THE PHILIPPINE STATISTICIAN
Volume XVII, Nos. 1/2, 1968

ON THE SPECIFICATION OF INVESTMENT FUNCTIONS[1]

by

J. ENCARNACION[2]

1. **Introduction and Summary.**— Ever since Keynes emphasized the role of investment demand in the determination of aggregate output, a considerable amount of theoretical and empirical research has been devoted to the formulation of investment functions. More recently, since investment is capital accumulation, widespread interest in the problems of economic growth and development has added to the volume of research in this area.[3] In principle, economic theory is supposed to guide the specification of an investment function, which is then estimated from empirical data. But while much work has gone into the theory and the estimation parts of this process, relatively little attention has been given to the specification problem. By this we mean the problem of stating an investment function in such a way that both **a priori** theoretical considerations and the limitations of data and practical estimation methods are taken into proper account. For obvious reasons, the standard procedure is to express investment as a linear or log linear function of the explanatory variables. In this paper we suggest a simple revision of this procedure in order to meet certain **a priori** theoretical considerations which have usually been ignored in the empirical literature. The revision will en-

[1] Work on this paper was supported in part by research funds made available by the Ford Foundation to the Wisconsin U.P. Program in Development Economics. The author has had helpful discussions with Professors R. W. Hooley, G. P. Sicat and J.G. Williamson on certain aspects of investment functions; but needless to say, responsibility rests with the writer.

[2] Professor of Economics, School of Economics, University of the Philippines.

[3] For reviews of this literature, see [5] and [11]. More recent references are given in [9].

tail some additional estimation problems, but these do not seem insuperable.

Empirical work on investment functions is necessarily a compromise with the requirements of existing theories of behavior. Data limitations may prevent the inclusion in the estimated equation of some variables considered relevant. In case the necessary data are available, collinearity between two variables may lead the investigator to eliminate one of them in order to get statistically significant results.[4] An investment function that has been specified for purposes of statistical estimation needs to be understood in a **ceteris paribus** sense, i.e. the function is meant to hold only for given values of the omitted (though relevant) variables. It is for this reason that different investment functions may be compatible with one another, in the same way that different relationships involving partial derivatives are compatible when they result from the same functional relation.

Although the rate of interest, for example, is presumably a relevant variable, its omission from an investment function may be justified if its variation is considered minor over the period of time under study. The estimated equation may give misleading results under major changes in the interest rate, but this is only to be expected. No single equation can be expected to represent **the** investment function, and any equation is based on the implicit assumption that the values of the omitted but relevant variables fall within certain narrow ranges. Thus the omission of a relevant variable is not itself a defect, and the failure of an estimated equation to perform well in quite different circumstances is not a valid objection to it, provided the specification is reasonable from a theoretical viewpoint. To evaluate a specification in this regard, it suffices to examine only the included variables.

In the following discussion, we confine our attention to the dependence of investment on the variables appearing in

[4] As T.-C. Liu [12] has pointed out, however, such decision may be quite wrong under certain circumstances; it may be that the right decision is to include another appropriate variable in the equation.

each equation. It will be clear that although the investment functions to be considered are linear, our basic criticism applies also to log linear formulations. Sections 2 to 4 discuss alternative formulations using one explanatory variable, lagged values of the same variable, and two or more explanatory variables. We argue that the investment functions considered are incorrectly specified, and that the reason in all three cases is the same: they ignore important parameters—either certain particular values of the variables or functions of them—that affect the investment decision. An explanatory variable will have different effects on investment depending on whether its value exceeds or falls short of the corresponding parameter. Section 5 makes a suggestion for taking this point into account by introducting such parameters into the specification of investment functions.

2. **One Explanatory Variable.—** We recall Chenery's work [2] on comparing a version of the acceleration principle (2.2) with what he calls the capacity principle (2.1). Let

$$k_t = \Delta K_{t+\theta} / K_{t+\theta-2}$$

and consider the alternative formulations

$$k_t = b(X_t/K_t) - b\lambda \qquad (2.1)$$

$$k_t = b(\Delta X_t/K_t) + C \qquad (2.2)$$

where K is capital stock measured in terms of capacity, X is output, b is a reaction coefficient, θ is the leg between output changes and investment, λ is the "capacity factor" or "optimum degree of utilization of plant," and C (a constant term) takes account of non-accelerator-induced investment. Time is measured in years, and $\Delta X_t = X_t - X_{t-1}$. For the sake of uniformity in the statistical procedure, Chenery assumed a common two-year interval for capacity change in the six industries in his sample. Both equations are expressed in ratio terms; otherwise, the capacity principle would have the form

$$\Delta K_{t+\theta} = b(\beta X_t - \lambda K_t)$$

where βX_t is the amount of capital required for current output. Since capital is measured in terms of capacity (defined theoretically by the point of tangency of the existing plant curve and the long-run cost curve), the acceleration coefficient $\beta = 1$.

From a **priori** considerations, Chenery argues that in an industry where large economics of scale exist it would be normal to operate at a relatively lower λ, excess capacity being maintained in anticipation of increasing demand. Moreover, b may be expected to be lower in such an industry, since any need to expand would be less pressing in view of the excess capacity. This **a priori** expectation seems to be confirmed by the empirical data, which show some correlation between b and λ. Chenery's main results suggest that the capacity principle gives a better explanation of investment behavior for industries with relatively low values of $b\lambda$, while the acceleration principle does better where $b\lambda$ is higher and accordingly excess capacity is less.

For our present purposes, Chenery's work illustrates two points. First, the two formulations appear useful in different circumstances. Since the acceleration principle takes no account of excess capacity, its weakness in describing investment behavior in industries that normally maintain excess capacity should not be surprising. On the other hand, it seems to do better (compared to the capacity principle) where there is little or no excess capacity. Again, one should probably expect such a result, for in this case, the explanatory value of λ as a parameter becomes rather small. To take the extreme possibility, suppose that $\lambda = 1$ and $b = 1$. Then equations (2.1) and (2.2) become, respectively,

(2.3) $$k_t = X_t/K_t - 1$$

(2.4) $$k_t = \Delta X_t/K_t + C$$

It is clear that (2.4) would give a better fit to time series data when there is an upward trend in fluctuating demand, as was the case for the period studied (1922-39). For the presence of the term C in (2.4) easily allows for negative values of $\Delta X_t/K_t$, but values of X_t/K_t less than 1 can be accommodated by (2.3) only through physical depreciation.

This brings us to the second point, viz. that the coefficient of the change-in-demand variable depends on the sign of this variable. Chenery has observed that the usefulness of "both principles would be improved by introducing different values for the reaction coefficient for negative predictions. The interpretation given to a decrease in demand in an industry with a long-run upward trend is quite different from the reaction to an increase" [2, p. 23]. The investment function (2.1) and (2.2) are thus incorrectly stated in terms of the included variables themselves. Chenery has therefore suggested estimating separate regressions for upward and downward changes in demand so that (2.2), for instance, would become

$$k_t = \begin{cases} b'(\Delta X_t/K_t) + C \text{ for } \Delta X_t \geqq 0 \\ b''(\Delta X_t/K_t) + C \text{ for } \Delta X_t < 0 \end{cases}$$

where the coefficients b' and b'' are generally different.

That the accelerator operates in different ways for upward and for downward changes in demand is well recognized, but the common practice is still to assume a linear relationship and then make adjustments when the statistical results are unsatisfactory. Adjustment are usually made by ignoring negative values of the variables, either registering them as zero or suppressing them altogether in estimating the equation. What would be preferable would be to have a correct specification in the first place, which requires incorporation of the fact that the value $\Delta X = 0$ is essentially a behavioral parameter. Investment behavior when $\Delta X_t < 0$ is different from when $\Delta X_t > 0$.

3. **Lagged Values of the Same Variable.** The naive specification of the acceleration principle—that investment depends only on the current change in output— has generally done poorly in a statistical sense; distributed lag formulations seem to do better, but not necessarily in an economic sense. In

Eisner's most recent study of the subject [4], he estimated several equations including the following:

$$(3.1) \quad i_t = b_0 + \Sigma_{t=1}^{7} b_j \Delta s_{t+1-j} + \Sigma_{j=8}^{9} b_j \, p_{t+8-j}$$

where i_t is the ratio of capital expenditures in year t to gross fixed assets at some base time; Δs_t is the ratio of the change in sales, from year t-1 to year t, to some average change in sales; and p_t is the ratio of net profits to gross fixed assets. It is not clear in what sense Δs_t, which involves sales during t, should be an explanatory variable for i_t, which involves capital expenditures during the same time period. There is a necessary time lag between knowing the sales figure for the year and the decision to invest on account of the current sales change, and a further lag between the decision and actual capital expenditures. But for present purposes, we shall suppose that Δs_t is relavant.

Eisner used data for 1955-62 from over 300 firms in 10 industries. The coefficients of the sales change variables were found to be significantly positive, but his "firm time series" regression shows the sum of these coefficients to be rather low (0.244). However, since the "cross section of firm means" regression gives a higher figure (0.629), Eisner remarks that "this result is consistent with the hypothesis that firms would view variations in their own sales experience over a relatively short period of time (the eight years from 1955 to 1962, in this regression) as in lesser part permanent than the differences between their own average sales experience and the average experience of all firms in the economy" [4, pp. 374-75]. Nonetheless, one would have expected rather higher figures on the basis of the acceleration principle.

Moreover, although one would expect—for annual data—that the sales change coefficients should be smaller in value the greater the lag, this is not always borne out by the various regressions reported by Eisner. For instance, in the "firm time series" results, $b_4 = 0.024$ and $b_5 = 0.030$, the standard error of estimate being 0.010 in both cases. It is also curious, as

Eisner has noted, that the current profits coefficient is negative in several regressions, in some cases quite significantly. His tentative explanation is that "perhaps higher current capital expenditures cause higher current depreciation charges and higher current interest payments, and also entail 'start-up' or other costs, all reducing current net profits" [4, p. 372]. These factors would reduce the size of the coefficient, but it is hard to see how they could make it negative. The more likely intepretation seems to be that current profit is **not** an explanatory variable, for it is contemporaneous with what is to be explained.

Our interest in the specification problem lies in a different direction, however, and to this we turn. Consider a firm whose investment behavior is represented by equation (3.1) and which has experienced an increasing rate of growth of sales, $\Delta s'_t > \Delta s'_{t-1} > \ldots$, the primes denoting specific values of the Δs. Let its percentage increase in capital be $i'_t > 0$. According to (3.1), we can have different sequences of Δs (taking account of the coefficients b_j) which would also produce the same value of i_t. In particular, there would be a sequence such that $\Delta s''_{t-1} < \ldots$ with the investment result $i''_t = i'_t$. But this implication of (3.1) is surely false for any reasonable magnitudes of the b_j. We expect that a firm's investment decision when its sales growth has been declining should be qualitatively different from when this has been increasing. Yet it would be easy to construct numerical examples where $\Delta s''_t < \Delta s''_{t+1}$ 0, which on the basis of theoretical considerations is likely to lead to some disinvestment, but $i''_t = i'_t > 0$ according to (3.1).

Thus we conclude that equation (3.1) is an incorrect specification, for while it considers changes in demand over several past periods as relevant, it ignores the radically different effects of upward and downward changes in the rate of growth of sales. Eisner's distributed lag formulation allows for (generally) decreasing weights attached to the lagged values of the sales-change variable, but their relative magnitudes are also relevant. The effects on the investment decision when

$\Delta s_t - \Delta s_{t+1} > 0$ and when $\Delta s_t - \Delta s_{t+1} < 0$ are inadequately represented by the coefficients of Δs_t and Δs_{t+1}.

4. **Substitution between Variables.** For our final illustration, we consider a formulation based on a Koyck-type distributed lag function [10]. Suppose that

$$(4.1) \quad I_t = r\alpha\Delta X_{t-1} + (1-r)r\alpha\Delta X_{t-2} + (1-r)^2 r\alpha\Delta X_{t-3} + \ldots$$

$$= \sum_{j=1}^{\infty} (1-r)^{j-1} r\alpha\Delta X_{t-j} \qquad 0 < r \leqq 1$$

where α is the desired capital-output ratio and r is a reaction coefficient. The assumption is that a fraction r of any discrepancy between the desired and the existing capital stock is corrected by investment in the following period, so that I_t is the sum of the terms on the right hand side of equation (4.1). The first term is due to the change in output between periods t-1 and t-2, which created a capital discrepancy of $\alpha\Delta X_{t-1}$ on its account, so to speak. The second term is due to the change in output between periods t-2 and t-3, which created a capital discrepancy that was partly reduced by investment during t-1 to the extent of $r\alpha\Delta X_{t-2}$, leaving still the amount $(1-r)\alpha\Delta X_{t-2}$ to be corrected. Accordingly, one gets the second term, and so on.

From (4.1) it follows directly that

$$(4.2) \qquad I_t = (1-r)\, I_{t-1} + r\alpha\Delta X_{t-1}$$

by writing the expression for $(1-r)I_{t-1}$ and then subtracting it from I_t. Equation (4.2) has the great convenience, which was part of the rationale for Koyck's original formulation of geometrically decreasing coefficients, that it can be readily estimated to yield estimates of r and α.[5] (There are certain pro-

[5] A.D. Brownlie [1] claims that Chenery's capacity principle is equivalent to (4.1 because $I_t = r(\alpha X_{t-1} - K_{t-1})$ can be put in the form of (4.2). But the very reason for introducing the capacity principle was that the "capacity factor" $\lambda < 1$ for industries with economies of scale, while Brownlie simply assumes that $\lambda = 1$.

blems about the bias of least-squares estimators here, but that is outside our present interest.)

Following Koyck's lead, A.D. Brownlie [1] added the profit rate as an explanatory variable and assumed that the investment equation (involving ratios) is

$$(4.3) \qquad i_t = \sum_{j=1}^{\infty} c^{j-1}(a_1 x_{t-j} + a_2 p_{t-j}) \qquad 0 \leqq c < 1$$

where $x_t = (X_t - X_{t-1})/X_{t-1}$. Outside of mentioning its possibility, Brownlie gives no reason why the same weight c^{j-1} should be associated with both x_{t-j} and p_{t-j}, $j=1,2,\ldots$. The economic rationale for the coefficient of x_{t-j} is based on the reaction coefficient r which indicates the fraction of capital discrepancy that is covered by investment in the period following, and the size of the reaction coefficient would be determined by a variety of factors—the degree to which the firm considers a change in sales as "permanent" rather than "transitory", the facility (technical and financial) with which changes in capacity can be made, etc.— that are generally different from those which affect the investment decision through lagged profits. A Koyck-type distributed lag function can be interpreted in terms of an adaptive expectation model [3, pp. 206-208], but in any event, that the declining weights would be precisely the same as those for lagged output changes is generally to be unexpected.

The convenience in assuming equality, of course, is that (4.3) gives

$$(4.4) \qquad i_t = ci_{t-1} + a_1 x_{t-1} + a_2 P_{t-1}$$

which can be estimated for c, a_1 and a_2. Brownlie obtained fair statistical results (e.g., $R^2 = 0.46$) from inter-industry cross-section data. As it stands, equation (4.4) might be rationalized by arguing that as an explanatory variable, the lagged investment rate represents the effect of past growth, including the effect of government policy on a particular industry; the lagged change in output takes account of the acceleration principle; and the profit rate is a measure of internal investment funds plus the accessibility of external funds.[6] Our concern now is whether or not the specification is admissible.

According to (4.4), x_{t-1} and p_{t-1} are substitutable in the usual sense that alternative combinations (as determined by their coefficients) of values of these two variables would result in the same value of i_t. Suppose that $x'_{t-1} = 0$ and $p'_{t-1} = p^*$, where p^* is the "normal rate of profit and the primes denote particular values of the variables. In this case, let us say that $i'_t = 0$, as might be expected from theoretical considerations. If (4.4) is correct, there is a pair of values $x''_{t-1} < 0$, $p''_{t-1} > p^*$ such that $i''_t = i'_t$. But we should expect that notwithstanding a profit rate better than normal, $i''_t < 0$ because of the decrease in output. A high level of profits constitutes merely a permissive factor in that greater investment is made possible if there is a prior decision to undertake investment, but the presence of the permissive factor does not necessarily lead to a positive decision.

5. **A Suggestion.** We have argued in the preceding sections that the usual investment function formulations are incorrect—in terms of the included explanatory variable themselves—because they do not account for the different effects on investment when the explanatory variables fall within certain ranges. The simplest way to represent these effects is to use min and max concepts in the specification.

Suppose, for example, that we are considering ΔX_t to explain I_{t+1}. Suppressing the subscript t, we can write

$$(5.1 \qquad I_{+1} = c_1 \max(0, \Delta X) + c_2 \min(0, \Delta X)$$

[6] A few years ago, R. W. Hooley and the author experimented with a somewhat similar equation using data published by the Philippine Statistical Survey of Manufactures. The disappointing results are summarized in Hooley and Sicat's important study [8, pp. 26-29] of investment in the Philippine manufacturing sector based on company records.

Thus if $\Delta X > 0$, the effect on investment is given by the coefficient c_1 since in this case, $\max(0, \Delta X) = \Delta X$ and $\min(0, \Delta X) = 0$. If $\Delta X < 0$, the relevant coefficient is c_2.

A profits variable can be included in the equation by adding

$$(5.2) \qquad c_3 \max(0, p - p^*) + c_4 \min(0. p - p^*)$$

to the right hand side of (5.1). (5.2) would give different roles to the profit rate, depending on whether this exceeds or falls short of the "normal" figure. If $p = p^*$, (5.2) is zero, so that investment would be due to the other explanatory variables.

We can consider the effect on investment due to a changing growth rate of output by using.

$$c_5 \max(0, \Delta X - \Delta X_{-1}) + c_6 \min(0, \Delta X - \Delta X_{-1})$$

or something similar in the specification.

There may be some value of a variable beyond which the marginal effect on investment is zero. For instance, if the profit rate exceeds p^{**}, say, the effect on investment attributable to this variable may conceivably be no greater than what it profits to expenditure on research and development, rather than to plant and equipment (cf. [6]). In this case we would have

$$\min [p^{**}, \max(0, p - p^*)]$$

in place of

$$\max(0, p - p^*)$$

as the appropriate term.

By means of such reformulations, we take account of nonlinearities in the causal relationships without sacrificing the convenient property of linearity in the equation itself. The main objective, of course, is to specify investment functions closer to economic theory, and a **min** and **max** formulation—

which allows for kinks and constraints—may help to explain why different empirical studies often come out with divergent results concerning the importance or non-importance of the various determinants of this paper, it all depends on the range in which an explanatory variable falls.

Parameters like p^* could be estimated independently, after which the problem would be the estimation of the coefficients c_j Suppose our equation is (5.1) with (5.2) added to its right side. One could try to estimate four separate regressions: one using observations: one using observations where $\Delta s > 0$ and $p > p^*$, another where $\Delta s > 0$ and $p < p^*$, etc. This procedure would yield two estimates for each of the coefficients c_1, . . ., c_4. In general, if the equation contains no constant term and each explanatory variable appears in one **max** term and one **min** term, we would have 2n coefficients for n variables. Computing 2^n separate regressions yields $n2^n$ coefficients, so that there would be 2^{n-1} estimates corresponding to each of the coefficients of the original equation. (It is only when n=1 that we would have a single estimate.) This is not a difficulty, however; we should expect to have different estimates of the coefficients depending on what ranges of the variables we are considering.

Some interesting problems are raised by this procedure. For example, how do we calculate the quantitative effect of a policy change that puts a firm (an industry, the entire economy) in a different regime? But this and other questions are for further study.

REFERENCES

[1] A.D. Brownlie, "An econometric analysis of investment in New Zealand manufacturing 1957-8", *Economic Record*, Vol. 38, No. 82 (1962) pp. 157-166.

[2] H.B. Chenery, "Overcapacity and the acceleration principle", *Econometrica*, Vol. 20, No. 1 (1952) pp. 1-28.

[3] C. F. Christ, *Econometric Models and Methods*. New York: Wiley, 1966.

[4] R. Eisner, "A permanent income theory for investment: some empirical explorations", *American Economic Review*, Vol. 57, No. 3 (1967) pp. 363-390.

[5] R. Eisner and R. H. Strotz, "Determinants of business investment", in D.B. Suits *et al.*, *Impacts of Monetary Policy*, Englewood Cliffs, N.J.: Prentice-Hall, 1963, pp. 59-337.

[6] J. Encarnación, "Constraints and the firm's utility function", *Review of Economic Studies*, Vol. 31, No. 2 (1964) pp. 113-120.

[7] Z. Griliches and N. Wallace, "The determinants of investment revisited", *International Economic Review*, Vol. 6, No. 3 (1965) pp. 311-329.

[8] R. W. Hooley and G.P. Sicat, "Investment demand in Philippine manufacturing", U.P. School of Economics Discussion Paper No. 67-2 (revised), June 1967. (129 pp., Mimeo.)

[9] D.W. Jorgenson, "The theory of investment behavior", in R. Ferber (ed.), *Determinants of Investment Behavior*. New York: National Bureau of Economic Research, 1967, pp. 129-155.

[10] L.M. Koyck, *Distributed Lags and Investment Analysis*. Amsterdam: North-Holland, 1954.

[11] E. Kuh, "Theory and institutions in the study of investment behavior", *American Economic Review*, Vol. 53, No. 2 (1963) pp. 260-268.

[12] T.-C. Liu, "Underidentification, structural estimation, and forecasting", *Econometrica*, Vol. 28, No. 4 (1960) pp. 855-865.