepts in the theories are defined either by earlier work or by reasonably objective criteria applied in advance.

Where our critics disagree with us is that they believe we have understated the independent role of autonomous expenditures operating through a multiplier process. They claim first, that the particular measure of autonomous expenditures we adopted was poor; second, that its use produced results that were biased against the income-expenditure approach; and third, that our procedure for comparing the relative stability of monetary velocity and the investment multiplier is too simple—too "one-equation" an approach.

We readily plead guilty to two of the three charges. Indeed, we emphasized in our original paper that our measure of autonomous expenditures is unsatisfactory and that our approach is too simple. However, in our opinion, our critics have neither established that our measure biased the results nor demonstrated that their alternative measures are more defensible in terms of relevant criteria. They have not shown what was wrong with the criterion we used to select a measure of autonomous expenditures and they have not developed alternative criteria capable of being applied to the data. The attempt in AM's Appendix to compare more complex models is commendable but unfortunately unsuccessful for reasons spelled out in our comments on their Appendix.

In the course of trying to satisfy ourselves whether our critics' objections are valid, we formulated more explicitly than in our article the income-expenditure model used in our analysis. We also clarified the theoretical requirements for a valid comparison of the two models in the simple way we compared them. These results, presented partly below in Sections III and IV.A. and partly in our reply to Hester, may contribute more to progress on the main problem than the specific answers to specific criticisms and the counter criticisms that fill the rest of our reply—but then the chaff is often more bulky than the grain.

Our initial draft of this reply was written in response to an earlier version of the AM paper. In our revision, we have tried to incorporate comments on the DM paper as well, but no doubt the final version still bears marks of its origin. Because the DM paper partly duplicates the AM paper, we have generally replied only to the additional subject matter introduced by DM. We have used throughout AM's notation, though we have had to add a few symbols. We found AM's excellent Table 1 invaluable in keeping track of the many balls being juggled.

We apologize for quoting extensively from our paper. We do so to let the reader judge for himself whether we are right in our belief that our critics have given a misleading impression of the aim and content of our paper. We hope also that these tidbits may entice readers who have not yet done so to read our complete paper.[4]

At first sight, the most persuasive argument in both critiques is the alleged statistical demonstration that altering the definition of autonomous expenditures "changes drastically" (AM's words) our conclusions. We shall therefore deal first with this allegation, waiving, for the time being, all questions about the theoretical validity of the procedures used by our critics. We shall then turn to some debating points raised by our critics that simply confuse the issue, and only then, in Sections III, IV, and V, to the substantive matters that are of Scientific importance.

I. Does Changing the Definition of Autonomous "Drastically" Change Our Conclusions?

From data for the period 1897–1958 (annual for the whole period and also quarterly for 1946–58) we concluded that "the income velocity of circulation of money is consistently and decidedly stabler than the investment multiplier except only during the early years of the Great Depression after 1929" (p. 186).[5] AM and DM maintain that this conclusion results from our use of an inappropriate concept of autonomous expenditures (A). AM assert that substitution of their alternative concept (Z^a+X^a, which we shall designate A^*) reverses our conclusion. As evidence, they present correlations for the years 1929–58, excluding 1942–46. For those 25 years, the correlation between consumption (C) and A^* is decidedly higher than between C and money (M), which in turn is decidedly higher than between C and A.

DM prefers two other concepts of autonomous expenditures, one on a net basis and one on a gross basis (which we shall designate A^{**} and A^{***} respectively). Their evidence consists largely of correlations between C and A^{**} or A^{***} for several periods since 1929. The correlation between C and A^{**} is generally somewhat higher than between C and A, and between C and A^{***}, a bit higher still. In most comparisons, the effect is to leave M more highly correlated with C than autonomous expenditures are, but to diminish the margin of superiority. DM also correlate components of A^{**} or A^{***} with C and find that two in particular, namely, gross fixed private domestic investment and this plus exports, are consistently highly correlated with C, generally more highly than M, A^{**}, or A^{***}. They are somewhat at a loss how to interpret this result.

A minor reason why AM's contention is not justified is because of the difference in periods covered. As we noted, "This [the early years of the Great Depression] is the only period when the relations between autonomous expenditures and consumption are impressively close either relative to the money relations or absolutely. This is, of course, also the period that gave rise to the income-expenditure theory in its present form and that sparked the conversion of a large fraction of economists to it" (p. 188). We also pointed out that even this exception was eliminated if the calculations were made for 1930–39 instead of for 1929–39. The exceptional period, 1929–39, which AM extended to 1941, contains more than half the years for which AM make calculations. Had we restricted our calculations to the years since 1929, we would have reached much less clear-cut conclusions than we did. The return of the results after World War II to the pattern prevailing from 1897–1929 enabled us to term 1929–39 the exception rather than the rule. Since data are readily available for the earlier years, we are puzzled why AM did not make their calculations for the time span we analyzed.

TABLE 1—CORRELATION BETWEEN CONSUMPTION AND ALTERNATIVE INDEPENDENT VARIABLES: TWO CONCEPTS OF AUTONOMOUS EXPENDITURES AND THE QUANTITY OF MONEY, FOUR PERIODS, ANNUAL DATA, 1929–58

Period	Correlation Coefficient Between C and			Squared Standard Errors of Estimate (S_E^2) for Regression of *C* on			Standard Error of Estimate as a Percentage of Mean Value of *C* for Regression of *C* on		
	A	*A**	*M*	*A*	*A**	*M*	*A*	*A**	*M*
1929–41	.936	.906	.910	16	23	23	6.3	7.5	7.4
1930–41	.939	.939	.965	14	14	8	6.0	5.9	4.5
1947–58, exc.	.808	.976	.989	704	98	47	11.7	4.3	3.0
1942–46	.964	.996	.989	601	69	174	16.9	5.8	9.2

The major reason why AM's contention is not justified is very different. They have confused secular shifts with the year-to-year changes that are relevant to the central issue. That issue, as we pointed out, "has to do with short-run relations.... Here the chief assertion [of the Keynesian view] is that . . . short-run changes in income can be regarded as largely caused by and reflecting corresponding short-run changes in investment expenditures" (p. 167). Hence, we stated, "it seems desirable to make the comparisons [between the income-expenditure and quantity theories] for relatively short periods" (p. 174).

AM apparently overlooked this caveat. All the calculations in their text are for the 25 years 1929 through 1941 and 1947 through 1958 taken as a whole. Since their calculations are all for aggregates in nominal values (we made calculations also in real terms) and since prices roughly doubled from the first set of years to the second and population rose by about 25 per cent, it is clear that their calculations are for two distinct sets of observations clustered about very different means. The results of any calculations that do not allow for trend are therefore dominated by the differences between the means for the two periods and have little to say about the short-run relations that are at issue.[6] When the relevant ones of their calculations are made for each period separately (1929 through 1941 and 1947 through 1958), the results turn out to be consistent with our earlier results and to require little change in our conclusions.[7]

These statements are documented in Table 1, which gives simple correlation coefficients for three regressions and four periods.[8] The reason for including 1930–41 is to check whether the conclusion we reached in our paper that the exclusion of the 1929 observation eliminated the exception also holds for the AM definition of autonomous.

If we had used *A** instead of *A* and had still correlated it with *C* (which we would not have done, see Section III.C, below), and if we had used 1929–41 instead of 1929–39, our conclusion about the greater stability of velocity than the multiplier would have been slightly strengthened for the prewar period, and greatly weakened for the postwar period. For the prewar period, the correlation between *C* and *M* is higher than between *C* and *A** even for 1929–41.[9] For the postwar period, the correlation between *C* and *M* exceeds that between *C* and *A** by a much smaller margin than it exceeds that between *C* and *A*. Since the difference between *A** and *A* would very likely be even less for the pre–1929 period than for 1929–41, presumably the results for still earlier years would be much the same for *A** as for *A*.[10] Hence, the substitution of *A** for *A* would have required only a minor change in our final conclusions, which were based entirely on the results for sub-periods and not on those for lengthy periods. The correlations for longer periods were included simply as additional information.

AM make much of the superiority of the squared standard error of estimate over the correlation coefficient as a measure of goodness of fit. This is purely a debating point that has no relevance to any of our results. *When the dependent variable is the same*, the squared standard errors for different sets of independent variables are directly proportional to squared correlation coefficients. In our paper, we made no comparison for which the italicized qualification is not satisfied. Hence, substitution of squared standard errors for the correlation coefficients we presented would have changed nothing of substance. When the dependent variables are not the same, both correlation coefficients and standard errors are likely not to be comparable. That is precisely why we insisted that every comparison between the alternative theories should involve the identical dependent variables, both in concept and years covered.

To keep our results as close to AM's as possible, however, we have supplemented the correlation coefficients in Table 1 with squared standard errors of estimate (their S_E^2). To bring out some of the issues involved in interpreting this measure, we have included also the percentage error (100 S_E/C), which perhaps comes closer than the absolute error to being comparable for different periods of time. It will be noted that only for *M* is the squared standard error larger for the whole period than for any of the subperiods. The reason, as is obvious from a glance at our Chart II-8 (p. 198), is that there was a sizable shift of the regression of *C* on *M* from before to after the war. The slope was not much altered but the position was.

TABLE 2—PARTIAL CORRELATION BETWEEN CONSUMPTION AND ALTERNATIVE INDEPENDENT VARIABLES: TWO CONCEPTS OF AUTONOMOUS EXPENDITURES AND THE QUANTITY OF MONEY, LAGGED CONSUMPTION HELD CONSTANT, FOUR PERIODS, ANNUAL DATA, 1929–58

Period	Square of Multiple Correlation Coefficient, *C* regressed on C_{t-1} and			Partial Correlation Coefficient (C_{t-1} Held Constant) Between *C* and			Standard Error of Estimate as Percentage of Mean Value of *C* for Regression of *C* on			
							C_{t-1}	C_{t-1} and		
	A	*A**	*M*	*A*	*A**	*M*		*A*	*A**	*M*
1929–41	.974	.940	.886	.96	.92	.84	11.1	3.0	4.6	6.3
1930–41	.978	.968	.947	.97	.96	.94	11.4	2.7	3.3	4.2
1947–58	.987	.990	.992	.02	.46	.57	2.2	2.3	2.1	1.9
1929-58 excl. 1942–46	.997	.997	.997	.61	.69	.66	4.6	3.7	3.4	3.5

In our paper, we supplemented correlations like those in Table I with correlations between first differences, in order to assure that our results were not produced simply by common trends.[11] For the same purpose, AM use an alternative technique: the inclusion of lagged consumption as an independent variable. This allowance for trend eliminates most of the distorting effect of combining two very different periods, but even so, the results for the period as a whole are less illuminating than those for the separate periods, as Table 2 shows. The only appreciable difference from Table 1 for the subperiods is that the exclusion of 1930 does not suffice to reverse the order of the correlations: that between *C* and either autonomous concept remains higher, though by a much reduced margin, than between *C* and *M*. For the postwar period, the trend of consumption is so regular that the variation over and above that accounted for by prior-year's consumption is very small (2.2 per cent). Allowing for *M* as well reduces this percentage error to 1.9 per cent, or by over twice as much as allowing for *A** reduces the percentage error. Allowing for *A* renders matters worse.[12] Given this situation, the quarterly data available for this period are especially valuable. The analysis of them in our paper—on both a synchronous and lagged basis—confirmed the results we obtained from annual data.

When account is taken of trend, the results for *A** are again much the same as those we obtained for *A* with a different method of allowing for trend. Again, the substitution of *A** for *A* worsens the income-expenditure correlations for the pre-World-War II years but still leaves them higher than the quantity-theory correlations, and improves the income-expenditure correlations for the post-World-War II years, but still leaves them lower than the quantity-theory correlations.[13]

DM do consider prewar and postwar periods separately, which is why they get no such striking reversals of our results as AM assert they get.

II. Some Debating Points[14]

A. Treatment of War Years

AM take us strongly to task because, they say, we "do not even once mention the possibility that any of [our] results might be distorted by the inclusion of the war years!" This is a debating point pure and simple, which AM are led to raise only by their own statistical error of treating the period from 1929–58 as one whole. It has no relevance to our paper, as the following quotations demonstrate.

> In choosing the periods of time, we have taken several considerations into account. First, since the question at issue is mainly the short-term stability of the relations being compared, it seems desirable to make the comparisons for relatively short periods. Second, since the relations may differ at different phases of the cycle, it seems desirable that any one comparison should cover one or more complete cycles.... Third, since most of the available data are annual, single business cycles generally provide too small a number of observations to yield statistically meaningful results. The compromise we have adopted among these somewhat conflicting considerations is to divide the period for which data are available into two sets of overlapping segments, one set marked out by the troughs of the major depressions during the period (1896, 1907, 1921, 1933, 1938) except for the postWorld-War II period, which we have marked off simply by the end of the war; a second set, by peaks intermediate between the troughs of major depressions, except again for dates separating out World War II.... The dates we have used are 1903, 1913 (to get a period excluding World War I), 1920, 1929, 1939, (to get a period excluding World War II), 1948, and 1957. We have made computations for the period as a whole as well as the separate segments (pp. 174–75).

The results turned out to be so consistent, except for the 1929–39 decade, that we had no occasion to discuss the results for other subperiods in detail: the peacetime subperiods alone gave the same results as did the subperiods including some war years; and our conclusions were based on the results for the subperiods, not for the period as a whole.

B. One-Equation-One-Independent-Variable Models in Search of the Highest Correlation

AM several times characterize our procedure in these or very similar words, describe us as "arbitrarily" imposing these conditions and as picking the variable *A* "arbitrarily," and say "there is no justification for FM's juxtaposition of the income-expenditure framework and the quantity theory model as mutually exclusive hypotheses."

A few quotations from our article will help to put these remarks in their proper light.

1. *Simple versus sophisticated models*. We wrote:

The central issue in dispute is not theoretical but empirical. It is easy enough to construct an analytical system that embodies both the relations between investment and consumption and the relations between money and income, that is, both the multiplier relations and the velocity relations. Economists who regard monetary changes as primary are divided from those who regard them as secondary much less by different theoretical systems than by different empirical judgments, different judgments about which set of relations in the more generalized theoretical system is (a) critical in the sense of being in practice the primary source of change and disturbance and (b) stable in the sense of expressing empirically consistent relations which can be depended on to remain the same from time to time....

In seeking to examine the relative stability of velocity and the multiplier, we faced an initial choice between two major approaches. The issue can be investigated by a sophisticated analysis involving many variables. Such an analysis must inevitably be restricted to a narrow segment of space and time.... Alternatively, the issue can be investigated initially on a rather simple level for a wide range of space and time. Our choice has been the second. It is our view that the issue that divides economists is extremely basic and one that should lend itself to a common answer over a wide range of circumstances. If it does not, it means that the dichotomy posed is much too simple, that the key issue is not which view to accept but rather the circumstances under which the one or the other view is likely to be the more fruitful. Moreover, in an investigation of this kind, it seems better to rely initially on a wide range of evidence interpreted on a rather simple level than on the more indirect and longer chain of connections inevitable in a sophisticated analysis resting on a narrower base (pp. 168–70.)

The fact that we have done so [neglected most refinements] makes it necessary to emphasize that our results cannot be decisive. On the simple level on which we propose to test the two theories, equation (1) [the quantity theory equation] might turn out to be better than equation (2) [the income-expenditure theory equation], or conversely; whereas on a more sophisticated level, when additional variables are introduced, the relative advantage of the two might be reversed. This possibility cannot be ruled out, although it seems a reasonable presumption that the relationship which explains the most in its simplest version is the relationship that will be most fruitful to explore further to convert into a more sophisticated form (p. 174).

2. *Picking A "arbitrarily."* We wrote:

One by-product of this investigation was the discovery that there is neither clear-cut agreement on the specific statistical definition of autonomous and induced expenditures nor any well established criteria for choosing particular definitions for a particular problem or period or body of data (p. 180).

In our actual empirical work, much the greatest amount of time was spent in trying to draw the appropriate boundary lines [between autonomous and induced

> expenditures] rather than in the calculations and analysis designed to compare and test the two hypotheses. We are by no means satisfied that we have used the appropriate criteria in drawing the lines. Neither are we satisfied with the precise lines we have drawn, some of which we regard as highly tentative. Much further work remains to be done on this fundamental problem, in particular in determining statistical tests for making the best choice (p. 181).

Pages 181 to 183 of our paper are then devoted to developing the criterion we used (see the excepts from these pages in the next subsection), and a 16-page Appendix A, "Selection of Variables," summarizes the statistical tests we made. In the course of our tests, we considered explicitly and rejected definitions very close to the one AM uses.[15]

We quite agree with AM and DM that if A^* or A^{**} or A^{***} is a better measure of autonomous than A, then the correlations we computed are not the right ones for comparing the goodness of the two alternatives. However, they must equally agree that if A is a better measure, their correlations are not the right ones. And AM and DM give no evidence, theoretical or empirical, that we have used the wrong criterion in selecting A or that we interpreted the empirical evidence we present incorrectly. AM merely assert, without evidence, that A^* is better, justifying its selection only by their intuition about what is "beyond reasonable doubt," and, for government expenditures and receipts, by reference to a paper by Ando, Brown, and Adams. The same intuitions are the sole basis for their first charge that our estimates of the investment multiplier have a downward least-squares bias. If A is a better measure than A^*, then of course their estimates, not ours, are biased.

Similarly, although DM refer to "theoretical and empirical grounds discussed in the preceding sections" for rejecting our measures, so far as we can tell from a careful reading of the preceding sections, no theoretical grounds whatsoever are given for such a conclusion, only references to "expenditures which are usually considered autonomous," to what "might be," and to how the treatment of inventory investment as autonomous if in fact it were induced "leads to wrong results," but with no theoretical (or other) evidence to justify the conclusion that inventory investment should be excluded from the total called autonomous. The only empirical evidence cited consists either of assertions about what items are or are not endogenous, or simple correlations with consumption expenditures (referred to later as "basic tests".)[16] As is clear from the quotation in the next subsection, this is not a valid test, whether or not it be "basic."

We must confess that we are puzzled by the smoke screen thrown up by both pairs of critics on this issue. If the concept of "autonomous" expenditures is to be useful as something more than an incantation for the faithful and an empty box for our ignorance, there must be some objective and operational criteria for giving it empirical content, something more clearly reproducible than intuition and unstated "theoretical and empirical" grounds. It is hardly satisfactory to have some score of alternative definitions all in the running at the same time. Our criterion (see next subsection) may not be a good one or we may have applied it incorrectly, but if so, we should appreciate it if our critics would explain what is wrong with what we have done.[17]

Incidentally, even if A^* or A^{**} or A^{***} were a better concept than A, the correlations AM and DM compute are not the right ones for comparing the goodness of the two alternative theories (see Sections III.C and IV.C below).

3. "*Search for the highest correlation*." On this point, which is closely related to the preceding, both AM and DM have simply confused, as they have throughout their papers, two different issues: (i) comparison of goodness of fit of two alternative models; (ii) selection of variables to use in the models.

As already pointed out, the correlation coefficient is an appropriate criterion of the relative goodness of fit of two alternative models when the dependent variables they predict are the same. This is the one and only purpose for which we used the height of the correlation coefficient as a test. As noted, the results would have been identical if we had used residual variance. On the other hand, in discussing the problem of distinguishing between autonomous and induced expenditures, we said in our original paper (pp. 181–83):

> What criterion should be used to fix the boundary lines? One simple method is to correlate alternatively defined measures of the independent variable with the dependent variable [income] and then select the concept which yields the highest correlation.... Applying this criterion to the definition of autonomous reveals, however, that it is not satisfactory.... It is possible to get a correlation as close to unity as desired, simply by including all items of income that vary much over time in "autonomous," which is to say, by correlating these variable items with themselves. The procedure adopted above to evade the difficulties raised by the spurious correlation, namely, correlating the rest of income with autonomous expenditures, is no solution for the present problem, since each definition of autonomous would then be correlated with a different variable, . . . [and] the resulting correlations would not be comparable....
>
> [An] alternative approach to the definition of autonomous expenditures can be illustrated by considering the question whether durable consumer goods should be included in consumption or in autonomous expenditures. Let *D* stand for consumption expenditures on durable goods, *N* on non-durable goods, *C* for their total, and *A* for autonomous, according to some tentative definition that excludes durable consumer goods but settles other doubtful items. The question to be decided is whether *D*+*A* or *A* alone is a preferable definition for autonomous expenditures. If *D* and *A* were perfect substitutes as autonomous or income-generating expenditures, then a shift of $1 from *D* to *A* or from *A* to *D* would have no effect on *N*. Hence *N* would tend to have a lower correlation with either *D* or *A* alone than with their sum. Consequently, this approach implies that a necessary condition for the inclusion of *D* in autonomous is that
>
> (1) $$r_{N(D+A)} \left[\begin{array}{l} r_{ND} \\ \text{and} \\ r_{NA} \end{array} \right.$$
>
> The requirement that the sum of autonomous and induced expenditures equal income gives rise to a similar test in the other direction, ... Suppose (1) is not satisfied. If this occurred because *D* was a part of induced expenditure along with

N, one might expect shifts between *D* and *N* to be independent of changes in *A*. Changes in *A* would affect only their sum. But this would imply that

$$(2) \qquad r_{A(D+N)} \begin{bmatrix} r_{AD} \\ \text{and} \\ r_{AN} \end{bmatrix}$$

This approach therefore yields the following criterion:

Possibility	*Condition (1)*	*Condition (2)*	*Conclusion*
(a)	Satisfied	Not Satisfied	*D* autonomous
(b)	Not satisfied	Satisfied	*D* induced
(c)	Satisfied	Satisfied	Ambiguous
(d)	Not satisfied	Not satisfied	Ambiguous

There is nothing about the arithmetic of the relations among the correlation coefficients that requires either (a) or (b) to hold. It is entirely possible, and in our work has frequently happened, that either (c) or (d) should hold. In consequence, this criterion is not one that is necessarily decisive.

Unfortunately, however, we have been able to devise no criterion that seemed better to us. Consequently, we have employed the criterion just outlined. When the results have been ambiguous, we have followed the procedure that seemed more in accord with the general presumptions in the literature about income-expenditure relations.

In applying the criterion, we have in each case set up the problem as in the above example. That is, we have tentatively decided all questions of inclusion or exclusion except one, leaving us with a division of total income into three parts, the treatment of one of which is in doubt.

Clearly, we did not use the highest correlation to fix the boundary lines between autonomous and induced expenditure or to choose the appropriate definition of money. Our criterion uses correlation coefficients but in a very different and more sophisticated way. It is AM and DM, not we, who implicitly use the highest correlation as a means of selecting the definition of variables, when they cite the higher correlations they get as relevant evidence of the superiority of their definitions over ours.

C. Use of Some Data for Choosing Definitions and Testing Hypothesis

DM criticize us for using the same data for defining variables and testing hypotheses incorporating these variables. In this criticism they presumably refer only to our tests for 1929–

58, since these are the only years which we use in constructing our definitions, though we tested the hypotheses for the whole period 1897–1958.

For the years 1929–58, their criticism would be entirely valid if indeed we had used the highest correlation as a means of selecting the definition of variables. In that case, we would have been using up degrees of freedom, as it were, in constructing definitions that we later used over again in testing hypotheses. However, as pointed out in the preceding section, we did not use the highest correlation as a criterion. In effect, we used different information for the same years in deriving definitions and in testing hypotheses. However, the two classes of information probably overlap. Hence, the criticism is partly valid. Certainly we would prefer to use independent evidence, but where are we to get it? DM talk about a priori definitions. But from nothing you can get nothing. They are simply using "a priori" as a euphemism for casual empiricism organized by unstated criteria, so, in effect, they, and AM too, are using, in a less explicit way, the same data to define terms and test hypotheses. This is inevitable in scientific work. The cure is ultimately in the cumulative weight of evidence.

It is amusing that this criticism of DM precisely cancels one of AM's. If we had done what DM criticize us for, the effect would have been to bias upwards the correlations for both the quantity theory and income-expenditure hypotheses, but to bias the latter more than the former because we considered a much wider variety of definitions for autonomous expenditures than we did for the stock of money. Yet AM criticize us for a procedure that they regard as biasing the income-expenditure correlations downwards.

III. The General Theoretical Issue

The theoretical exercise of AM's Section I(iii) is unnecessary. They are led into it by two factors: (1) They attempt to derive our equation from a consumption function we did not use. (2) They and DM attribute to us an unexplained and unexplainable desire to correlate consumption with other things regardless of the definition of autonomous expenditures. We had no such desire. What we used as a dependent variable was critically dependent on our concept of autonomous expenditures; had we chosen a different concept of autonomous expenditures, we would have used a different dependent variable.

Since AM and DM as well as Hester have misunderstood us on both points, we obviously failed to make it clear that our theoretical structure for income-expenditure analysis is the standard structure used by all of them and by most other economists. Instead of following the detailed convolutions of AM's analysis, we shall try instead to rectify the defects of our prior exposition by setting forth explicitly the general theoretical structure we used, even at the risk of repeating what is well known.

A. *The Simple Model*

The central feature of simple income-expenditure models is that they distinguish between two categories of expenditures entering into the income stream: expenditures that are closely linked to (induced by) current income flows, and expenditures that are autonomous, not in the sense of being random or arbitrary or unexplainable but simply in the sense of not being determined by current income flows. This distinction between induced and autonomous expenditures—not that between consumption and investment—is fundamental. The tendency has arisen to use the term "consumption" as synonymous for "induced" because of the empirical judgment that consumption (or, better, change in consumption) is predominantly induced and, conversely, that the bulk of induced expenditures consists of consumption. Because our definitions led to this result, we adopted the common terminology. We now believe that it would have contributed to less misunderstanding if we had been pedantic and had used the term "induced" throughout instead of simply including it in a parenthesis after our first use of the word consumption, and then dropping it (p. 175).

To formalize this simple model, let *U* stand for induced expenditures (since AM use the more mnemonic *I* for imports), *A* for autonomous expenditures, and *Y* for income. We then have, neglecting disturbances,

(1) $$U + A = Y,$$

(2) $$U = f(Y),$$

or

(2a) $$U = g(A).$$

No essential complications are introduced by subdividing *U* into components, provided all can be treated as induced by the same income total, or by subdividing *A* into components. Equations (1) and (2a) may be reduced-form expressions for much more complicated models.[18]

Similarly, no essential complication is introduced by having more than one income total, provided the differences between them consist of items that can themselves be regarded as induced.[19] The situation is different, however, if different components of *U* are induced by different income totals, which themselves differ by noninduced components, or if we are interested in forecasting, or determining, an income total which differs by a noninduced component from the one that can be regarded as inducing *U*. The model must then be made more complex.

B. *A Slightly More Complex Model*

To illustrate, suppose it is desired to predict *Y* of equation (1), but that *U* is regarded as induced by some other income total, say *Y′*. Equation (2) then becomes

(3) $$U = f(Y').$$

Let us define

(4) $$A_2 = Y' - U$$

(5) $$A_2 = A - A_1,$$

so that we can rewrite (1) as

(1a) $$U + A_1 + A_2 = Y$$

and the counterpart of (2a) as

(2b) $$U = g(A_1).$$

If either A_1 or A_2 can itself be regarded as induced by either Y or Y'—which is not ruled out by the fact that their total is autonomous—this model will reduce to the one of the preceding section. Otherwise, there is no way of solving (1a) and (2b) simultaneously, except by specifying A_1 and A_2 separately. Of course, in a fuller model, this may be done by making A_1 and A_2 functions of variables other than Y or Y'. But on the present level, there is no alternative to having separate figures for A_1 and A_2. AM's theoretical exercise in their Section I (iii) comes from trying to force a model like the one of this subsection into the mold of the one of the preceding subsection by treating A_2 as induced.

C. *The Use and Testing of a Simple Model*

The use of a simple model to "predict" or "forecast" or "determine" income involves four steps: (i) forecasting in some way the value of A (or of A_1 and A_2 separately); (ii) estimating in some way, presumably from past experience, the form and parameters of equation (2a) [or (2b)]; (iii) using the value of A (or of A_1) in equation (2a) [or (2b)] to forecast U; and (iv) adding U and A to forecast Y.

It is obvious that the income-expenditure theory per se is used only in steps (ii) and (iii). Some other approach (which could, for example, be the quantity theory) is used in step (i) to forecast. A. The *usefulness* of the model depends on how easy it is to forecast or, for policy purposes, determine A, and on how large U is relative to Y, but the *goodness* or *validity* of the model depends only on how accurately it forecasts U. If we are to compare the goodness of this model with an alternative, we must either (a) use the alternative to forecast only U also and see how well they do the same job, assuming for both that the values of the independent variables they contain are known with certainty; or (b) append to both models a procedure (i.e., another model) for forecasting their independent variables (A for the income-expenditure model, M for the quantity-theory model) and compare how well the two alternative pairs of models forecast Y. Needless to say, if procedure (b) is used, there is no way of attributing any difference in results to the separate parts of the joint models (hypotheses) tested except by reverting to procedure (a).

In our paper, we used procedure (a). The quantity theory, we said, implies a relation between M and Y. However, we have no "standard" alternative theory for forecasting A and, in any event, our aim is to compare the quantity theory with the income-expenditure theory, not with a joint theory. Hence, we said, we shall revise the quantity theory by taking it to imply a relation between U and M. This puts the quantity theory at a disadvantage, but it enables us to make a valid comparison; if that comparison happens to be in favor of the quantity theory, then the "true" superiority of the quantity theory is greater yet. If it happens to be in favor of the income-expenditure theory, the verdict will strictly speaking have to be "unproved," though obviously if the difference in favor of the income-expenditure theory were large, there would be some presumption against the quantity theory.[20]

It follows from this discussion that, in the simple model, what should be correlated with the autonomous variable depends on how the autonomous variable is defined. We used C as the dependent variable because that turned out to be our U. Had we used A^* as the autonomous variable, we would have had a different U and would have used a different dependent variable.

It follows also that, by using the more complex model, this procedure can be rendered meaningless in a way like that outlined at the beginning of the lengthy quotation given above from pages 181–83 of our original paper. Designate as A_2 all items of income that are highly variable and not closely related to other components of income. Separate the rest of income into two parts, U and A_1, which show smooth and regular movements. These two parts will be highly correlated, so knowledge of A_1 will give excellent estimates of U. If we then *assume* that A_2 as well as A_1 is known with certainty, the error in estimating total income from A_2 and A_1 will be small. However, what has been done in the process is essentially to scrap the income-expenditure theory as an explanation of fluctuations in income. The major fluctuations are relegated to A_2 and this is *assumed* to be known by some other unspecified theory and to exert no influence on income through the multiplier process. The great appeal of the income-expenditure theory is precisely that it professes to explain movements in aggregate income through the multiplication of fluctuations in autonomous expenditures. The whole question that the theory is designed to answer is simply begged if the part of income that is hard to predict is assumed to be outside the multiplier mechanism.

We stress this possibility because it is precisely what DM have done and explains the high correlations that puzzle them. In their Table 2, the concepts of autonomous expenditures in the final two columns (gross fixed private domestic investment and this plus exports) cover only a small part of the items that need to be added to consumption (their dependent variable) to get income (however defined). And the omitted items are the most variable and hardest to predict.[21] Instead of finding an answer to the initial question, they have in effect found a question to which there is an answer. This same comment applies to their so-called “basic tests” of 20 different definitions of autonomous expenditures.

IV. Some Specific Issues

Four specific issues require attention: (a) alternative income totals, (b) our consumption function, (c) AM’s and DM’s specific models, (d) the introduction of M^* to replace M.

A. *Different Income Totals*

The existence of several income totals that may be of interest derives primarily from the undistributed income of corporations, and the absence of any way to evaluate governmental output analogous to the market valuation of private output. Three such totals were implicitly considered by us and explicitly considered by AM.

1. Y^d = Personal disposable income on a cash basis. On the factor-payments side, this is the current income of consumer units on a largely cash receipts basis after payment of direct personal taxes.[22] On the product side,

$$Y^d = C + S, \tag{6}$$

where C is personal consumption and S is saving defined by equation (6) and hence is a cash concept of saving, excluding any saving in the form of undistributed corporate earnings. Note that consumption is at market prices, so has as its counterpart not only factor payments but also indirect and direct business taxes.

2. Y = Personal income after direct taxes on an accrual basis. This total is connected to Y^d by

(7) $$Y = Y^d + R + W,$$

where R is the undistributed income of corporations, and W is the excess of wage accruals over wage payments.[23] On the factor-payments side, Y, like Y^d, is also the current income of consumer units after payment of direct taxes, but on a largely accrual basis, R adding the increase in the equity of consumer units in corporations, as measured by undistributed net corporate income after taxes, and W converting wage payments from a cash to in accrual basis. On the product side,

(8) $$Y = C + A,$$

where A is what we treated as saving of consumers in our analysis. On the product side, A can also be regarded as the sum of private capital formation plus governmental expenditures on goods and services financed by borrowing (or, on one interpretation, government capital formation).[24]

3. N = Net national product. This total is connected to Y by

(9) $$N = Y + T,$$

where T can be regarded as governmental expenditures on goods and services financed by taxes.[25] It is not easy to give N a simple interpretation on either the factor payments or product side.

On the factor-payments side, insofar as the taxes are direct taxes (personal income taxes, and corporation profit taxes), N includes total factor payments to individuals on an accrual basis and before taxes. However, it includes more than that, namely, indirect business taxes. There is no simple way to identify these with factor payments which is why they are deducted in computing still a fourth income total used by neither AM nor us, currently termed national income, sometimes termed national income at factor cost, to distinguish it from net national product, sometimes termed national income at market prices. Hence, N has no clear interpretation as a sum of factor payments to individual consumers or income recipients. An alternative possible interpretation in terms of factor payments is to regard government as an ultimate income recipient strictly parallel to the class of individual consumers rather than as an intermediary serving them. T is then the government's net income, $C + A$ the income of the private sector.[26] The total has no special significance for the class of ultimate individual income recipients.

On the product side, the situation is no better. We can write N as

(10) $$N = C + (K + E - I) + G,$$

or as the sum of the market value of (i) consumption expenditures, (ii) private capital formation (net private domestic investment plus the net foreign balance) and (iii) government expenditures on goods and services. The first two have a clear meaning. Both are values of what are regarded as final products. But the third is a mixture of expenditures on (1) intermediate products, which should be excluded to avoid double counting, (2) final products for consumption, which in principle should be treated as part of *C*, and (3) final products for governmental capital formation, which should be added to private capital formation. The difficulty is that it has proved virtually impossible, despite many attempts, to classify actual governmental expenditures into these categories. The result has been what so often happens: the statisticians measure what they can rather than what economic theory calls for, and then they and economists who use the figures attempt to rationalize what can be measured as what "really" should be measured.[27] The current wide use of *N* reflects this process, not any resolution of the basic theoretical ambiguities about its meaning.

In an accounting sense, *N* can be regarded as including indirect business taxes twice: once as part of the value of goods and services included in consumption and private capital formation; and once as part of governmental expenditures. This is the item mentioned above as not capable of being identified as a factor payment and which is therefore deducted (along with a number of other items) from Net National Product to compute National Income.

Despite its ambiguities, *N* (adjusted for price change) may sometimes be a better index of total "activity" or "output" than *Y*. Although *N* includes too much, *Y* includes too little because some allowance should be made for final output of government. The truth, therefore, seems to be in between. DM implicitly add a fourth income total, Gross National Product, derived from *N* by adding capital consumption. This is a hybrid concept that has no clear theoretical meaning.

B. *Our Consumption Function and Model*

Y is the income total that we treated as inducing consumption expenditures. AM imply that we should have used Y^d and state that the use of *Y* involves "grievous misspecification" of the consumption function—though in their footnote 1, they acknowledge that recent work suggests including in the relevant income total one component of the difference between *Y* and Y^d (R = corporate savings) and three paragraphs later in their text they dismiss the other components of the difference ($H + W + T^f$) as "minor reconciliation items." We share the view described in their footnote but not accepted by them, that recent work by them, by one of us, and by others recommends the use of *Y* rather than Y^d. However, we were led to use *Y* not by such considerations but by the empirical evidence and theoretical criterion we used in choosing a concept of autonomous expenditures.

We are not sure we fully understand AM's criticism. Perhaps their point is not that Y^d would have yielded an empirically better consumption function but that we should have used Y^d to be faithful to the prior Keynesian usage.[28] If that is their point, it has much merit. While we relied primarily on our own empirical tests, and only secondarily on the literature, we did not in fact test an autonomous concept corresponding to Y^d. In view of the widespread use of Y^d as the argument of consumption functions, we should have done so.

Y is also the income total that we treated as the one that it is desired to forecast or determine. Finally, *A* is the total that we designated as autonomous. Hence our decisions made our model of the simple type described above in Section III.A.

If we had decided to treat *N* as the total that it is desired to forecast or determine, our model would still have remained of the same simple type, since we explicitly regarded *T* as induced, not autonomous. Hence in that case, we would have defined *U* as equal to $C + T$, and have correlated $C + T$ alternatively with *A* and with *M*.[29] We have tried this for periods since 1929, and our conclusions are not much affected.[30]

C. *AM's Consumption Function and Model*

AM explicitly assert that they treat consumption as induced by Y^d rather than *Y*, which is what gives rise to most of the complexities of their discussion. They describe *N* as the income total they desire to forecast or determine, and $Z^a + X^a$ as the total of expenditures they regard as autonomous. The model they finally use can be restated in terms of the simple model described above in Section III.A, though like them we have found no easy way to rationalize it. We have

$$\begin{aligned} N &= (C - X^a + Z^i) + (Z^a + X^a) \\ &= U^* + A^*, \end{aligned} \tag{11}$$

where U^* is what AM implicitly regard as the net total of induced expenditures, and A^* is what they explicitly designate autonomous expenditures.[31] If *Y* rather than *N* is treated as the income total to be forecast or determined, $U^*—T$ becomes the induced component of *Y*, since AM regard the bulk of *T* as induced. Minor items aside, U^* equals consumption plus net inventory change minus imports minus transfer payments.

AM describe $C^f = C + Z^i$ as the total of induced expenditures, but this is clearly wrong in the context of a model that purports to explain the whole of income by autonomous expenditures plus multiplier reactions. As just noted, induced expenditures are either $(C-X^a+Z^i)$ or $(C-X^a+Z^i-T)$ according as *N* or *Y* is the income total to be determined. AM report that they were led to use C^f by a comment made in a letter by one of us. The point made in that letter is the one made above in Section III.C, namely, that in using the income-expenditure theory to explain income, only induced expenditures are explained by the theory. Unfortunately, they interpreted the wrong total as induced. For testing their concept of autonomous expenditures neither *C* nor C^f is the relevant total and we can see no advantage to using the one rather than the other. C^f would be the relevant total if Z^a alone were regarded as autonomous.

DM do not recognize that the whole problem is to test a model for predicting or determining a concept of income, but rather treat the problem as one of predicting consumption, presumably for its own sake (contrast their specification of the rival hypotheses in their Section III with our Section III.C above). They assert (in their Appendix) that they regard Y^d as inducing consumption. They are not explicit about what income total they wish to determine and do not recognize that what is relevant is total induced expenditures and that this total shifts with a change in the concept of autonomous expenditures. Interpreted in our framework, the induced component that is implicit in their procedure, if *N* is the income total to be predicted is as follows:

Concept of Autonomous	Induced Component
A^{**}	$U^{**} = U^{*} +$ State and local government expenditures on goods and services
A^{***}	$U^{***} = U^{**} -$ Capital consumption allowances
Gross Fixed Private Domestic Investment + Exports	$U^{****} = U^{***} +$ Federal Government Expenditures on income and product account
Gross Fixed Private Domestic Investment	$U^{****} +$ exports

In each case, these are the induced expenditures that should be correlated with autonomous expenditures. If Y is the income total to be predicted, T should be subtracted from each induced component.[32]

It may be worth introducing a few numbers to show how AM and DM have emasculated the income-expenditure model. Let us take the mean values for the period 1947–58 for illustrative purposes and compare the division between induced and autonomous elements for their models and our model, and for Y and N as alternative income totals to be determined (Table 3).

By our model, we in effect treated the income-expenditure theory as saying: if you know from other sources what is going to happen to roughly one-tenth of Y or N, then the multiplier analysis will tell you (or give you an estimate of) what will happen to the other nine-tenths. AM convert the model into one that says: if you know from other sources what is going to happen to nearly half of Y or over one-third of N, then the multiplier analysis will tell you what will happen to the other half of Y or two-thirds of N. DM's two models treat only slightly less of total income as autonomous. If AM and DM were to continue along this line of "improving" the model by having it predict a smaller and smaller percentage of income more and more accurately they would soon arrive at the point where it is predicting nothing—perfectly! In the old saw, with such friends, the income-expenditure theory hardly needs any enemies.

Properly to test the relative goodness of the AM model or the DM models and the quantity theory we must, as noted earlier, either force the quantity theory to predict $U^{*}-T$ (or $U^{**}-T$ or $U^{***}-T$), if Y is to be determined, or U^{*} (or U^{**} or U^{***}), if N is to be determined, or else supplement the income-expenditure theory by some model that will predict A^{*} (or A^{**} or A^{***}), so that we can test the ability of the pair of models to predict Y or N compared with the ability of the quantity theory to do so. We do not know how to do the second, so we must perforce do the first. In doing so, as we shall for the AM model but not, for want of patience, for the DM models, it should be emphasized how much we are stacking the cards against the quantity theory. It is one thing to change the quantity theory from a relation with income to a relation with a total that accounts for some 90 per cent of income. If the quantity theory works well for the first it should also work well for the second. It is a much more drastic wrench to convert the quantity theory into a relation with half to two-thirds of income. There is no theoretical reason why, even if the quantity theory works well for income as a whole, it should work well for half to two-thirds of income. Hence, this time, there is far more point to the caveat that if the quantity theory under these handicaps does better than the income-expenditure theory, it must have a much greater "true" superiority but that if it does worse, the evidence is inconclusive.

TABLE 3—DIVISION BETWEEN INDUCED AND AUTONOMOUS EXPENDITURES, FOUR MODELS, MEAN VALUES, 1947–58

(Y=\$254.4 billion)

(N=\$319.7 billion)

	Model			
	FM	AM	DM	
	A	A^*	A^{**}	A^{***}
Autonomous expenditures (billions of dollars)	27.2	115.0	80.4	106.2
Induced expenditures (billions of dollars) in determination of				
Y	227.1	139.3	174.0	148.2
N	292.4	204.7	239.3	213.5
Percentage of expenditures autonomous in determination of				
Y	10.7	45.2	31.6	41.7
N	8.5	36.0	25.1	33.2
Percentage of expenditures induced in determination of				
Y	89.3	54.8	68.4	58.3
N	91.5	64.0	74.9	66.8

Table 4 shows correlation coefficients and the percentage errors of estimate for regressions of induced expenditures on A^* and M, for the two alternative totals to be determined (Y and N) and for the four periods we used earlier. These are the relevant correlations for comparing the AM version of the income-expenditure theory with the quantity theory.[33]

TABLE 4—CORRELATION BETWEEN MONEY OR AUTONOMOUS EXPENDITURES AS DEFINED BY AM AND INDUCED EXPENDITURES, FOR TWO ALTERNATIVE INCOME TOTALS AND FOUR PERIODS, ANNUAL DATA, 1929–58

Period	Correlation Coefficient Between U^*-T and		Standard Error of Estimate as Percentage of Mean Value of U^*-T for Regression of U^*-T on		Correlation Coefficient Between U^* and		Standard Error of Estimate as Percentage of Mean Value of U^* for Regression of U^* on	
	A^*	M	A^*	M	A^*	M	A^*	M
1929–41	.790	.814	10.6	10.0	.929	.928	7.1	7.1
1930–41	.858	.921	8.0	6.0	.960	.979	5.2	3.9
1947–58	.904	.958	6.3	4.2	.980	.976	3.7	4.0
1929-58 excl. 1942–46	.984	.991	9.0	6.9	.996	.990	5.2	8.6

Symbols:

M=Money

*A**=Autonomous expenditures as defined by AM

U^*-T=Induced expenditures, when Y is income total to be determined

U^*=Induced expenditures, when N is income total to be determined

T=Taxes less government transfer payments, net interest paid by government, and subsidies less surpluses of government enterprises

If *Y* is taken to be the total to be determined, the quantity-theory correlations are consistently higher than the income-expenditure correlations, even for 1929–41, though the difference for that period is negligible. If *N* is taken as the total to be determined, the quantity theory is in a dead heat with the income-expenditure theory for 1929–41, has a slightly higher correlation for 1930–41, and a slightly lower one for 1947–58. Clearly, these results, even waiving all our doubts about whether A^* is a valid definition of autonomous, if anything reinforce our earlier conclusions and certainly do not contradict them.

One interesting feature of Table 4 is the sharp difference between the performance of the income-expenditure theory in determining Y and N. The correlations are much lower and the standard errors much higher when $U^* - T$ is the induced component than when U^* is. There is a difference in the same direction for the quantity-theory correlations, but of appreciably smaller magnitude. The reason for this feature of the table is the point made earlier in discussing different income totals: the double counting of indirect business taxes in N. These give U^* and A^* a common element. The taxes (or part of them) are already included in the market value of consumption items entering into U^*; they are also included in A^*, via government expenditure, as *AM* defines it. The result is to make the correlation between U^* and A^* spuriously high. On the other hand, when T is subtracted from U^* to give $U^* - T$, an opposite bias may be introduced, since while the common element is properly subtracted out, so is the rest of T, so that errors of measurement, which enter positively into A^* enter negatively into U^*. It is likely therefore that the correlations for the Y income total are biased against this version of the income-expenditure theory; the correlations for the N income total are biased in its favor. Hence a valid estimate of its performance would be in between.

D. *The Use of M* Instead of M*

We accept as valid the theoretical idea leading AM to substitute M^* for M, namely, that some part of the movement in M may itself be induced. We do not, however, accept their particular expression of the idea. For the period since the end of World War II, M^* and M must be very highly correlated, though we do not have the estimates of M^* with which to check this conjecture. For the period before World War II, M^*, if we understand its construction, is a poor variable on theoretical and empirical grounds. M^* is described as "the estimated maximum amount of money (in the conventional definition) that could be created by the banking system on the basis of the reserves supplied by the money authority (except in response to commercial bank borrowings), account being taken of reserve requirements and currency holding habits," where apparently reserve requirements are interpreted as meaning *legal* reserve requirements but currency holding habits as meaning *actual* habits.[34] This inconsistency is a fatal flaw. The theoretical difficulty it raises can be readily seen by applying the concept to a banking system with no legally required reserves (which, certainly until recent years, has been the situation in most of the world). M^* on this basis is limited only by currency holding habits and completely unaffected by banking behavior; if the ratio of currency to deposits were to approach zero, or banks could issue currency, M^* would approach infinity. For M^* to be meaningful, the reserve requirements for which it is calculated must be the banks' desired (perhaps long-run or secular desired) reserve requirements, not legal reserve requirements. For the period since World War II, and also most years prior to 1929, these did not differ greatly although there were significant changes in the relation between the two.[36] But beginning in 1932, increasingly after 1933, and until World War II, desired reserves differed widely from legally required reserves. This was not a difference in any sense "induced" by contemporaneous changes in income. It was a reaction to the failure of a third of the banks from 1929 to 1933, the resulting disillusionment with the Federal Reserve System as a "lender of last resort," and the slow acceptance of the FDIC as an effective substitute. Banks decided that required reserves were no reserves at all, and that they had to provide their own liquidity by keeping reserves in excess of requirements. M^* during this period must be much larger than the actual money stock and is not at all comparable with its values before 1929, from 1929–32, or after World War II. To regard it as the variable relevant to the quantity theory makes about as much sense as to regard the investment variable relevant to the income-expenditure theory as being the maximum investment that could have been made if business had borrowed the funds that banks could have loaned if they had eliminated legal excess reserves. Hence, we regard all of AM's correlations containing M^* as worthless and have made no attempt to repeat them for the separate subperiods.

V. Comments on AM Appendix

This Appendix, which is the second to which we have written a reply,[36] represents a step in a highly desirable direction—the formulation of more complex alternative hypotheses of income determination than those we considered and the comparison of their conformity with experience. Unfortunately, it is as yet an unsuccessful experiment because the models formulated, while complex, are unsatisfactory.

The Appendix falls into two rather distinct parts: (1) the first two-thirds which presents the models and a qualitative analysis of them; (2) the final third, which attempts a quantitative analysis and comparison.

1. *The first two-thirds*. This part of the Appendix (through the discussion of Table A.1) strikes us as a rather extended exegesis of a few sentences in our original paper, namely:

> In practice, M and A [money and autonomous expenditures] are positively correlated and that is to be expected under either of the theories under consideration. Under the quantity theory, changes in M involve changes in all components of income, including both consumption and other components. Under the income-expenditure theory, changes in M influence rates of interest and thereby affect investment and consumption; in addition, changes in investment may induce corresponding changes in M via their effect on banks. Hence, a positive simple correlation between A and C may simply be a disguised reflection of the effect of M on C; alternatively, a positive simple correlation between M and C may simply be a disguised reflection of the effect of A on C (p. 177).

Hence we agree with the general conclusion of this part, that it is generally not possible to discriminate among the kinds of hypotheses in question simply be examining the statistical significance of individual correlations, or the signs and statistical significance of separate regression coefficients. That is why we instead *compared* relations computed in as similar a way as possible and for as many subperiods as possible.

Our only disagreement with this part is that AM do not carry their analysis to its logical conclusion. They assert that the EDO (effective demand only) and MO (money only) models are contradicted: the EDO model because the coefficient of M^* in their equation (4.4a) is not zero; the MO model because the quantity of money demanded and the quantity supplied have been shown elsewhere to be affected by interest rates. Both results do contradict the specific models in question as they formulate them. But at least the assertion about the MO model simply reflects a defect in the way AM formulate it. They are required to rule out interest-rate effects in their MO model because they have left out the price level. Restore the price level to the model, as it should be restored, and it is entirely consistent with an MO model that both the *nominal* quantity of money supplied and the *real* quantity demanded should be responsive to interest rates. Indeed, we doubt that there has ever been a reputable quantity theorist who did not believe that both these effects were present, at least in the short run. Such a theorist would have gone on to deny, and correctly, that this implied that interest rates were, in the long run, determined by monetary factors. It is this proposition, very different from the one considered by AM, that distinguishes the quantity theorist from the income-expenditure theorist—or, perhaps better, the liquidity-preference theorist.

2. *The final third*. The rest of the Appendix is a different matter. Here AM quite properly go beyond the coefficients of each equation separately and, in effect, compare different equations. Here they try to find quantitative comparisons comparable to those that we used that can discriminate among the different theories but that are based on more complex and general models than those we used. This is an admirable direction in which to proceed. However, the reliability of the results depends critically on how satisfactory the more complex models are. Unfortunately, the particular models AM use seem to us so unsatisfactory as to render their numerical results of no interest for the purpose for which they are intended.

The chief defects in their models seem to us to be:

(a) *The omission of the price level*. We noted in our paper: "Both sets of equations [quantity theory and income-expenditure] have a major gap, and incidentally, the same gap. . . . Neither . . . says anything about the proportion in which changes in income will be divided between output and prices." Yet this division "may have a significant effect on the numerical values of the multiplier and of the velocity and on their stability" (pp. 171–72). If the simple models are to be elaborated, the introduction of a price level seems like the first and most important elaboration that is required. (Some of our calculations are among real magnitudes but we did not satisfactorily integrate our equations involving real and nominal magnitudes.)

(b) *The nonhomogencity of the equations*. As AM correctly note in their footnote 31, "In order for the subsequent developments to be strictly valid, the functions c, f, L, and β must be homogeneous of degree one in the flow variables and the stock of money." This is related to the preceding point, since, on a theoretical level, one way to justify leaving out the price level is to assume that the relevant clasticities are unity with respect to both prices and real magnitudes and hence that the response can be calculated from nominal magnitudes alone. Yet the equations (B.1) to (B.4) that AM use as their approximation to a more complex and general system are not homogeneous because of the presence of interest-rate terms. According to that system, doubling all flow variables and the stock of money would imply doubling the interest rate! And this is so regardless of whether the change consists of a change in prices alone or in real magnitudes alone.

(c) *The use of $M^* \cdot M^*$* plays a critical role in the equations and the tests. Yet for reasons outlined in our text, M^*, as AM have apparently defined it, is so logically defective and so empirically misleading a variable that it is hard to attach much meaning to results in which it plays a critical role.

(d) *The use of $C^f \cdot C^f$* also plays a critical role in the equations. Yet this is a strange construct that AM were led to only by a misunderstanding on their part of a suggestion one of us made to them in connection with some of their earlier work. It has no clear economic significance in the context of comparing different models such as U^* or U has.

VI. Conclusion

What is the net outcome of this interchange? In one respect, it is rather surprising. A reader of modern economic textbooks would conclude from them that the Keynesian Revolution produced a carefully formulated, logically coherent theory of income determination that had the great virtue of using concepts that have direct and readily available empirical counterparts. He might also conclude, though more from the tone than from any explicit statements, that the theory had been tested by evidence and had passed the tests with flying colors.

The reader of this interchange will come out with a very different impression. The central concept of the income-expenditure theory—autonomous expenditures that are the driving force of income—apparently has more symbolic than substantive content. All of us use words to describe it—like "independent," "uncorrelated with the residual error," "exogenous"—that are figurative rather than operational. We proposed an operational criterion for choosing its empirical counterpart that we regard as far from satisfactory and that our critics apparently regard as beneath mention. They rely on what is "beyond reasonable doubt" or "plausible." One of our critics states flatly that: "Theory or 'intuition' is necessary to specify components of autonomous expenditure,"[37] without, of course, specifying what the content of the one theory is that is to give content to the other theory, or how different "intuitions" are to be reconciled.

Even in the absence of an unambiguous criterion, there might be a large measure of agreement about what is the most useful empirical counterpart for autonomous expenditures, as there is about the counterpart for money. However, that is clearly not the case, in view of the richness of the alternatives that have been served up.

In our paper, we may have settled on a concept that was particularly unfavorable to the income-expenditure approach and as a result may have been led to too sweeping a conclusion about its lack of conformity with experience. Certainly, the fishing expedition conducted by our critics has turned up some constructs that are highly correlated with some components of income for some periods. However, our construct at least had the merit that it was derived by a specifiable and reproducible procedure, and that it was fitted into a system of constructs that accounted for the whole of the income total we regarded the income-expenditure theory as trying to explain. The alternative constructs have neither merit. They all leave loose ends of income dangling. They all treat a high fraction of income as autonomous and so limit sharply the scope of the income-expenditure theory. What the calculations of our critics do is to establish a presumption that further research along similar lines may be more rewarding than we thought was likely on the basis of our earlier work.

It is worth recording that our critics have all accepted the monetary relations we presented and have not experimented with any wide range of alternative concepts of money. All remark on the closeness of the monetary relations, and with one possible exception (Hester), do not question that monetary changes play an independent role that is more than the disguised reflection of changes in autonomous expenditures. Ando and Modigliani have gone farther and experimented with a different monetary concept, M^*, designed to separate out the autonomous monetary element in income change. We agree that this is a desirable direction in which to move, but question the validity of their particular formulation.

The task we set ourselves was to compare, in as unambiguous and unbiased a fashion as possible, the relative stability of monetary velocity and the investment multiplier. None of the calculations made by our critics for supposedly the same purpose is correct because they omit some components of income for the income-expenditure calculations, set the two theories different tasks, or use lengthy periods combining two different subperiods. We have made some of the correct calculations for one of the alternative concepts of autonomous expenditures (Ando and Modigliani's). Though less clear-cut, the results are in the same direction as those from our original calculations. Hence, we are left with no reason to change our earlier conclusion that "so far as these data go [and, we may now add, those adduced by Ando and Modigliani, DePrano and Mayer, and Hester] the widespread belief that the investment multiplier is stabler than the monetary velocity is an invalid generalization from the experience of three or four years. It holds for neither later nor earlier years" (p. 166).

Notes

[1] The only other systematic attempts preceding ours to assess empirically the relative usefulness of the two leading theories for interpreting short-term economic change with which we are familiar are: Clark Warburton, "Monetary Expansion and the Inflationary Gap," *American Economic Review* (June 1944) pp. 302-27, and "The Volume of Money and the Price Level Between the World Wars," *Journal of Political Economy* (June 1945) pp. 150-63; Milton Friedman, "Price, Income, and Monetary Changes in Three Wartime Periods," *American Economic Review* (May 1952); Milton Friedman and Gary Becker, "A Statistical Illusion in Judging Keynesian Models," *Journal of Political Economy* (February 1957) pp. 64-75; Karl Brunner and Anatole Balbach, "An Evaluation of Alternative Monetary Theories," *Proceedings*, Western Economic Association (1960).

[2] See Donald D. Hester, "Keynes and the Quantity Theory: A Comment on the Friedman-Meiselman CMC Paper," "Reply to Donald Hester," and Hester's "Rejoinder," *Review of Economics and Statistics* (November 1964).

[3] To some extent this is unfortunate. We believe that further experimentation with measures of money, particularly in the direction of weighted aggregates of different categories of monetary assets, may well be extremely fruitful. See Roy Elliott, "Savings Deposits as Money," unpublished Ph.D. dissertation, University of Chicago (1964); J. G. Gurley, *Liquidity and Financial Institutions in the Postwar Period*, Study Paper No. 14, U.S. Congress, Joint Economic Committee, Employment, Growth, and Price Levels, Washington (1960); Edward J. Kane "Money as a Weighted Aggregate," *Zeitschrift für Nationaleconomic* (1964), pp. 221-43.

[4] There is some duplication between this reply and our reply to Hester, whose critique, as AM and DM note, to some extent overlaps theirs. We have tried to reduce duplication by cross-reference but have not been able to eliminate it entirely. See Friedman and Meiselman, *op. cit*.

[5] All otherwise unidentified page number references are to Milton Friedman and David Meiselman, "The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958," *Stabilization Policies*, Commission on Money and Credit (New York 1963), pp. 165-268.

[6] Of the total variance of consumption for the 25 years, 88 per cent is accounted for by the differences between the means for the two subperiods.

[7] In the present version of their paper, AM refer in footnotes 23 and 39 to calculations for the periods 1929-41 and 1947-58 separately. These references did not appear in the original version, though the present reference to unspecified alternative "subperiods" in their footnote 8 did.

[8] We are indebted to Ando and Modigliani for making some of their data available to us.

[9] However, for 1929-39, $r_{CA}*$ is higher than r_{CM}, though lower than r_{CA}.

[10] The correlation between A and $A*$ is .977 for 1929-41, .883 for 1947-58.

[11] Though we plotted all the first differences and gave some correlation coefficients in the text, we did not publish tables of our first difference correlations comparable to those we published for correlations among levels. This was a mistake that we have tried to rectify by publishing the basic tables in our reply to Donald Hester, Tables 1 and 2, pp. 375 and 376, *op. cit*.

[12] The *sum* of squared residuals is, of course, reduced by including *A*, but by a smaller fraction than the number of degrees of freedom is reduced. Hence the mean square residual per degree of freedom is raised.

[13] Many of AM's other results are also drastically changed by considering each period separately. For example, they conclude that including Z^a and X^a in the correlation separately rather than only their sum reduced the residual variance by about one-third. For the separate periods, the residual variance is sometimes larger, sometimes smaller. For some periods, the coefficient of X^a turns out to be negative.

[14] We leave out of this section the charge that we have misspecified the consumption function because we deal with that in Section IV. B, below.

[15] AM do not even mention our criterion for choosing the concept of autonomous expenditures, or our explanation of why it is that the decision to include the net foreign balance rather than exports in *A* is entirely consistent with our regarding imports as being induced, and similarly, our decision to include the net deficit rather than government expenditures, with our regarding taxes as being induced.

DM at least note in their footnotes 15 and 16 that we explicitly considered alternative definitions, but they do not discuss how we did so.

See our reply to Hester for a fuller discussion of these points, *op. cit*., pp. 374-75.

[16] Casual empiricism runs rampant throughout both papers, as the following examples, for which no evidence is cited, testify. From AM:

1. "The extreme short term variability of undistributed corporate income makes it unlikely that consumption would significantly respond to it from year to year" (footnote 1).
2. "Since *S* includes terms correlated with *a*, it will in general be itself correlated with *a*." This is an empirical not logical statement since the correlations of separate terms with *a* may be such as to cancel.
3. "The expression in brackets . . . is positively correlated with *a*." This and the preceding item exemplify the kind of untested assertions on which AM base their charge of "least-squares bias."
4. A hypothetical model for which no evidence is given and which is for an economy in which there are no corporations and no foreign trade is cited as one of three items justifying the proposition that "the laborious battery of tests presented by FM is basically irrelevant for the purpose of assessing the *empirical* usefulness of the income-expenditure framework." (italics added).
5. "The condition of *given* income is more likely to be binding in the full employment situation, even in terms of current prices," i.e., prices are constant at full employment (footnote 6).

From DM:

1. "These two consumption functions are likely to show substantial difference in predictive power."
2. "If, as seems highly probable, consumption is a function of the previous year's income as well as of current income. . . ."
3. "An increase in consumption has a positive effect on inventory investment predominantly in the same year. On the other hand, an increase in consumption can lead to unplanned disinvestment in inventories"—so anything can happen. AM are less indefinite—they come out solidly for the negative relation.
4. "The government deficit . . . is . . . partially endogenous . . . tax receipts are clearly endogenous . . . exports [are] largely autonomous."
5. "Since an increase in domestic consumption" raises "prices" (footnote 20). This footnote contains several other fine examples.

[17] In his rejoinder to our reply, Hester offers a criticism of our criticism of our criterion that has some merit. He points out that our criterion handles only one questionable item at a time and may give misleading results because of omitted items. This is true, but he has suggested no alternative that is free of this defect.

[18] A simple example will perhaps clarify the point in question. To avoid complications, assume there are no corporations and no government but there is foreign trade. Using AM's notation, let

$$(a) \qquad Y = C + K + E - I,$$

$$(b) \qquad C = a + bY,$$

$$(c) \qquad I = c + dY,$$

so *C* and *I* are both induced and both by *Y*.

If both *K* and the net foreign balance are treated as autonomous (as we did in our paper) then, in the notation of equations (1) and (2),

$$(d) \qquad U = C$$

$$(e) \qquad A = K + E - I,$$

and equation (2a) becomes

$$(f) \qquad C = \frac{a}{1-b} + \frac{b}{1-b} A.$$

In addition, *I* can also be expressed as a function of *A*, yielding

$$(g) \qquad I = c + \frac{ad}{1-b} + \frac{d}{1-b} A.$$

Alternatively, if *K* and *E* are treated as autonomous (as AM do), then in the notation of equations (1) and (2)

$$(h) \qquad U = C - I$$

$$(i) \qquad A = K + E.$$

Equation (2a) becomes

$$(j) \qquad U = \frac{a-c}{1-b+d} + \frac{b-d}{1-b+d} A.$$

C and *I* can also be expressed separately as functions of *A* yielding:

$$(k) \qquad C = \frac{a-bc+ad}{1-b+d} + \frac{b}{1-b+d} A.$$

$$(l) \qquad I = \frac{c-bc+ad}{1-b+d} + \frac{d}{1-b+d} A.$$

Note that the same basic model [equations (a), (b), and (c)] is compatible with alternative specifications of autonomous expenditures, but that these affect the specific character of the reduced form model. The model specified in AM's equations (4), (5)′, and (6)′ is another example. See also our reply to Hester for a fuller discussion of this point.

[19] In the example of the preceding footnote, one can readily derive gross domestic output, *Y + I*, as well as *Y* from the model.

[20] We do not understand AM's footnote 24, which seems simply wrong. If, as they say, "the only meaningful measure for performance of these models is the residual variance of NNP given the autonomous variables," and if, as in the present instance, some of the autonomous variables are components of NNP, then, as we pointed out in our paper (see quotation above from our pp. 181-83), this residual variance can easily be made as small as anyone likes by choosing the autonomous variables to be the counterpart of a larger and larger part of NNP. We can make sense out of their statement only if the proviso is attached that the autonomous variables are items other than components of NNP which is not how they apply it. For example, the quantity theory model would satisfy this proviso, the quantity of money being the autonomous variable. See also Friedman and Becker, *op. cit*.

[21] For example, inventory investment has a correlation of only 445 with *C* for 1953-63.

[22] We say "largely" because of the difficulty of getting from the national income accounts any totals strictly on a cash or accrual basis.

[23] AM include two additional items on the right-hand side, namely, *H*, the statistical discrepancy, and T^f, foreign transfer payments by government. The first has nothing to do with the concepts involved, only with the particular method of measurement. The Department of Commerce measures some items from the payments side, some from the value of products. Given the accounting identities, this means that there are in fact two different estimates of each item. The inclusion of the statistical discrepancy is a signal that some of the items have been taken from one set of estimates, some from the other. Its inclusion serves to see to it that all items are in effect measured from the same side.
AM are correct in saying that we overlooked T^f. It should not conceptually be included in either *Y* or *A*, yet our numbers apparently include it. Since the item is small, indeed non-existent for most of the period our data covers, its inclusion cannot materially have affected our results.

[24] This interpretation is somewhat arbitrary because it implicitly involves matching specific governmental expenditures with specific receipts. The governmental item included in *A* is the deficit on income and product account. Its calculation implicitly treats all government transfer payments as being paid out of taxes rather than borrowing.

25 The *casual* of the preceding footnote is relevant here also. As an accounting matter, T is total taxes less government transfer payments, not interest paid by government, and subsidies less current surpluses of government enterprises.

26 Along the lines of the preceding footnotes, this requires treating transfer payments, etc., as a prior charge against receipts in computing the net income of government.

27 The use of gross rather than net national product is the most glaring other example of this process in national income accounts. See the early volumes of *Studies in Income and Wealth* for extensive discussion of these issues.

28 Incidentally, as an aside on the "search for the highest correlation," Y^d is generally more highly correlated with M than Y. See our Appendix Table II-Al, p. 244.

29 There is a typographical error on our page 256 that misled us and also DM. On line 14, the expression $C+T$ should have been $C - T$.

30 The standard errors of estimate as a percentage of the mean value of $C+T$ are as follows for the regressions of $C+T$ on A and M:

	$C+T$ on	
	A	M
1929-41	5.0	6.0
1930-41	5.3	4.2
1947-58	11.7	4.3
1929-58, excl. 1942-46	17.7	11.0

31 Minor items aside, their total A^* differs from our A in three main ways: by treating (1) inventory investment as induced rather than autonomous; (2) total exports rather than the net foreign balance as autonomous; (3) total government expenditures for goods and services rather than the deficit or surplus on income and product account as autonomous. See our Appendix A for an explicit discussion of the alternatives in (2) and (3) and for the evidence that led us to decide these as we did. We did not consider item (1). DM's autonomous A^{**} is the same as A^* except that they include only federal government expenditures on income and product account, omitting entirely state and local expenditures. Their A^{***} adds capital consumption to A^{**}. Neither AM nor DM, as noted, give any empirical evidence justifying their decisions.

32 Another way to interpret what DM have done is in terms of the more complex model of Section III.B above, with their autonomous concepts as A_1, and consumption regarded as their U and as induced by Y'=C+A_1. But this leads to such unusual consumption functions that we have preferred the alternative of interpreting their calculations in terms of the model of Section III.A (e.g., for A^{***}, which is the concept they generally prefer, the alternative interpretation would lead to regarding consumption as a function of N plus capital consumption allowances plus imports less inventory investment and less state and local government expenditures for goods and services).

33 First-difference correlations, or correlations using a lagged value to adjust for trend, would be also, but we have not calculated them.

34 This interpretation is based on the paper by Teigen referred to by AM. Since it was contained in the first draft of our reply and AM have not indicated that we have misinterpreted them, we assume that it is correct.

35 See Phillip Cagan's forthcoming NBER monograph, *Determinants and Effects of Changes in the Stock of Money*.

36 We mention this for the benefit of those readers who may have seen AM's original manuscript which was widely circulated in mimeographed form. Their present Appendix is almost entirely new.

37 Hester, *op. cit*., p. 377.

10/3/12