Mario Arcelli's selected papers on economics (1967-1977)

Edited by Federico Arcelli
Foreword by Giorgio di Giorgio

2004

Federico Arcelli **Giorgio Di Giorgio**

~~**Appunti Europei**~~

Mario Arcelli's selected papers on economics (1967-1977)

Federico Arcelli **Giorgio Di Giorgio**

FOREWORD

(nota introduttiva da fare – lasciare due pagine)

Federico Arcelli **Giorgio Di Giorgio**

MARIO ARCELLI: A LIFE FOR ECONOMICS

As a professor of economics, Mario Arcelli has been indeed, since his very early teaching years, in his favourite environment. The strong bond with the academic world represents the key to understand a whole life of research efforts joined to several important achievements as civil servant and in the business community. In fact, when we look to all the public appointment he got, as well as all the positions that he held in corporate boards, we discover that those were always a part-time commitments for him, and the only real continuing activity has been university. Mario Arcelli was descendant of an old Northern Italian family, originated from the town of Piacenza. He was indeed born in Milan on May 21st, 1935. He got married in 1967 with Maria Gabriella Legrenzi and they had two sons.

His bond with economic studies started with the degree he got in 1957 from Università Commerciale Luigi Bocconi – Milan, being awarded of top marks and distinction. In the following years he was assistant researcher in the same university, starting a long track of publications with some articles for a review promoted by Assolombarda; at the same time he was involved in research and consulting projects in collaboration with several banking and industry groups based in Milan.

In 1963 he got the first step in the academic "cursus honorum" being appointed "libero docente", a more stable professor role, at the University of Trieste, and few years later, in 1967, being only 32, he became full professor in Economics in the same university. In the same period his scientific works were published on several reviews in Italian, English and French. In particular he was author of some "selected papers" for the Review of Economic Politics, and for Quality and Quantity; in Italian he wrote several articles for the "Rivista di Politica Economica", for "Economia Internazionale delle Fonti di Energia", for the "Giornale degli economisti e annali di economia", and, mainly, for "L'Industria" and "Il Risparmio".

From 1969 to 1973 he took the chair of Economics at the Università di Padua, quitting that place for a one-year visiting period at the Massachusetts Institute of Technology (M.I.T.) of Boston (USA), where he met Paul Samuelson, Kenneth Arrows and Franco Modigliani, with whom he maintained contacts in the following years.

In 1974 he became full professor of economics at the Università "La Sapienza" of Rome, where he served for about fifteen years, until 1989, and directed the Department of Economics between late Seventies and early Eighties. Being in Rome, he joined several editorial committees and, since 1979, he took the position of editor of "Economia Italiana" and its English homologue, the "Review of Economic Conditions". He maintained both this position until his death in 2004. At the same time Mario Arcelli was regularly writing on some Italian reviews (among others "Bancaria", "Banche e Banchieri" and "Rivista di Politica Economica") and he was columnist, as economic expert, for Rome's newspaper "Il Tempo", between 1974 and 1986. Since 1986 and up to 1997 he was columnist for "Il Messaggero".

Between the late Seventies and early Eighties Arcelli was member of several government committees, and acted as advisor of the Banca d'Italia, the Italian Central Bank. He also was member of the scientific advising committee of Confindustria, the association between Italian industrialists. Two of those committees were of particular importance: the so-called "comitato Spaventa" (after the name of its chairman) on the conditions of public debt in Italy and the "Comitato Tecnico Scientifico" of the ministry of Budget and Economic Planning, that was the main consultive body for economic matters within Italian government. Of this latter he got his first appointment in 1984, resigning only when he served as Minister of Budget and Economic Planning 1996. From 1993 he has been appointed chairman.

Mario Arcelli has been for a decade one of the main economic advisors of Italian governments. He served as Head of the Economic Affairs Department at the office of the Prime Minister (Presidente del Consiglio dei Ministri) of the Republic of Italy between 1981 and 1983 (during Fanfani V and Spadolini I and II cabinets). He was also economic advisor of the Prime Minister in 1987 (Fanfani VI cabinet) and 1988-89 (De Mita cabinet). In those roles he participated to the Italian delegation to five G7 summits, the ones held in Versailles (France), Williamsburg (USA), Venezia (Italy), Toronto (Canada), Paris (France).

Is commitment as civil servant got a special place aside to his academic activity. His major role in advising Italian governments on economic matters led him to be appointed in February 1996 to the position of Minister for

Budget, Economic Planning and European Affairs. The appointment was the natural consequence of the resignation of the previous minister, Rainer Masera, whom Mario Arcelli was advisor to. He took this role in a really delicate moment of Italy's history, during the Italian six-month turn as president of the Council of Europe and when the position of Minister of Treasury was vacant (the Prime Minister, Lamberto Dini, was also acting Minister of Treasury). As a representative of the president of the European Union, he chaired the "Ecofin" summit in Venice (May 1996) and continued the process that led, two years afterwards, to the birth of the Euro.

In 1989 he started teaching at LUISS (then LUISS Guido Carli) university of Rome. At the same time he took the direction of the newly established "Osservatorio e Centro di Studi Monetari" ("OCSM") in the same university. He contributed to the creation of this new centre, and this has been one of the main reasons for him for quitting "la Sapienza". Few years later, in 1992, he was appointed rector (chancellor) of LUISS, by a board of directors still chaired by Guido Carli (who will die in 1993); Carli was the man who leaded the change from the old university "Pro Deo" in the LUISS in 1978, and he maintained the position of chairman of the administrative body until his death. Some years later, being Mario Arcelli the rector, the university has been renamed after Guido Carli, becoming "LUISS Guido Carli". Mario Arcelli, who linked to LUISS University the most relevant part of his academic path, quitted the post of rector only when he retired, in 2002. At the end of October 2003 he has been appointed to the honorary post of "professore emerito" by the university.

During his long academic activity, Arcelli published a number of articles and books. Among them we may recall the eight editions of his textbook of economics, but also many others. Since the publication of *Variazioni qualitative dei fattori e progresso tecnico*, he wrote close to fifteen books and around 150 articles (many of them for the two newspapers published in Rome, Il Tempo and Il Messaggero, in the period comprised between 1975 and 1997), including his papers for the OCSM of LUISS Guido Carli and for the Review of economic Conditions in Italy, of which he was editor.

In consideration of his major contribution to the development of economic studies, Mario Arcelli has been appointed in 1995 member of the Accademia Nazionale dei Lincei, one of the world most important academic institutions. He has been also deputy chairman of the Italian Economic Society (S.I.E.) and member of the board of the Istituto per la Enciclopedia Italiana, the editor of Treccani Encyclopaedia. He was also for many years in the awarding committee of the Fondazione Romeo ed Enrica Invernizzi.

Mario Arcelli had also a really important managerial track record. He has been member of the board of directors of ENI, the Italian oil giant (1980-82), member of the board of the Ufficio Italiano Cambi, or "UIC" (1985-88), member of the board of directors of Finanziaria Ernesto Breda, a leading mechanical group (during the first half of the Eighties) and deputy chairman of Banco di Roma (1986-1992). For this latter Bank he served as acting chairman in 1992, before the merger with the Cassa di Risparmio di Roma and the establishment of the new Banca di Roma (now Capitalia). After the merger he quit all responsibilities in the new group. In1989 he was appointed chairman of newly established Rasbank, leading it to be fully operational as a Bank – getting all licences and authorisations in 1990 – and continuing to chair its board until the last day of his life, for over 15 years. Rasbank has been a pioneer in Italy for internet banking and trading, and it is now the main non-insurance subsidiary of RAS group. Mario Arcelli has also been member of the board of directors RAS SpA (1998-2004), of the Società per la Bonifica dei Terreni Ferraresi (1995-2004), of Pininfarina SpA (1995-2002) and of Italcementi SpA (1998-2004).

In 1998 he was elected chairman of Assogestioni, the Italian association between investment funds and fund managers. His appointment coincided with a period of expansion of the market in Italy and around the world and with a new role taken by this body, that become a relevant voice in the debate over financial investments, rules and market control in Italy. After four years, the statutory limit to chair the association, he stepped down in 2002.

At the Christmas 2001 holy mass in St Peter, in Rome, Mario Arcelli red the address to the Pope as representative of the Italian universities.

Federico Arcelli **Giorgio Di Giorgio**

~~DA FARE~~

~~CARICARE CORREZIONI SETTEPANI (ANCORA DA RICEVERE)~~

~~CARICARE CORREZIONI DE ANGELIS (CI SONO è DA FARE, prima di martedì 11 feb simonetta non può)~~

Federico Arcelli **Giorgio Di Giorgio**

~~CA~~CHAPTER I
TECHNOLOGY AND CAUSALITY

Federico Arcelli **Giorgio Di Giorgio**

FOREWORD

The two following papers were written by Mario Arcelli in the mid-Sixties. They are among the first articles written by Mario arcelli and touch quite broad and different aspects of the economic theory, with main focus on quantitavie methodology.

The first one, “Causal order and dependencies in linear economic structures” is a study on the development of econometric research and especially about the fruitfulness of a definition of economic structure in mathematical terms. The second paper benefited of the work that he did to comment and translate into Italian some studies of Andreas Papandreou, who treated amply the subject of the method of economics.

The object of “Quality Changes in Factors and Decomposition of Technical Progress in Aggregate Production Functions” is to analyse the components of technical progress in an aggregate production function, considering first of all the effects of changes in the quality of factors.

In the first paper Mario Arcelli studies in depth the close connections existing between the causal ordering and causal inference. He points out that the associations between causal ordering and identification and the extension of the analysis of regression into the most comprehensive analysis of causal dependence have contributed to making these concepts increasingly operational and have permitted their satisfactory inclusion in an organic theoretical framework. He stresses, anyway, the importance of the judgement of the economist, also in order to avoid a misuse of the mathematical and statistical tools. In fact Mario Arcelli recalls that “the results obtained should not, however, give an excessive confidence in the instruments available. In this sense, we should like to conclude by harking back to a sentence in one of our previous essays: the improvement of econometrics often lies, more than in the use of the most refined technical instruments, in a more accurate collection of data, under the control of a strict scientific methodology”.

The work made in the second paper aimed “to place in evidence the links existing between a series of preponderantly qualitative variables and the rate of change in productivity, while at the same time providing a clearer idea of some economic development mechanisms. A critical assessment of the empirical repercussions of the acceptance of certain hypotheses to interpret reality will be able to teach us something about the solution of economic policy problems. The existence of a number of criticisms disputing the very foundations of aggregate production functions makes it necessary to consider the effects of changes in the quality of factors and to break down technical progress into its most significant components.”. Mario Arcelli idea is to demonstrate that failure to consider the effects of changes in the quality of factors already produces biased relations between variables at the level of the classical Cobb-Douglas equation, the partial elasticities of output with respect to inputs, and returns to scale are either underestimated or overestimated.

Federico Arcelli **Giorgio Di Giorgio**

CAUSAL ORDER AND DEPENDENCIES IN LINEAR ECONOMIC STRUCTURES [1]

Introduction.

The development of econometric research and especially the studies of the Cowles Commission[2] have underlined the fruitfulness of a definition of economic structure in mathematical terms. With the aid of the logic of classes and relationships this concept has been taken up with particular rigour by Papandreou who treated amply of the subject in the course of a volume on the method of economics[3]. The distinctions between the various types of structures: self-contained structures, sectional structures, complete structures) and the correlative concepts of endogenous variables and exogenous variables (in some cases more simply dependent variables and independent variables) permit the construction of an operational definition of the notion of causal ordering which links up with the studies of Simon and Blalock[4].

Econometricians have pointed out the connections existing between the determination of the causal ordering and the identification of the parameters of linear structures. The complete identifiably of a linear structure requires conditions even more restrictive than the operational definition of the causal ordering. The identification of the parameters permits their successive evaluation by one of the existing different methods of evaluation and in such a case not only the causal ordering but also the causal dependence between the variables is measured by the coefficients of regression calculated between the dependent variables and the independent variables. From these coefficients, with the corrective factor of the ratio between the variances, are obtained the coefficients of correlation, which represent another measure of the linear relationship existing between the dependent and the independent variables. How can the causal dependence between the variables be measured where the causal structure is not identifiable? An important result in this regard specifies that in stochastic linear structures with specific random variables, even in the absence of identifiably, coefficients of dependence that represent an extension of linear regression coefficients can always be determined. In the course of this essay we intend to develop the themes stated in this introduction, illustrating them with some examples drawn from recent economic models.

The structure as a system of linear equations

Discussion of the concept of structure in economics can refer to stochastic structures or more simply to structures in which casual variables do not appear. In the first approximation we shall discuss the second class of structures. The discourse can start from the nature of economic variables. As Papandreou[5] has stated "There is no doubt that they share with all other kinds of variables the property of being generic elements of sets, that is to say they may be viewed as mere placeholders into which constants – names of elements – may be substituted... The very fact, however that we wish to differentiate the economist's variables from other kinds of variables, necessitates that we go beyond this formal property and consider in general terms the nature of their

---

[1] From: Quality & Quantity, vol. 1, 1967, pages 166-191 – Marsilio Ed. Padua. The present essay develops some notes of M. Arcelli on chapters 4 and 5 of Papandreou's work *Economics as a science*, Lippincott, New York, 1958, This essay has been published in Italian in the review "L'industria", n 3, 1965, pp. 3-27.

[2] See above all the monographs 10 (Statistical Inference in Dynamic Economic Models) and 14 (Studies in Economic Method), Cowles commission, 1950 and 1952.

[3] See A. G. Papandreou, Economics as a Science, Lippincott, New York, 1958 translated in Italian by M. Arcelli (L'economia come scienza, L'Industria, Milano, 1962).

[4] With regard to H. A. Simon in particular, see Causal Ordering and Identifiability (14th monograph of Cowles Commission), some essays in the book Models of Man, J. Wiley and Sons 1957 and "On the definition of the Causal Relation", in Journal of Philosophy, 48, 1952, With regard to H. M. Blalock see "Four variable Causal Models and Partial Correlations", American Journal of Sociology, 68, 1962, pp. 182-194.

[5] See A. G. Papandreou, Economics as a science, p. 57.

differentiating characteristics". It is therefore necessary to refer to their empiric interpretation, that is to say, to define classes of acts of observations that establish correspondences with numbers that represent prices, investments, incomes, consumption etc. and more generically economic quantities.

Let $0 = \{o_1, o_2 ...\}$ be a certain set of observations that specify prices of cloths. To each corresponds a price that we shall indicate generically by *x*. It follows that the set of prices of cloths is definable by the function F, subset of the Cartesian product *OX*. In the language of the logic of relationships.

Inserire formula

Therefore not all the elements of *X* are prices of cloths but only those that form the subset corresponding to the acts of observation o.

If a variable cannot be defined by acts of direct observation, its inclusion in economic theory means to represent the place-holder of a theoretic term, the operational meaning of which will result from the context of the economic theory interpreted as a whole.

The empiric interpretation of economic variables is relevant for the purpose of defining their field of variability. An economic theory requires a consistent set of relationships among economic variables those posses an empiric interpretation.

Supposing that any formalizable economic theory could be constructed on the foundation of the classes X_1, X_2, ..., X_n, the set of all the relationships of economic space requires that all possible Cartesian products formable with the said classes should be considered. Since a relationship is a sub-set of a Cartesian product, the set of all the relationships of the economic universe will be given by all the sub-sets of all the possible Cartesian products that can be constructed with the classes X_1, X_2, ..., X_n, always bearing in mind the problem of empiric interpretation. As we deal with relationships among numerical variables, they can be defined by implicit functions. We agree to indicate with

Inserire formula

the generic relationship **Inserire lettera** with X_1^j, X_2^j, X_n^j, which are generic elements of the classes X_1, X_2 ... X_n. Let us consider a set of n relationships definable, as we have seen, through n implicit functions.

If we examine the intersection of the n equations, and this intersection is not empty, which is to say if there exists at least one n-tuple that satisfies every relationship, we shall have found the solution-set of the system of relationships. A set of relationships, the solution-set of which is not empty, defines itself as a structure. In order for the solution-set to be not empty, it is necessary that the system be consistent, that is, determined or determinable, and this guarantees also the independence of the equations that contribute to the solution.

Marschak[6] defines as structure "all properties of the equations, including properties not known a priori. A model[7] is a class of structures. Each structure is defined by the functional forms of the equations and the values of the parameters occurring in them. When the equations are fully specified, we call them structural equations". In particular, a linear structure is composed of an independent and consistent set of linear equations. In the case of economics, however, a concept of structure that disregards considerations of casual variables can be considered unsatisfactory.

In economic relationships there is always an unexplained residue that may be interpreted in terms of probabilistic distributions: this because of errors of measurement; or as consequence of numerous unidentified variables the cumulative effect of which is assimilable to a casual variable; or finally, as expression of purely casual events that disturb the fundamental economic relationship. The second is the most frequent reason for the insertion of a casual variable in an economic relationship, which transforms itself into a stochastic relationship.

We shall therefore define a structure that contains one or more stochastic relationships as a stochastic structure.

One of the assumptions most often made for them, as a working hypothesis, with regard to the problem of identifying the parameters, or for the determination of the coefficients of dependence, or for the derivation of important theorems concerning the probabilistic distribution of the dependent variables, is the independence and the absence of correlations of the casual variables, between them and with regard to the other independent variables. In economic reality this condition is not always verified, so that there are often insuperable obstacles to the solution of the problems stated above.

These are arguments to which we shall revert in successive paragraphs. It is now important to investigate the various types of structure that allow us an operational definition of causal ordering.

---

[6] See J. Marschak, Statistical Inference in Economics, 10th. monograph of Cowles Commission.

[7] The meaning of the term "model" is therefore different from Papandreou's (see Economics as a science) which contrasts it to the term theory, from which it is distinguished by a different degree of adherence to empiric reality. In the model, for Papandreou, the definition of the relevant social space would be lacking.

Federico Arcelli **Giorgio Di Giorgio**

Self-contained structures and Sectional Structures: an operational definition of causal ordering

Let there be a system of *n* relationships that form a structure. Our discussion is pertinent to linear structures. We shall say that a structure is self-contained if the number of variables is equal to the number of relationships. If, vice versa, the number of variables is more than the number of equations, the structure is sectional.

For example:

Inserire formula

is a sectional structure because the variables (x_1^j, x_2^j, x_3^j) are more than the equations. One of the possible self-contained structures, that include the preceding sectional structure, could be the following:

Inserire formula

At this point it is possible to introduce the concept of the minimum self-contained sub-set of a structure. It is a self-contained structure that does not include any autonomous structure. In the case of the system *B*, the third relationship, by itself, constitutes a self-contained structure that is also a minimum self-contained sub-set of the self-contained structure *B*. The third relationship immediately provides the solution of the variable x_3^j. By substituting the solution-value of this variable in the second equation, it too becomes a self-contained sub-set, which makes possible the solution of the variable x_1^j. Considered in isolation, however, the 2nd equation is a sectional structure. Therefore while the 3rd equation is a minimum self-contained sub-set before any substitution and can therefore be defined as of zero order, the 2nd equation, which constitutes a sectional structure, becomes a self-contained sub-set only as the result of a substitution. It can therefore be defined as a derived self-contained structure of first order. The value of x_1^j being known and carrying out the substitution in the first equation, it too becomes a minimum self-contained sub-set that allows us to determine x_2^j. We can therefore speak of a derived self-contained structure of the second order.

It is evident, therefore, that a sectional structure can assume the shape of a sub-set of a self-contained structure: we shall say complete sub-set because making the opportune substitutions it becomes a self-contained structure. In the particular case of the system *B*, the solution takes place progressively by substituting x_3^j in the second equation, from which x_1^j is obtained, which, substituted in the first equation allows us to determine the value of x_2^j. Different cases can easily be conceived where the minimum self-contained sub-set is constituted by numerous equations to be resolved simultaneously.

The distinction between self-contained structures and sectional structures also allows us to clarify the significance of endogenous variables and exogenous variables. A variable the value of which is given externally to the structure being considered can be defined as an exogenous variable. Thus x_3^j is exogenous with respect to the derived self-contained structure of the 1st. order, constituted by the second equation. The same variable is endogenous when it is determined on the inside of a structure that comprehends it, thus x_1^j is endogenous in the structure *B* or in a larger structure that includes the system *B*.

The distinction between exogenous variables and endogenous variables can in many aspects lead back to the one between independent variables and dependent variables. Exogenous variables and endogenous variables are therefore bound by asymmetrical relationships that proceed from the exogenous variables towards the endogenous variables. And it is precisely this relationship of asymmetry that is at the base of a definition of causal ordering, proposed by Simon[8] that has gained numerous followers among students of different social sciences. As is known, the concept of causality proposes themes that can be analysed according to different perspectives of research. Philosophical analysis, for example, can investigate the empiric correlations of the notion of cause; the temporal sequence of phenomena, recognizing or not recognizing in this sequence the presupposition of causality; the importance of the asymmetry of the relationship between the quantities considered etc.

The methodology of the sciences can dwell, for example, on the problems of causal inference starting from a certain group of observations: these are problems that are at the centre of the construction of any scientific theory. However, in view of epistemological criticism of the notion of cause, from Hume to the latest probabilistic interpretations of reality, the concept of cause has increasingly assumed the significance of causal ordering understood as asymmetric relationship. The concept of structure and the distinction between exogenous and

---

[8] See H. Simon, Causal Ordering and Identifiability. For other definitions of causal ordering and of the concept of cause, according to the thought of different economists see the exhaustive critical review by G. Garb, "The Problem of Causality in Economics", Kyklos, 1964, n. 4, pp. 594-611.

endogenous variables permit us, with Simon[9], a severe and original definition of this new notion of cause. Simon states, however, that the definition of causal ordering on the basis of the concept of structure, although assimilable to what "common sense" considers as causal, refers to a model – a system of equations – and not to the "real" world that the model intends to describe. The temporal sequence is no longer, therefore, essential to the definition of causal ordering, although the temporal sequence often provides a base to the asymmetry.

What is relevant is the asymmetry of the relationship and not the temporal sequence. Formulated in this way the notion of causal ordering is strictly based on mathematical relationships, nullifying all the criticisms of an epistemological type which are irrelevant with regard to the logical dimension of this definition; a definition, however, that has a precise operational meaning, as Simon's analysis has made clear and as we shall establish in a successive paragraph. This operational aspect can be connected with the logic of causal inference, as we shall see. In this way the mathematical equations of the model undergo an adequate empiric interpretation, which permits their empirical testing.

The logic of causal inference, Boudon[10] observes, "has been studied separately by a good number of methodologists of different disciplines, economists, sociologists, and biologists. The economists have analysed above all the problem of the identification of linear causal structures; the sociologists have concentrated in particular on authentic causal relationships and on the discovery of fictitious causal relationships deduced from spurious correlations; the biologists have provided methods which are able not only to infer the presence or absence of causal relationships between variables, but also to measure the causal dependence".

A first link between causal ordering and causal inference is found in the study of the conditions, which are necessary if the causal ordering is to possess operational significance. The conclusion will be reached that they are analogous to those necessary for the solution of the problem of the identification of the parameters of the structure: in fact these latter conditions can be even more restrictive. As this link has been particularly studied by economists, we shall dwell on this aspect at length. Proceeding with order, in the first place we shall expound Simon's analysis of the causal ordering, secondly, we shall consider the conditions of identifiably of the parameters according to Koopmans, and the collateral studies of Marschak, Rubin, Leipnik Hurcwiz and Wald[11]; it will then be easy to establish the link between causal ordering and identifiability. Finally, a concrete example concerning an econometric model used by the Papi Commission in Italian programming practice will serve to reveal the fruitfulness of the concepts analysed.

Causal ordering and identification of the parameters of a linear structure

Simon's reasoning on the subject of the causal ordering starts with an objection to that definition of this concept that is founded on the asymmetry of the relationship between the variables. This definition would be "essentially artificial, since the same set of observations could be represented by different structures with different causal orderings of the variables"[12].
Simon[13] in fact considers the following three systems of two equations

Inserire formula

that, as can be verified, possesses the same solution-set.

Inserire formula

However, the causal ordering of the three structures is different. The equations (l*a*, 2*a*) considered together, constitute a minimum self- contained structure of zero order. In the second system the equation (2*b*) is c complete sectional structure, or, by substituting the value of y1 obtained from (1*b*), it becomes a derived minimum self-contained sub-set of first order. In the third system both the equations (1*a* and 2*c*) constitute minimum self-contained sub-sets of zero order.

---

9 H. Simon, Causal Ordering and Identifiability.

10 See R. Boudon, Revue Francaise de Sociologie, "Méthodes d'analyse causale", VI, 1965, pp. 24-43.

11 See in particular the monographs 10 and 14 of the Cowles Commission.

12 See H. SIMON, *Causal Ordering and Identifiability*, p. 63.

13 *See H. Simon, Causal Ordering and Identifiability, pp. 63-65.*

We can also say that the structure (l*a*, 2*a*) is completely integrated; that is to say founded on interdependencies; the second structure (1*b*, 2*b*) is causally ordered with an asymmetric relationship (1*b*) → (2*b*); and the third structure is disintegrated: in fact y_1 is obtained from (l*c*) and y_2 from (2*c*) without any substitution.

If we consider the three structures as having equal significance in that each of them can be derived from the others with the substitutions indicated at the side of the systems, the causal ordering would not have any operational meaning. In fact, the asymmetric relationship (1*b*)→(2*b*) of the second system could change itself into an interdependency (1*a*) ←→ (2*a*) or into a completely disintegrated structure (1*c*), (2*c*). A more careful examination of the structures, however, allows the enucleation of a basis of distinction between the three structures. Take the first structure: altering either of the two equations (l*a*, 2*a*), y_1 and y_2 change value. In the second structure, however, both the solutions change only if (1*b*) is altered. Finally, in the third structure in order to change both the solutions the two equations (1*c*, 2*c*) have to be changed at the same time. The causal relationships expressed by the causal ordering, that is, by the asymmetry between the variables, possess therefore operational meaning only in so far as changes or interventions concerning the structure can be associated with specific complete sub-sets of the system of equations under consideration.

In other words, a person is considered endowed with decisional power (politico-economical for example) who can choose the non-zero elements of the matrix of coefficients of a linear structure, but who cannot substitute zero coefficients with other coefficients. He directly controls, therefore, the values of the non-zero coefficients. When the matrix is specified, the values of the *n* variables in the *n* linear equations of the structure are univocally determined. Whoever has decisional power, therefore, controls indirectly the values of the variables, and the causal ordering specifies which variables will undergo changes in consequence of interventions in a particular complete sub-set of the structure. The causal ordering involves, therefore, knowledge of the limits imposed on the decisions-maker as regards the alteration of the equations of the system and therefore assumes a really operational significance. An intervention that cannot substitute zero coefficients for other coefficients and vice versa, alters the structure but leaves the causal ordering unchanged: therefore, the causal ordering is an invariant property with respect to certain types of intervention in the system. The structural equations that are completely specified equations permit the singling out of the specific possibilities of intervention that do not change the causal ordering.

After all, therefore, the conditions on the basis of which the causal ordering of a linear structure can be fixed depend on the a priori determination of the zero coefficients whatever the type of expected intervention in the sub-sets of the system of equations. And it is precisely on the basis of these considerations that it is possible to single out a connection with the problem of the identification of the parameters.

The identification of the parameters of a linear structure requires that the constraints imposed on the equations exclude the existence of an equivalent structure that is a structure the equations of which are totally or partially linear combinations of the original equations. This involves the establishment a priori that certain coefficients of the matrix of the system are null. In other words, it is specified which variables will appear in every equation of the system. In a self-contained structure this is equivalent to determining the causal ordering. The stating of restrictions or constraints on the form of the equations is therefore determinant in the process of identification of the parameters. This problem has been amply dealt with by Marschak, Koopmans, Rubin, Leipnik, Wald and Simon[14].

The existence of equivalent structures renders indeterminate the problem of the estimation of the parameters. Koopmans defines a necessary condition and a condition necessary and sufficient for the identification of a structural equation of a linear system. The necessary condition is that the number of the variables excluded from the equation to be identified (or, more generally, the number of linear restrictions imposed on the parameters of the equation) be at least equal to the number N of the structural equations minus one In the number of the excluded variables, one may include the exogenous and the predetermined variables.

A necessary and sufficient condition is that it is possible to form at least one non-vanishing determinant of *N-I* order with the coefficients of the variables excluded from the structural equation. It is therefore clear that for equations in reduced form, that is, for equations that express only one endogenous variable as a function of exogenous (or predetermined) variables, identification is always possible.

The necessary and sufficient condition for the identification of the parameters implies the impossibility of forming any linear combination of the structural equations of the system, without contradicting the original structural hypotheses.

Boudon[15] expresses this condition in the following way: be the fundamental system in symbols of matrices and vectors.

---

[14] See J. Marschak, Statistical Inference in Economics; T. Koopmans, H. Rubin, Leipnik, Measuring the Equation Systems of Dynamic Economics; A. Wald, Note on the Identification of Economic Relations; L. Hurwicz, Generalization of the Concept of Identification; 10th monograph of Cowles Commissioni T. Koopmans, Identification problems in Economic Model Construction; H. Simon, Causal Ordering and Identifiability, 14th monograph of Cowles Commission.

[15] R. Boudon, "Methodes d'analyse causale", op. cit.

Inserire formula

If there does not exist any non-diagonal matrix *P* of the type

Inserire formula

so that *PA* has the same structure as *A*, and *PAx* = *Pb*, that is to say, it does not contradict the a priori hypotheses concerning the nullity of the coefficients, then the system is completely identifiable. The existence of a matrix *P* having such properties makes identification impossible for at least one equation of the system.

Simon[16], for his path, with an elegant geometric treatment demonstrates how the conditions of identification determine the coefficients, starting from a group of empiric observations. Possessing numerous observations of the simultaneous values of the *n* variables of a system in *k* linear equations where $k<n$; every observation, obtained by varying the exogenous variables, must satisfy the *k* equations. The observations can be considered as points of an n-dimensional space having as coordinates the values of the *n* variables. If every observation satisfies the *k* equations, all the observations are situated in a hyperplane of *n-k* dimensions, which represents the intersection of the *k* hyperplanes of *n-I* dimensions, each of which is the geometric image of an equation of the structure.

The criteria of identification are sufficient to determine univocally the coefficients, in that any other series of *k* hyperplanes of *n-I* dimensions that forms an intersection in the same hyperplane of *n-k* dimensions, is a linear combination of the original equations, and, as such, is excluded on the basis of the restrictions imposed.

In the case of stochastic structures, it is necessary to bear in mind the probabilistic distribution introduced by the casual variables present in the structure. “For any specific set of values of the exogenous variables, the distribution of the latent variables entails or generates through the structural equations a probability distribution of the endogenous variables”[17]. This conditional distribution, in function, that is, of the exogenous variables, is the starting datum for the problem of identification.

Two stochastic structures are equivalent if they present equal conditional distributions of the endogenous variables for any value of the exogenous variables[18]. The conditions of identifiability of the parameters that we have recognized as necessary ones for deterministic structures are therefore valid also for the stochastic structures. In their case, however, it is generally necessary to admit the additional condition of absence of correlation between the casual variables. If the structure is uni-equational, with normal distribution of the casual variable, the structural coefficients as well as being identifiable will coincide with the coefficients of regression. Only the dependent variable will have a probabilistic distribution, while the other variables will be considered free from errors of measurement.

We can now ask ourselves why we make the effort to theorize the existence of a structure underlying the probabilistic distribution of the variables, bearing in mind that this necessarily brings us to the difficult problems of identification of the parameters. In other terms, could it not be enough, even for forecasting purposes, to know the probabilistic distribution of the variables? The answer would be in the affirmative in the case of validity of the protopostulate of invariance of reality; but the laws of economics change in space and in time as an effect of changes of reality, so that the variations of the structural parameters, whether natural or due to conscious economic policy interventions, make knowledge of the structure and its changes necessary in order to be able to make reliable forecasts.

The comprehension of the structure, which is obtained from the identification of the parameters, permits a deeper knowledge of the relationships between the variables and, therefore, of the determination of the new probabilistic distributions when changes have intervened. Therefore, whenever a structure remains unaltered, the identification of the parameters is not necessary, but when a change has occurred, the correctness of the forecasts requires the evaluation of at least some fundamental parameter. The conditions on the basis of which the causal ordering assumes operational meaning are therefore generally the same as those that give rise to the problem of the identification of the parameters. It should however be added that the conditions for complete identifiability of the parameters can be more restrictive than those required for the determination of the causal ordering. In some cases, in fact, a causal ordering can exist even if the structure is not completely identified; this because the causal ordering is concerned with the changes that intervene in specific sub-sets of the structure, and what is important is that the order between the various sub-sets does not change, even though admitting that a particular sub-set can

---

16 H. Simon, Causal ordering and Identifiability, op. cit., pp. 69-70.

17 See T. Koopivians, Identification Problems in Economic Model Construction, op cit., p. 36.

18 Ibidem, p. 36.

be substituted by an equivalent one. The complete identification of the parameters, on the other hand, requires that not even one sub-set of the system can be substituted by an equivalent sub-set.

Causal ordering and identification in the model of the Papi Commission

The concepts of causal ordering and of identification of the parameters find efficacious illustration in the study of the econometric model proposed by the Papi Commission[19]. This model is set forth as follows[20]

Inserire formula

(equation of employment)
(equation of total national consumption)
(equation of the gross accumulation of capital)
(equation of total imports)
(equation of net revenues from abroad)
(equation of definition of the national budget)
(equation of definition of the national budget)
(equation of definition of the gross accumulation of capital)
(equation of definition of the expansion of the gross national product)
(equation of definition of the gross national product)
(equation of equilibrium of the balance of payments)

With regard to the economic significance of the endogenous variables and the parameters contained in the pluri-equational system under examination, refer to the summary below; it should be noted that the variables referring to the time *o* should be understood as exogenous variables. The time *o*, in the application of the model, is made equal to the year 1960 (60); the time *t* to the year 1970 (70).

Inserire formula

Since the number of the equations (among themselves independent) is equal to that of the endogenous variables, it can be said that our model is determinate.

From analytical knowledge of the model it is easy to pass to the determination of its causal ordering. We were personally concerned with this problem, as members of the econometric group, which was part of the larger Work Group I of the Papi Commission. We therefore reproduce, developing it where opportune, our contribution appearing in the document that synthesizes the studies of this Work Group[21]. The graphic representation of the system shows the causal ordering of the econometric model and, therefore, the steps to be followed to determine the values of the endogenous variables.

Inserire Fig.

With reference to this, we observe that:

a) the equation (1) is a complete minimum sub-set of zero order with P_{70} (endogenous variable), which is determined as a function of P_{60} (exogenous variable);

b) the equation (2) is a complete minimum sub-set of first order with L_{70} (endogenous variable), which is determined as a function of P_{70} (now exogenous variable);

c) the equation (7) is a complete minimum subset of zero order with y'_{70} (endogenous variable), which is determined as a function of y'_{60} (exogenous variable);

d) the equations (4), (5), (9), (12), with the substitutions in (5) based on the equations of definition (10) and (11) and in (4) and (9) with the identity (13), form a complete minimum sub-set of first order in four endogenous variables (Y_{70}, I_{70}, C_{70}, y_{70}) and the exogenous variables (Y_{60}, I_{60}, and y'_{60} now exogenous variable).

The endogenous variables ate simultaneously determined;

---

19 See the document of Work Group I, the Papi Commission, Proiezioni Macroeconomiche al 1970, beginning October 1961, mimeographed.

20 Symbology and notes taken from the document cited.

21 See, Work Group I, Proiezioni Macroeconomiche al 1970, op. Cit.

e) the equation (6) constitutes a minimum complete sub-set of second order with M_{70} (endogenous variable), which is determined as a function of Y_{70} (now exogenous variable);

f) the equation (3) is a complete minimum sub-set of second order with Lo_{70} (endogenous variable), which is determined as a function of y_{70} (now exogenous variable);

g) the equation (8) is a complete minimum sub-set of third order with Ld_{70} (endogenous variable), which is determined as a function of Lo_{70} and L_{70} (now exogenous variables).

The model is therefore divisible into two fundamental sets. The first is constituted by the demographic equations connected according to relationships of causal asymmetry; the second consists of the interdependencies that connect the equation of total national consumption (4), the equation of gross accumulation of capital (5), the equation of definition of the national budget (9) and the equation of definition of gross national product (12). The second nucleus of the system, however, constitutes only a sectional structure that is integrated with the exogenous variables Y_{60}, I_{60} and y'_{70} obtained from the equation (7), and with the equations of definition of gross accumulation of capital (10), of expansion of the gross national product (11) and with the condition of equilibrium of the balance of payments (13). The last three equations allow convenient substitutions: for the simultaneous solution of the second nucleus of the system[22].

The second nucleus of the system, determining among the endogenous variables the gross national product, in its turn gives rise to a causal chain with the equation of employment (3), which connects this variable with the gross national product. This is the starting point for a successive link with the equation of definition of the labour force (8) that constitutes the point where this system is joined to the nucleus of demographic equations.

It was possible to determine the causal ordering that there had been prior ascertainment of the conditions that impeded consideration of equivalent structures obtained through linear combinations of the equations of the system. In other terms, the zero coefficients that could not be substituted by non-zero coefficients, whatever the politico-economic intervention on the structure, were fixed a priori.

We shall soon see the operational meaning of the causal ordering determined in this way. Proceeding strictly, however, we wish first to ascertain that the conditions of identifiability of the parameters of the structure have also been satisfied. Since these conditions are more restrictive than those required for the determination of the causal ordering, their examination will exempt us from showing the discussion of the conditions relative to the causal ordering.

In the case of a non-stochastic structure, Koopmans' necessary condition for the identifiability of an equation, that the number of the variable excluded from the equation to be identified should be at least equal to the number N of the structural equations minus one, also becomes a sufficient condition. We shall therefore confine ourselves to examining whether this condition is satisfied, remembering that in the number of variable s excluded are comprehended also the exogenous and predetermined variables.

With regard to type of structural equations, the system can be divided into two large categories; the first seven equations concern behaviour, the last six are equations of definition. For the latter, obviously, the problem of identifiability does not arise. We can, therefore, concentrate our attention on the first seven equations, remembering that the system of thirteen equations includes thirteen endogenous variables and four predetermined variables (P_0, I_0, Y_O and y'_0), that are a total of seventeen variables. Since not more than four variables figure in any of the first seven equations, the excluded variables in every equation are always at least thirteen. So, apart from any other consideration, Koopmans' condition is always satisfied and all the equations are therefore identifiable.

We must now examine what operational significance the determination of the causal ordering had in the econometric model of the Papi Commission. We can say immediately that it was very important. Work Group I in fact decided to assume alternative values for some parameters of the system, concentrating its attention, above all, on d (average propensity to consume) and on e (incremental capital-gross national product ratio). The purpose was to reveal what would have been the changes in the principal endogenous variables of the model, with particular reference to the gross national product and to unemployment. As alternative hypotheses the values 0.775; 0.785; 0.800 for the propensity to consume, and 5.2; 4.2; 3.8; for the capital-product ratio were assumed.

On the basis of these values the alternative rates of growth of the gross national income and the different levels of unemployment were determined. We shall not dwell on such calculations[23]. What it is important to underline in the ambit of the present research is that the fact of having assumed alternative values for the parameters was equivalent to postulating particular interventions in specific sub-sets of the structure; interventions that transform the structure but do not alter its causal ordering. They do not change the causal

---

[22] In fact one obtains the following reduced system of easy solution: Inserire formula

[23] With regards to these calculations and to pertinent comments and observations, see: Work Group I, Proiezioni Macroeconomiche al 1970, op. cit., pp. 31-37; F. di Fenizio, La programmazione economica in Italia, (1946-1962), UTET, Torino, 1965, pp. 211-214.

ordering because they do not substitute non-zero coefficients for zero coefficients. In other terms, to change d and e means to alter the equations (4) and (5), that is, intervene on the nucleus of interdependencies, modifying the values of the endogenous variables Y_{70}, I_{70}, C_{70} and y_{70} and, therefore, by means of the causal chain, influence the values of employment and the determination of the number of unemployed. The way the effects of alternative values of the parameters propagate themselves on the system of equations can easily be followed by means of the causal order graph that we have already reproduced. In conclusion, therefore, we can well affirm that "the possibility of instrumenting an econometric model, even a fairly aggregated zone, as has been so far discussed"[24] depends on the study of the parameters and on the investigation of the possible forms of intervention on the structure. All this is founded on the knowledge of the causal ordering the operational meaning of which we have therefore amply demonstrated.

Identification and causal dependence in linear structures

Having ascertained the identifiability of the equations of a structure, the next step consists in the estimation of the coefficients. The estimation of the coefficients is done by one of the many methods now commonly used. In stochastic structures, by applying the logic of likelihood, we can use the method of the maximum likelihood with complete information or with partial information. The first method also employs restrictions exceeding the minimum requisites required for identification. The second method can be applied (and is generally applied) to sub-sets of the structure, regardless of information (restrictions) relative to the equations that do not belong to the sub-set being studied. It can also be applied to the single equation and in that case is equivalent to the indirect method of the least squares, that is to say, the method of the least squares applied to the equation expressed in reduced form, from which it is possible, in certain conditions, (identifiability of the structure) to revert to the equation in its normal form.

The determination of the coefficients permits the singling out not only of the causal ordering but also of the causal dependence between the variables. If the structure is uni-equational, apart from the cases of high intercorrelation between the variables, which render indeterminate the process of valuation of the parameters because of multi-collinearity, the multiple regressions will provide the value of the single coefficients. Although, therefore, there are circumstances even for uni-equational structures where the determination of the coefficients appears insoluble, one cannot in these cases talk about non-identification. In fact, only through the linear combination with at least one other equation would it be possible to substitute non-zero coefficients for zero coefficients and vice versa. The uni-equational structure represents, therefore, a particular aspect of causal inference.

It is possible also to effect a comparison between multiple linear regression and causal analysis of a pluri-equational structure[25].

Wold[26] has dwelt at length on this problem In multiple linear regression it is possible to ascertain the effect of every independent variable on the dependent variable, the other variables being supposed constant. There does not, however, exist a network of relationships among the variables that shows certain relationships of causality or interdependence hypothesis. This is an assumption, which is resorted to frequently enough between them.

In the case of a pluri-equational structure, if it satisfies the conditions of identifiability, it is possible to determine the coefficients and the model that is obtained can be considered as an extension of the analysis of regression[27]. But the conditions of identifiability are often lacking and in t his case it is not possible to determine the coefficients of the system. The problem, which then arises, is how to measure the causal dependence between the variables.

The solution is indicated by analysis of the coefficients of dependence.

In order to determine these coefficients of dependence, which offer a measure of the weight of one variable on another, on the basis of a relationship of causality, it is necessary, however, to introduce a supplementary hypothesis. This is an assumption which is resorted to frequently enough and its acceptance, therefore, does not raise particular problems.

The linear relationships of the structure are integrated by causal variables that represent the effect of the latent variables. The validity of the conclusions depends, however, on the behaviour of the casual variables: it is supposed that each of them is specific to a X_i variable. In other words the factors of disturbance of X_i cannot be correlated with the factors of disturbance of X_j that is to say the co-variance of the casual variables must be null.

---

[24] See F. Di Fenizio, La programmazione economica in Italia, op. cit., p, 214.

[25] See V. Capecchi, "Causalità e correlazione nella problematica sociologica", Studi di Sociologia, II, 1964, p. 251.

[26] See H. O. Wold, "Causal Inference from Observational Data", Journal of the Royal Statistical Society, 1956, pp. 29-50.

[27] See R. Boudon, "Methodes d'analyse causale", op. cit., p. 26.

In non-experimental situations, as in the case of studies on cross-sections, this hypothesis is not so plausible as in the case of controlled experiments. In the first case it is therefore necessary to express the greatest possible number of variables, subtracting them from the group of latent variables[28].

Once the hypothesis of non-correlation and of specificity of the casual variables has been accepted, it is always possible to calculate some coefficients of the pluri-equational structure, which, although they do not coincide with the coefficients of regression, in that the structure is not identifiable, measure the causal dependence of the variables[29]. These coefficients are the coefficients of dependence. They can be determined in the following way. If the generic equation of the structure is $x_i + a_{i1}\, x_i + a_{i2}\, x_2 + \ldots a_{ni}\, x_n\, u_i$ (the variables measure deviations from average values). Multiplying both members by x_i and introducing the operator E (mean) one obtains

Inserire formula

The terms preceded by the operator E represent the covariances of the variables in brackets. Given the hypothesis of the specificity of the: casual variables, it will be $E\,(u_i\, x_j) = 0$. From which

Inserire formula

Likewise for all the other equations of the structure. Thus is easily obtained a system of equations, which expresses the coefficients of dependence a_{ij} as a function of the variances $E\,(x_j\, x_j)$ and of the covariances $E\,(x_i\, x_j)$ of the express variables[30]. It is now a matter of analysing the statistical significance of these coefficients. An interpretation has been given by the biologist Wright, outside the context of the problem of pluri-equational structures. We refer to synthesis of Boudon's, expounded in the study we have already quoted several times. Let there be any linear equation with more than one variable, with x_j as endogenous variable

$$x_i = a_{i1}\, x_i + a_{i2}\, x_2 \ldots + a_{ij}\, x_j\, u_i$$

Let one determine the influence of the variance of a variable (for example x_1) on the variance of the endogenous variable x_i, the values of the other variables being supposed unaltered. With the symbology of the linked variances one will have:

Inserire formula

Using d_{i1} to indicate Wright's coefficient and bearing in mind the preceding expression, we will have:

Inserire formule

If x_1 does not depend on other variables, its variance is equal to its linked variance. It follows that d_{i1}^2 **Inserire formula**. Wright's efficient expressed in this way measures the part of the variance of the variable x_i due to x_1. If x_1 is not independent, it is necessary to correct the coefficient with the second factor, which takes into account the fact that the variance of x_1 can increase as a result of the variations of the other variables.

Therefore the coefficients of dependence a_{ij} corrected with the ratio s_j/s_i constitute a type of measure of the causal influence of the independent variable on the dependent variable. Their usefulness derives from the fact that they are always determinable, even in the absence of the conditions of identifability of the structure to be analysed. The corrected coefficients of dependence can therefore be considered an extension of the coefficients of correlation. In fact, if we consider the case in which x_1 does not depend on other variables, for which reason its variance is equal to its linked variance, Wright's coefficient appears as a normal coefficient of correlation.

From this consideration also emerges the possibility of approaching the coefficients of dependence to the coefficients of regression. It is a matter of demonstrating that in an identifiable structure the coefficients of dependence coincide with the coefficients of regression. It is enough to demonstrate that the equations from

---

28 See H. M. Blalock, "Four Variable Causal Model and Partial Correlations", op. cit., p. 183.

29 See R. Boudon, "Methodes d'analyse causale", op. cit., p. 31. For the demonstration of the possibility of determining the coefficients of dependence in non-identified structures, the hypothesis of the specificity of the causal variables being valid, pages 33-35 of the same work. See R. Boudon, "Methodes d'analyse causale", op. cit., p. 31. For the demonstration of the possibility of determining the coefficients of dependence in non-identified structures, the hypothesis of the specificity of the causal variables being valid, pages 33-35 of the same work.

30 See R. Boudon, "Methodes d'analyse causale", op. cit., pp. 34-35.

which the coefficients of dependence are obtained are the same ones that serve to determine the coefficients of regression. We already know the equations from which the coefficients of dependence are obtained. We shall therefore concentrate our attention on the equations from which the coefficients of regression are obtained[31]. These coefficients are determined by the method of least squares. The method of least squares applied to a single equation consists, as is known, in minimizing the sum of the squares of the deviations from the interpolated values. This is the equivalent of minimizing the following expression, where the variables are expressed as deviations from the mean:

Inserire formula

For this expression to be a minimurn, it is necessary that the partial derivatives with regard to the parameters are null. In particular it must be $dz/d_{ij} = 0$, that is:

Inserire formula

that can be written

Inserire formula

or also

Inserire formula

But this last expression is identical with the generic expression from which the coefficients of dependence are obtained, whence it is demonstrated that in an identified structure the coefficients of regression and the coefficients of dependence coincide.

In conclusion, therefore, the coefficients of dependence can be understood as a generalization of the coefficients of regression. These latter are valid only for identified structures; the former show their fertility, also in the case of non-identifiable linear structures, provided that the casual variables satisfy the condition of specificity.

Likewise, the coefficients of dependence corrected by the ratio between the standard deviation of the explanatory variable and the standard deviation of the explained variable can be understood as a generalization of the coefficients of correlation that show their limited validity with regard to identified structures.

Conclusions

From the preceding theoretical synthesis and the analysis of the experience effected in the ambit of the Papi Commission it is possible to draw some fundamental conclusions, which connect up with the introductive discourse.

The development of econometric research has underlined the operational fertility of a definition of structure as a system of consistent economic relationships. The classification of the types of structure (self-contained, sectional) with the corresponding concepts of variable (endogenous, exogenous) and the analysis of the sub-sets constituting the system of equations, have permitted us to single out asymmetric relationships between the variables, to which, in accordance with Simon's definition, we have made the concept of causal ordering correspond. This concept is bound to the possibility of interventions on specific sub-sets of the structure, which produce alterations in it without producing changes in the asymmetry and in the order of the relationships between the variables. In other words, the zero coefficients of the system are fixed a priori, so that any change in the value of the other parameters changes the solution-values of the system but does not alter the line of propagation of the effects of these changes through the structure. The concept of causal ordering, with permits consideration of the effects on the variables of certain interventions on the parameters of the system, has therefore operational significance in that these interventions can be associated with decisions of the experimenter (in our case of the political economist).

If the zero coefficients of the system are not fixed a priori, there would always be the possibility of constructing an equivalent structure by means of linear combinations of the equations, in such a way as to satisfy the same group of empiric observations. In this case the concept of causal ordering would become evanescent, at least from the empiric point of view.

---

31 A strict treatment of the analysis of regressions applied to the general linear model is to be found in J. Johnston, Econometric Methods, McGraw Hill, 1963, pp. 106-115. We shall limit ourselves to the consideration of a generic equation of the structure and the application to it of the method of the least squares. See for instance p. 56 of the same volume.

It has been seen that the considerations that permit determination of the causal ordering are the same on which is based the possibility of identifying the parameters of the structure. And even more restrictive conditions than those required for the causal ordering are sometimes necessary for identification. In the absence of identification, the evaluation of the coefficients of a structure becomes indeterminate.

The fruitfulness of the notions of causal ordering and identifiability of the parameters was empirically verified in the course of the analysis of the econometric model proposed by the Papi Commission. It was concluded then that "the possibility of instrumenting an econometric model, even a fairly aggregated one" depends on the study of the parameters and on the investigation of the possible forms of intervention on the structure. All this is founded on the knowledge of the causal ordering, the operational meaning of which therefore appeared to be amply demonstrated.

An identified structure permits the estimation of the coefficients, the knowledge of which indicates not only the causal ordering but also the weight that every variable exerts on the others. Therefore, the study of the coefficients connects the problem of the causal ordering to that of causal inference. In this regard we remember that in a uni-equational structure the coefficients are determined through the analysis of regression. It is possible to attribute an analogous significance to the coefficients of a completely identified structure. But in a structure in which it is not possible to satisfy the conditions of identification, is it possible to measure the causal dependence between the variables, since the coefficients of regression are indeterminate? The reply is in the affirmative, and the analysis of causal dependence provides the road to the solution of this problem. It is demonstrated that, provided that the casual variables of the structure satisfy the conditions of specificity and of non-correlation between them, it is always possible to calculate the coefficients of dependence in stochastic linear structures, even if they are not identifiable. The statistical significance of these coefficients is revealed in an extension of the analysis of regression: in identified structures coefficients of dependence and coefficients of regression coincide. The correction of these coefficients of dependence with the ratio between the standard deviation of the explanatory variable and the standard deviation of the explained variable, gives rise to Wright's coefficients. In the same way that the coefficients of dependence can be interpreted as an extension of the coefficients of regression, Wright's coefficients must be understood as an extension of the coefficients of correlation.

We can therefore summarize the results of the study by pointing out the close connections existing between the causal ordering and causal inference. 'The associations between causal ordering and identification on the one hand, and the extension of the analysis of regression into the most comprehensive analysis of causal dependence, on the other, have contributed to making these concepts increasingly operational and have permitted their satisfactory inclusion in an organic theoretical framework.

The results obtained should not, however, give an excessive confidence in the instruments available. In this sense, we should like to conclude by harking back to a sentence in one of our previous essays:[32] the improvement of econometrics often lies, more than in the use of the most refined technical instruments, in a more accurate collection of data, under the control of a strict scientific methodology. The depuration of observations should demand all the economist's attention.

---

32 M. ARCELLI, "Crisi dell'econometrica o errata impostazione metodologica", *L'industria*, 1961, n. 1, p. 10.

Federico Arcelli **Giorgio Di Giorgio**

QUALITY CHANGES IN FACTORS AND DECOMPOSITION OF TECHNICAL PROGRESS IN AGGREGATE PRODUCTION FUNCTIONS [33]

MANCANO LE NOTE DAL TESTO ORIGINALE – INSERIRE!!

Preamble

The object of this study is to analyse the components of technical progress in an aggregate production function, considering first of all the effects of changes in the quality of factors. This will make it possible to place in evidence the links existing between a series of preponderantly qualitative variables and the rate of change in productivity, while at the same time providing a clearer idea of some economic development mechanisms. A critical assessment of the empirical repercussions of the acceptance of certain hypotheses to interpret reality will be able to teach us something about the solution of economic policy problems. The existence of a number of criticisms disputing the very foundations of aggregate production functions makes it necessary to consider the effects of changes in the quality of factors and to break down technical progress into its most significant components.

As will be demonstrated, failure to consider the effects of changes in the quality of factors already produces biased relations between variables at the level of the classical Cobb-Douglas equation. The partial elasticities of output with respect to inputs, and returns to scale are either underestimated or overestimated.

When one passes from the static production function to a dynamic function of the kind proposed by Solow[1], the bias in the elasticities of factor inputs remains and, if the sum of these elasticities is posited a priori as equal to the unit, the returns to scale and the inherent deviations merge in the trend term that should measure the influence of technical progress. But it is for this very reason that Solow's function is of but slight heuristic significance. If, according to the model of construction, all unexplained deviations are embodied in the trend term that should express technical progress, the result must necessarily reveal a stable relationship between the output per man/ hour and the per capita capital once the function has been freed from technical progress. In other words, the function appears to be tautological once a constant relationship has been established between the per capita product and capital and the residue is confined to the term trend. Although this criticism[2], in its strictest formulation, does not do full justice to Solow's[3] empirical analysis, it has the merit of stressing the scant significance of the estimate of technical progress term, since it embodies the effects of any qualitative and quantitative variable with the exception of the relationship between product and per capita capital, the meaning of which, however, appears most uncertain. In fact, not only do the variables appear to be insufficiently defined, but, as has been seen, the relationship between the variables is also biased in the specification of the elasticities whose sum is posited a *priori* as equal to the unit.

Moreover, the failure to consider differences in the quality of factors and to break the trend term up into its most significant variables renders the stability of the relationship precarious owing to the problem of aggregation.

It is now widely accepted, in fact, that the stability of the aggregate production functions, since they are constructions based on an essentially microeconomic theory, depends on the invariance of the underlying, microeconomic relations and on the absence of changes in the relative importance of the sectors forming the aggregate. It depends, therefore, on the constancy of the structural relations in the existing sectors and on no new sectors of production being added to the existing economic activities. The greater the difference in the quality of the factors, in relation to coordinates of space or time (better vocational training in the case of labour, more advanced technology in the case of capital), or merely owing to a problem of shortage (first the more productive factors are used and then the less efficient ones), the more difficult it will be to fix a stable macroeconomic relationship.

Not only can a trend term that embodies the effects of a set of unspecified variables, qualitative variables included, not be considered stable; it does not even provide an instrument enabling us to understand how the process of development takes place. Its heuristic fecundity is consequently practically null and its efficacy as an operative instrument extremely limited.

More recent economic analysis has therefore devoted attention to overcoming the foregoing criticisms by adopting different research approaches.

Firstly, the production functions have been disaggregated so as to allow progress – and the consequent increment of productivity – due to the redistribution of factors among the different branches of industry to be

---

[33] From: Rivista di Politica Economica, Selected Papers, No. 1 – 1967.

distinguished from progress due to an improvement in efficiency inside each industry. The same has been done with regard to area redistribution problems.

Secondly, the criticism that differences in the quality of capital deriving from technological improvements embodied in the newest capital goods are neglected is overcome by constructing the so called «vintage» models, in which the hypothesis of the homogeneity of capital is discarded and the link between technical progress and investment stressed.

Thirdly, a re-assessment is made of the qualitative factors connected with the standard of training and the improvement in the efficiency of the labour factor. As a matter of fact, whereas in former analyses the importance of investments in capital goods was insisted on as essential to the achievement of a high rate of technical progress, recent currents of thought are inclined to re- assess the importance of the factor «man» and of investments in human capital.

The more recent evolution of contemporary economic thought concerning the origin of technical progress may here be summed up by emphasizing the passage from an exogenous conception to an at least partly endogenous view In the first models technical progress is generally of the neutral kind, not embodied in capital goods. It develops autonomously through time. Later, however, there is a tendency to stress the close connection existing between investments in new capital goods embodying the latest technologies and the increase in the efficiency of the production process. The classical production function is replaced by vintage models or technical progress functions that are a particular configuration thereof. In the latter case, the distinction between factor substitution and the shift of a production function is discarded as insufficiently realistic.

The lesson to be learnt from vintage models is that investment in new capital goods is a first-rate strategic variable for obtaining a rapid increase in the rate of growth of output and productivity and that the tempo of technological innovations may be suitably stepped up by an appropriate scientific research policy.

The discussion of a recent research that tends to redimension this assertion – at least as far as the technologically more advanced countries are concerned – will provide us with an opportunity for a refinement of the analysis that will, I think, make a useful contribution to the vintage model theory.

In this connection I shall have occasion to insist on the fact that the hypothesis of the embodiment of technical progress in capital goods has not only a heuristic meaning. Even if, when it comes to testing the models empirically, the difficulty of distinguishing embodied from disembodied technical progress may cause some epistemological scepticism.

I shall not, however, deny the importance of a better knowledge of the role played by improvements in the factor «man» and by the redistribution of factors in the achievement of a higher rate of productivity.

Models have recently been built that explicitly introduce a variable expressing changes in the quality of labour. And in recent studies attention has been centred on the effects of research and education on productivity, including these variables in the C.E.S. production function.

The controversy (embodied or disembodied technical progress) – less determinant than the suggestive theoretical presuppositions might lead one to anticipate – is thus tempered by the consideration of other strategical factors as regard the increment in productivity.

In this article my intention is to illustrate some of the themes briefly enunciated in the preamble, pointing out new trends of analysis and empirical research and offering food for meditation regarding some of the results of a survey recently carried out by the Italian Institute of Statistics.

Effects of the Exclusion of Quality Changes in Factors from a Static Aggregate Production Function

As I said in the preamble, failure to take into account differences in the quality of factors in a production function of the Cobb-Douglas type produces distortions in the relations between the variables, which are reflected by a bias in the calculation of the elasticities of the output with respect to factors. We generally estimate statistically the parameters of the P = **Inserire formula** function, whereas the true function to be estimated should be the following: $P = (q\ L)^{l}\ (t\ K)^{m}$, in which «*q*» and «*t*» represent the weights that serve to transform heterogeneous inputs of labour and capital respectively into equivalent quantities. «*q*» is a variable that renders the quantities of labour homogeneous, with different degrees of vocational training. «*t*», is a variable that renders homogeneous the capital goods of different ages embodying different technologies.

By estimating P = **Inserire formula**, we exclude two of the variables (*q* and *t*) of the true relation. We ask ourselves whether it is possible to express the parameters **Inserire lettera** and **Inserire lettera** in terms of l and m in order to establish the direction of the bias.

Going on to the logarithms, the production function becomes a linear relation whose parameters can be estimated with the method of least squares. In this case, «if we know what the true equation should be like, but we estimate something different from it, it is possible to express the resulting coefficients in terms of the coefficients of the true equation. Each of the estimated coefficients will be a weighted sum of all the coefficients in the true equation. The weights will be the coefficients in the auxiliary regressions of each of the true variables on all of the variables included in the estimated equation»[5].

The statistical estimate theory also teaches us that if we exclude from a relation a relevant variable not correlated with the other independent variables, the regression coefficients of the variables included will not undergo any change. But if a correlation exists between the variable excluded and the other independent variables, the estimates of the parameters of these variables will be biased. In our case, since the correlation exists, **Inserire lettera** and **Inserire lettera** will be affected by bias.

Let us suppose for a moment that capital is homogeneous. Whereas labour presents differences in quality, we wish to determine the effect on the parameters of failure to consider the changes in the quality of labour. Therefore let $P = K^m (q\ L)^l$ be the true function, while the estimate is made on the function P = **Inserire lettere**. Ws are therefore excluding variable *q* with coefficient l. The estimate of parameters **Inserire lettere** and **Inserire lettere**, expressed in terms of the coefficients of the true equation, will therefore be respectively:

Inserire Formula

in which zm and zl are the regression coefficients of the auxiliary equation $q = K^{zm} L^{zl6}$.

The elasticity of inputs will therefore be upwards or downwards biased according to the signs of zm and zl: zm expresses the correlation between the quality of labour and the capital invested. Everything allows us to assume that this correlation is positive. On the other hand, zm expresses the correlation between the quality and the quantity of labour and in this case quality may legitimately be expected to substitute for the quantity of labour. This would therefore be a negative correlation. In conclusion, therefore, if differences in the quality of labour are neglected, the estimate of the contribution made by capital will be overestimated and that of the contribution made by labour will be underestimated.

The same procedure may be followed to determine the effects of failure to consider differences in the quality of capital when labour is assumed to be homogeneous.

In this case $P = (t\ K)^m$ L **Inserire lettera** is the true function, while P = K **Inserire lettere** is the fuction estimated.

The estimate of **Inserire lettera** and **Inserire lettera** will then be

Inserire formula

where *s*l and *s*m are the regression coefficients of the auxiliary equation $t = K^{sm} L^{sl}$

The sign of *s*l and *s*m will therefore determine the direction of the bias of the elasticity of the product with respect to the factor.

*s*m expresses the correlation between the quality and quantity of capital invested, which we need not hesitate to assume to be positive. The estimate of the elasticity of the product with respect to the quantity of capital will therefore be overvalued.

*s*l expresses the correlation between the quality of capital and the quantity of labour. A more modern technological process generally requires less manpower so that we may expect a negative relation. The elasticity of the product with respect to labour will therefore be underestimated.

Failure to consider qualitative improvements in labour and capital simultaneously will strengthen the overestimation of the contribution made by capital and the underestimation of the contribution made by labour.

However the discussion cannot be considered ended unless we consider the effects on returns to scale. If the sum of the coefficients of the estimated function is larger than in the case of the true function, we shall say that returns to scale are overestimated, or conversely that they are underestimated.

If the qualitative difference in labour is neglected, the bias of returns to scale will be equal to l (zm + zl–1), equivalent to the difference between the sum of the estimated coefficients and the sum of the coefficients of the true function. Since zm + zl is manifestly less than one, returns to scale will be underestimated. If, on the other hand, it is the qualitative differences in capital goods that is neglected, the bias in returns to scale will be equal to m (*s*l + *s*m – – 1), equivalent to the difference between the sum of the coefficients of the estimated function and the sum of the coefficients of the true function. If coefficient *s*m is particularly high, meaning that technological innovations proceed with very marked changes in quality, the sum of *s*m and *s*l might even exceed the unit and in this case returns to scale would be overestimated. However it is not possible to fix a law that is universally valid.

Leaving aside multicollinearity problems, which often render the estimate of parameters indeterminate[7], and aggregation problems[8], which introduce errors of aggregation and instability in aggregate production functions, we have thus seen how errors of specification, which consist of neglecting qualitative variations in factors produce bias in the estimation of elasticities and consequently distort the evaluation of returns to scale.

From a heuristic point of view, moreover, the problem of the causes determining the rate of growth of the product eludes us so that it is not possible to assess the operational consequences of alternative economic policies that insist either on one factor of development or on another.

Passing from a static to a dynamic function and setting the restriction that the sum of the elasticities of the product with respect to factors shall be equal to the unit, returns to scale and their over-or under-estimations due to failure to consider qualitative differences in factors merge in the trend term which – designed to express the effect of technical progress – in reality conglobates the total increase in the productivity of factors due to this cause.

In a dynamic function of the kind proposed by Solow[9] technical progress, as is known, is identified in an autonomous factor that develops neutrally through time as a result of multiplicative shifts of the isoquant of production. The foregoing considerations, however, lead one to inquire into the components of this factor for the purpose of isolating what is truly autonomous (organizational progress for instance) from what depends on other causes. In other words, it must be placed in evidence that the growth of productivity also depends on the technological progress embodied in the most recent capital goods, that is to say on improvements in the quality of capital; on the greater efficiency of labour as a result of improved skills and, lastly, on the redistribution of production factors among sectors of industry. This is what leads us to propend for analysis with a disaggregation of the whole production function and not only of the term representing technical progress.

It is my intention, in this paragraph, to deal with the first of these problems, the one referring to the relations between technical progress and investments on the one hand and changes in productivity on the other. This will enable us to understand the meaning of vintage models without going into them particularly thoroughly.

Let us consider an economy in expansion: the capital does not consist of a stock of homogeneous goods; there will be different kinds of capital goods with which different degrees of technological perfection can be associated[10]. Given a certain level of technical and scientific knowledge, we may consider the best technology to be embodied in the capital goods most recently introduced. If we suppose a complementary relation to exist between new technologies and new investments, that is to say between technical progress and capital accumulation, we may then affirm that a direct relation exists between the increase in productivity and the rate of growth of investments. One might even say, as a corollary, that a reduction in the average age of the capital stock tends to accelerate the rate of growth of income.

Let us assume $H(t)$, the best technology, to be an exponential function of time. Considered as a function of time, $H(t)$ stands for the flow of inventions in a society and therefore points out the ideal technological path that would be followed if the new technologies were adopted instantaneously – the state that would prevail if all existing capital embodied the latest technological developments. A continuous function being assumed, the capital stock's average age would be equal to an infinitesimal quantity and would therefore coincide with its marginal age.

If A (t), the cumulate rate of Solow's technical progress function includes the technical progress embodied in the capital goods, it is possible to give an interpretation thereof in terms of $H(t)$.

$A(t)$ stands for the technological level of the economy as a whole and therefore takes into account the composition of the capital stock by age, representing in substance a weighted average of H (t).

The weights of this average are the investments in the different periods. If, therefore, $H(t)$ is the function of inventions, $A(t)$ is the function of innovations, that is to say of those technological improvements actually applied to production processes.

Following the course outlined in a study by Massell[11], we can then write $A(t) = f(j)\,H(t)$ where $f(j)$ takes on values between zero and one and is therefore an indication of the age of the capital stock, while j represents a set of parameters (e.g. exponential rate of growth of gross investments, rate of growth of the labour force, etc.) characterizing the growth of production in a certain economic system, with the introduction of some suitable simplifying hypotheses. If these parameters remain constant through time, we shall say that the economy develops along a path of equilibrium. At each change in the parameters, there will be a passage from one path of equilibrium to another, with an asymptotic movement. Applying the logarithms to the expression $A(t) = f(j)\,H(t)$ and taking the derivative with respect to time, assuming $f(j)$ to be constant, we shall obtain the result that $\mathbf{A}/A$, the rate of change with respect to the innovation function, is equal to $\mathbf{H}/H$, the rate of change with respect to the invention function. Therefore, if the economy proceeds along a path of equilibrium, the rate of equilibrium of innovation is equal to the rate of equilibrium of invention and independent of the rate of investment.

But if investments have no effect on the gradient of the technological path, they do have one on the absolute level of technology $A(t)$. The rate of growth of investments, included in $f(j)$, contributes in fact to determine the intercept of the path of equilibrium, the expression of which in logarithmic scale is:

$$\log A(t) = \log f(j) + \log H(t)$$

It must be borne in mind that since $0 < f(j) < 1$, $\log f(j) < 0$ and therefore, as is natural, the path of invention lies above the path of innovation.

In conditions of equilibrium, the two paths proceed parallelly, however, and the gradient, of the technological path is equal to the rate of change in the flow of inventions. Since the flow of inventions may be held to be influenced by the sums spent by the collectivity on scientific research, it ensues that the technological path may

also be modified. Here we have a theme I shall have occasion to develop farther on: the relation between scientific research and the increase in productivity.

The degree of an economy's technological development may therefore be explained partly in terms of decisional variables: this is an approach departing from the traditional formula that considers technical progress a purely exogenous variable.

While, in conditions of long-run equilibrium, the flow of inventions conditions the flow of innovations; in the short run, a change in the rate of investments, that is to say a change in f (*j*) influences the rate of growth of productivity. This is understandable if one bears in mind that the average age of the capital stock is thus decreased and – since the latest capital goods embody the most advanced technologies – there will be an increase in productivity.

If, therefore f (*j*) remains constant through time, that is to say if the average age of the capital stock does not change, the recognition of the interdependence of investments and technical progress does not call for a qualification of the results obtained with Solow's equation $P = A(t) L^{l} K^{1-l}$. It will only be necessary to consider the particular meaning of *A* (*t*) resulting from the interpretation given in this paragraph.

In other words, *A* (*t*) does not only include organizational technical progress, but also the progress embodied in the new capital goods.

However this does not appear to distort the results of the analysis. We shall find confirmation of this when analysing a vintage model proposed by Solow.

But if the composition of the capital stock by age varies through time, then the estimation of **A**/*A* with the equation $P = A(t) K^{1-l} L^{l}$ is no longer correct. In particular, it is possible to show that if $d f / d t > 0$, the importance of the changes in technical progress is generally overestimated to the detriment of investment. Conversely in the opposite case. In these cases, the change in the quality of capital makes it necessary for analysis to pass to vintage models.

Embodied Technical Progress and Vintage Models.

When changes in the quality of capital are taken into account, the model $P = A L^{l} K^{1-l}$ becomes $P = A' L^{l} t^{1-l}$, where t is a weighted sum of the capital goods, with weights varying in relation to their age. The weights are higher for the latest goods since these are assumed to embody the most modern technologies. Model $P = A' L^{l} t^{1-l}$ is the classical vintage model[12] proposed by Solow. It differs from the preceding model not only because t has replaced *K*, but also because *A'* differs from *A*. In fact *A* includes both *A'* and the technical progress embodied in t. Moreover the variables P and L must be understood to be integral values of *P* and of *L*, calculated for the different vintages, that is to say for investments in different periods.

Passing to the logarithms and deriving, we obtain the relative rate of growth of the product:

Inserire formula

in function of the different components. Passing from the continuous to finite differences, the expression becomes:

Inserire formula

Let us suppose the rate of technological change in capital goods to be quantifiable in an «r» percent a year. t (*t*), expressed in terms of *K*, will then be equal to:

Inserire formula

where K_{vt} is the investment in year «v» still active at time «t». Formula 2 may, with acceptable approximation, be translated into the following:

Inserire formula

where ã0 and ãt are the average ages of the capital at time «0» and at time «*t*» and

Inserire formula

For low *r* values and for at not much removed from ã0, we then get with successive approximation[13]:

Inserire formula

This expression allows an interesting interpretation. The first term of the right-hand member of equation [4] expresses the increase in the capital stock; the second term the weight of the change in quality, that is to say the technological change embodied in the investment, the average age of the capital stock being assumed to be unvaried; the third term express the influence of the change in the average age of the capital.

By substituting [4] in [1] we get:

Inserire formula

If the average age of the capital stock does not change, the last term of [5] is annulled and we then discover that **Inserire formula**. So that, in conformity with the conclusion reached at the end of the preceding paragraph, we are provided with evidence that the interdependence of investments and technical progress does not in this case call for a qualification of the results obtained with the equation $P = A(t) L^{l} K^{1-l}$,

Analysis with the vintage model has therefore enabled us to distinguish that part of total technical progress ascribable to the superior technologies embodied in the new capital goods. If, however, the average age of the capital stock changes too, then it becomes imperative to take the differences in the quality of capital into account to avoid distortions in the results of the analysis. Failure to consider the qualitative differences in factors, which already implies distortions in the calculation of elasticities in a static function, affects the technical progress term also in a dynamic function. The fact of having considered that, in an economy in expansion, capital does not take the form of a stock of homogeneous goods since the more recent investments embody more advanced technologies, allows us to distinguish realistically an important component of the global rate of growth of factor productivity. What we now have to do is to evaluate the weight of this at the empirical level.

On the Need to Refine Analysis to Evaluate the Empirical Importance of Technical Progress Embodied in New Capital Goods.

The conclusions reached in the preceding paragraph provide valuable guidance for the critical assessment of some statements concerning the greater or lesser empirical importance of the technical progress embodied in new capital goods. However the analysis has to be further refined, clearing the field of possible commixtures with other problems. In order to avoid any possible confusion, it will be well to state in advance that this question is not incompatible with (although different from) that of ascertaining the statistical significance, as hypothesis for the interpretation of reality, of a function of the $P = A(t) L^{l}$ Km kind, in which technical progress is neutral.

As has been seen, when the average age of the capital stock does not change, the complementary nature of technical progress and investment does not call for a qualification of the results of the analysis made with a function of the $P = A(t) L^{l} K^{1-l}$ kind. The importance of $A(t)$ is not distorted with respect to that of capital investments.

It will simply be **Inserire formula**

This last equality has, however, the advantage of placing in evidence the component $(1–l)\,r$, which expresses the progress embodied in the new capital goods. The equation $P = A(t) L^{l} K^{1-l}$, in which technical progress is neutral, is not therefore incompatible with the hypothesis that part of this progress is embodied in the new capital goods.

In the presence of neutral technical progress, therefore, if the average age of the capital stock does not change, the decomposition of technical progress into disembodied and embodied has only a purely heuristic meaning.

However this does not exclude the possibility of technical progress embodied in capital goods assuming conspicuous dimensions.

Had we adopted a C.E.S. production function, the conclusions reached concerning the uselessness of the decomposition of technical progress when the average age of the capital stock does not change could not have been maintained.

In terms of rates of growth of the product, of technical progress and of the factors a C.E.S. function may in fact be written:

Inserire formula

If K is now substituted by t [1] becomes

Inserire formula

and substituting to D t/t the approximate formula

Federico Arcelli **Giorgio Di Giorgio**

Inserire formula

In the case of r = 0, [3] becomes equal to [5] in the preceding paragraph. But if r ≠ 0, it is no longer possible to write **Inserire formula** *r*, when the last term of [3] is annulled by the constancy of the average age of the capital stock. In fact **Inserire formula** differs from **Inserire formula** because t > *K*.

If r > 0, then D *A* /*A* will be underestimated if the rate of increase in labour is lower than the rate of increase in capital.

Conversely D *A* /*A* will be overestimated if there is a higher rate of increase for labour than for capital.

If r < 0, D *A* /*A* will be overestimated if the rate of growth of capital exceeds the rate of growth of labour, conversely in the opposite case.

Since e expresses the elasticity of factor substitution[14], (r > 0 means s < 1, and r < O means s > 1), if r differs from zero – *i.e.* if the elasticity of substitution is other than one – it will no longer be possible to accept the equivalence of the analysis made respectively with a production function and with a corresponding vintage model, if the condition of the constancy of the average age of the capital stock is maintained. And that is not all: the contribution of embodied technical progress to the rate of growth of the product will now depend also on changes in the relative share of capital used in production.

The case of the complementary nature of factors, *i.e.* when s ∅ 0 and r ∅ ∞, therefore comes within the more general case of r ≠ 0: consequently if, in this event, the capital goods are heterogeneous, the analysis made with a production function in which capital is considered homogeneous will produce a bias in the evaluation of the technical progress magnitude. The results reached in cases where the elasticity of factor substitution is other than the unit are also illustrative in respect of cases of non-neutral technical progress.

From the preceding analysis we know that when s ≠ 1, the elasticities of the product with respect to factors are not fixed as in the Cobb-Douglas case, but depend on the ratio of factors employed in the production process. Now the characteristic of non-neutral technical progress is precisely that it alters the elasticities of the product with respect to factors.

The above discussion therefore provides evidence that, also in the case of non-neutral technical progress, the constancy of the average age of the capital stock is not a sufficient condition to allow the decomposition of technical progress into its components to be neglected.

At the end of this paragraph we may therefore reach the important conclusion that it is only in the particular case of neutral technical progress and of unit elasticity of factor substitution that the interdependence recognized between technical progress and investment makes it unnecessary to qualify the results obtained with a production function; supposing also the average age of the capital stock to be constant.

In all other cases; *a*) change in the average age of the capital stock; *b*) elasticity of factor substitution other than one; *c*) non neutral technical progress, the analysis made with a production function in which capital is assumed to be homogeneous leads to an over- or under- estimation of the technical progress term. The acceptance of the embodiment hypothesis makes it necessary, then, to break the trend term down into its components, not only for heuristic reasons but above all to avoid distortions in the analysis.

Is the Hypothesis of the Embodiment of Technical Progress in Investment Really Unimportant?

We are now better able to judge some of Denison's statements[15] concerning the unimportance of the hypothesis of the embodiment of technical progress in capital goods. According to Denison, the hypothesis of the embodiment of technical progress shows its effects through changes in the average age of the capital stock.

Supposing the average age of the capital stock can be decreased by one year in the period between 1960 and 1970, if technical progress is embodied in investments this means that in 1970 we shall operate with the technique we should have applied in 1971 if the average age of capital had not changed.

Supposing the development of technology produces an increase of 1.6% a year in productivity, the 1970 product will be 1.6% higher than it would have been if the average age of the capital stock had not changed. Translated into annual terms, the yearly rate of increase would differ by 0.16% (assuming a simple rate of interest). And that is not all: for the whole of the period in which the new average age was maintained, the levels of the national product would differ constantly by 1.6% as compared with the hypothesis denying the embodiment of technical progress.

The demonstrative strength of the calculation is increased, according to Denison, by the fact that the average age of the capital stock is not very variable so that a change of one year must already be considered important, particularly in the case of plant and machinery whose active life is relatively short. The calculation made by the U.S. Commerce Department assigns an average age of 7.2 years to machinery in 1961 and a range of variation of 1.5 years from 1929 on. An alternative calculation fixed the average age of machinery at 5.6 years and the range of variation at 1.2 years. If, in addition to machinery, we consider plant too, the extreme limit of variability rises respectively to 2.6 and 3.6 years according to alternative calculations. Altogether, however, the change in the

average age between 1929 and 1961 was only 0.3 years in the first case and 1 year in the second case. These figures would suffice to demonstrate the insensitiveness of the variation in the average age of capital to changes in the rate of investment. Denison recognizes, however, that in economic systems less technologically advanced than the U.S.A., in countries where the gap between the best and worst technologies used simultaneously may be very wide, the change in the average age of the capital stock is more sensitive to changes in the rate of investment. But even after making this concession, Denison insists on the unimportance of the embodiment hypothesis.

The considerations made in the preceding paragraph allow us, however, not to share Denison's conclusions. As has been pointed out, it is only in particular cases that the interdependence recognized between technical progress and investments does not call for a qualification of the results obtained with a function of the P = A (*t*) *f* (*L*, *K*) kind, when the average age of the capital stock does not change. But also in these cases, **Inserire formula** the rate of relative change in the innovation function is linked with the rate of relative change in the invention function. So that, even if the average age of the capital stock remains constant, A (*t*) will increase at a brisker rate if the flow of innovations (linked with that of inventions) in the economy increases. Consequently the embodiment hypothesis has the effect not only of accelerating the rate of development of the product when the average age of the capital stock diminishes, but also of bringing about changes in the rate of innovation. If the existing rate of innovation were to cease – in which case the embodiment hypothesis would be excluded – D *A/A* would become equal to **Inserire formula** and the rate of growth of technical progress would therefore diminish at the rate of growth of the disembodied component.

Consequently it is not correct to say that it is only changes in the average age of the capital stock that make the embodiment hypothesis differ from the hypothesis denying embodiment.

In some periods and for some economic systems the accent may fall more on this aspect of the question, whereas in other periods and for other economic systems it will be the flow of innovations that attracts the attention of the student of development. And it is not necessarily true that a brisker rate of innovation always requires a higher rate of investment. Radical innovations in production processes may be introduced without substantially altering the rate of investment and therefore without changes occurring in the distribution of the capital stock by age.

The increase in the rate of growth of the product will then be accounted for by the greater improvement in quality of the more recent technological appliances. Among other things, this means that a link exists between development problems and scientific research problems.

Insistence on the importance of changes in the quality of capital does not however amount to a denial of the other strategic factors of the increase in productivity and particularly of changes in the quality of labour and the redistribution of resources among the various sectors of production. Attention will be devoted to these problems in the following paragraphs.

Continuing our analysis of the embodiment hypothesis, it will be well to consider some empirical evaluations of the dimensions of the disembodied component of technical progress and of the component embodied in capital goods in order to assess their relative magnitudes. It must however be explained that, in making the empirical estimates, the hypotheses of disembodied and embodied technical progress were at first assumed to be alternatives that excluded each other.

A first alternative measurement of disembodied technical progress and embodied progress was made by Solow[16] with reference to data concerning the American economy. Solow reached an estimate of 21/2% of disembodied technical progress a year.

Adopting the embodiment hypothesis instead, Solow built a vintage model chosen indirectly. Alternative series of tr were built, based on different assumptions regarding the weight to be assigned to machines and plant of different ages. Multiple correlation coefficients were then considered, together with the standard errors of the coefficients of the various regressions, and the statistically most satisfactory regression was chosen. This choice led embodied technical, progress to be estimated at 2% a year for machinery and 4% for plant. Later, however, disembodied and embodied technical progress was considered to be components existing together in a same function. A first synthesis was made by Phelps[17] and a more recent study has been made by Intriligator[18]. Intriligator's study is also distinguished by his having taken into consideration changes in the quality of the labour factor. His function is of the P(t) = A exp (e t) t(t)a $M(t)^{1-a}$ kind where t and M are capital and labour inputs weighted by changes in the quality of factors. The method of estimation is similar to the one followed by Solow in his study published in 1962. The regressions show clearly that disembodied technical progress and embodied progress are unsuppressible components of a same function: the former is estimated to be 1.67% a year, the latter 4% a year[19].

Lastly, a study by Mc Carthy[20], always based on American data, uses a C.E.S. function to measure the two components of technical progress. Disembodied technical progress is set at 1.2-1.8% a year and embodied progress is estimated to be of the same size.

A considerable dispersion of results is to be noticed in the inquiries rapidly passed in survey, especially as regards the embodied technical progress component. It must be added moreover that, although some regressions prove preferable to others from a statistical point of view, there are nevertheless many statistically significant

regressions that accentuate the importance of one component of technical progress or the other or actually exclude one of them.

The doubt may therefore legitimately be felt that, as a test bench, empirical evidence is not sufficient to render definitive the choice between one hypothesis or the other or to accept them both in proportions ascertainable with a high degree of reliability. It is this path of subtile epistemological scepticism that is pursued in a recent study by Jorgenson[21].

In the introduction to his study Jorgenson says in fact: «It has frequently been suggested that embodied and disembodied technical change are two different aspects of reality...

The purpose of this research is to examine the proposition that embodied and disembodied technical change may be distinguished by an appeal to evidence. After dropping the highly restrictive assumption that technical change proceeds at a constant exponential rate, we are able to show that one can never distinguish, a model of embodied technical change from a model of disembodied technical change on the basis of factual evidence. Both kinds of technical progress have precisely the same factual implications».

If this were so the consequences for economic policy would be considerable as disembodied and embodied technical progress would not set different problems in the strategy of development.

What is important is that the assumptions concerning reality do not prove incompatible; once this is ensured it would always be possible to test the equivalence of alternative models. It would merely be a question of finding out what rate of embodied technical progress corresponding to a given rate of disembodied technical progress is required, in a certain investment, for the rates of growth of the product to coincide. As can easily be understood, if Jorgenson's conclusions were to be accepted, a conspicuous part of what has been said so far would be devoid of interest. It is therefore my intention to make a critical assessment of Jorgenson's analysis for the purpose of showing clearly the meaning and scope of the results he reaches.

Jorgenson's Theoretical Construction as Criticism of the Distinction Between Disembodied and Embodied Technical Progress

Jorgenson considers a simple theoretical framework encompassing both disembodied and embodied technical progress: he is thus able to demonstrate the equivalence of the hypotheses. I will now explain briefly how he reaches this synthesis.

Given a two-sector model (respectively for consumption and investment goods), the value of the total output and that of the total input must be identical in any period, namely

$$q_c C + q_i I = Pk K + p_k L$$

where the «*p*»'s are the prices of factors and the «*q*»'s the prices of the goods produced – both consumption and investment goods.

From this expression it is easy to obtain the relative shares in value of the consumption and investment goods produced, which we shall indicate with v_c and v_i.

In the same way we can obtain w_k and w_1 which express the relative share of factors.

An index of the increase in output may therefore be:

Inserire formula

and an index of the growth of the global input will be:

Inserire formula

Since these are division indexes, we can then obtain the index of total factor productivity with the following logarithmic difference:

Inserire formula

The calculation of this index would be comparatively simple were it not for the problem of measuring capital services. If the capital input magnitude is obtained from data relating to investment transaction, there is the possibility of making mistakes of measurement in the attempt to separate prices from quantities.

This separation is necessary because it is generally only data for value that are available. We must therefore evaluate the consequences on the preceding formulas of systematic errors in the measurement of prices.

If *Q* represents the error in the prices of goods «*I*», the bias in the rate of increase in investment goods output will be:

Inserire formula where I^{+} stands for the investments measured. whereas I stands for the quantities actually invested.

Supposing that investment replacement takes place according to an exponential law of rate d, the K^{+} measured will then be equal to

Inserire formula

and the bias in the rate of growth of capital services will have the following formula:

Inserire formula

The index of productivity, calculated using the preceding magnitudes biased by systematic error may therefore be written:

Inserire formula

Making the suitable substitutions with the preceding formulas, we get correspondingly:

Inserire formula

If, suppose, we posit the bias of the rate of growth of productivity as equal to the reverse of its true rate of growth, which is equivalent to making **Inserire formula**, we shall have

Inserire formula*

This formula lends itself to interesting considerations. In formula [8] it is possible, in fact, to individuate a biunivocal correspondence between the true rate of growth of productivity and the particular error in the measurement of investment prices that annuls the estimated index of productivity. But just because of this biunivocal correspondence, any index of the growth of productivity can always be interpreted in terms of errors in the measurement of capital services. It therefore appears impossible to distinguish what is an increase in productivity from what is an error of measurement. This destructive conclusion, linked with the assumptions made, cannot be logically confuted.

However a heuristically fecund alternative exists, which consists of attributing a different interpretation to formula [8]: it is in fact in this way that Jorgenson sets about demonstrating the equivalence of disembodied and embodied technical progress. $\mathbf{P}/P$ stands for the rate of disembodied technical progress, while $1/Q$ now defines the changes in the quality of capital. The rate of growth of $1/Q$ therefore coincides with the rate of embodied technical progress.

I^{+} and K^{+} may then be taken to represent investments and capital stock corrected by changes in quality. Re-interpreted, relation [8] therefore links the rate of disembodied technical progress with the rate of embodied technical progress. A rate of embodied technical progress having been fixed, we are able to calculate the corresponding rate of disembodied technical progress, and conversely. The biunivocal correspondence guarantees the equivalence of the two rates of technical progress.

Now let us analyse the conceptualisation of Jorgenson's embodiment hypothesis in order to point out its distinctive features with respect to the categories we accepted previously.

In the preceding analysis, consumption and investment goods were always considered perfect substitutes in production; in other words the models adopted were one-sector models. In terms of a two-sector model, this amounts to supposing that the production functions are identical in both sectors. Furthermore, quality adjustments were limited to the capital stock used in production, while the output of new capital goods was not corrected except in terms of value. But in this way the analysis revealed some internal contradictions. There can be no doubt, therefore, that the passage from a one-sector model to a two-sector model and the introduction of corrections also as regards output renders the procedure of evaluating the increase in productivity more coherent even if, as has been seen, the problems involved by measurement call for caution with regard to the interpretation of results.

However the same critical sense is required also in assessing the result of Jorgenson's analysis. It is necessary to understand the meaning of the equivalence of the hypotheses of disembodied technical progress and of embodied progress. It is important to remember, when comparing factual assumptions describable in terms of embodied progress with factual assumptions describable in terms of disembodied progress, that the rate of technical progress and the rate of capital input growth are closely correlated. It ensues that the equivalence between rates of embodied progress and rates of disembodied technical progress is only meaningful once the time path of the capital input has been specified. If this condition is not respected, we cannot speak of alternative

hypotheses in relation to an identical factual reality. In other words, one cannot obtain a rate of disembodied technical progress corresponding to a given rate of embodied technical progress unless the path of capital stock is specified at the same time.

That a given factual reality may be explained by different hypotheses, all of them compatible is not a prerogative of economic science only; consequently it is not epistemologically surprising. In this case, however, Jorgenson's analysis has the fault of offering us a drastically alternative view concerning the hypothesis of technical progress, ignoring the possibility that the coexistence of components of disembodied progress and components of embodied technical progress may provide a more comprehensive explanation of reality. Jorgenson's epistemological scepticism is moreover a menace to any thorough explanation of the mechanism of development. The preceding models strike me, therefore, as being heuristically more fecund, even though Jorgenson's lesson should not be forgotten. From an operative point of view, moreover, I do not feel that I can fully share Jorgenson's conclusion that it is a matter of indifference which hypothesis is chosen. If it may not matter which alternative is chosen for an *ex post* explanation, it certainly does matter in the case of forecasting. Since the rates of technical progress are linked with the time paths of capital input, information concerning the past will only be valid for the future too, on condition that no change occurs in the time path of this input. Otherwise, given a certain rate of embodied technical progress, different rates of disembodied technical progress will correspond thereto according to the different possible paths of capital inputs. In other words, accepting the hypothesis of embodied technical progress, a rate of disembodied technical progress obtained from past statistics will not throw much light on the future unless the condition of reality repeating itself is posited. As can be seen, therefore, the acceptance or denial of the validity of the embodiment hypothesis is not without consequences on forecasts even if, a posteriori, it would always be possible to explain reality alternatively with one model or the other. Jorgenson's analysis, therefore, rather than being fecund from an operational point of view, lends itself above all to interesting epistemological reflections. Although we must not forget Jorgenson's warnings concerning the ambiguous meaning of increases in productivity, owing to possible errors of measurement, we cannot for this reason stop trying to individuate the different components of technical progress. At this point the doubt may legitimately be raised that empirical evidence is not sufficient as test bench for our hypotheses. But this must not prevent us from seeking to acquire a more thorough knowledge of the mechanism of development: I do not think the threshold of decreasing yields of analysis has been reached in this field yet. Proof of this is provided by recent studies placing in evidence the effects on productivity of the improvement in the quality of factors. Having discussed the improvement in the quality of capital, therefore, we may now consider the effects of improvements in the quality of labour.

Improvements in the Quality of Labour.

In the preceding paragraphs we were able to value critically the contribution made by embodied technical progress – that is to say by improvements in the quality of capital – to the rate of growth of global productivity. The trend term of the Cobb-Douglas dynamic model was then broken up into disembodied technical progress and embodied progress.

This enabled us to write the equality:

Inserire formula

the remaining disembodied component of technical progress.

We must now ask ourselves whether the latter component cannot undergo a further scission for the purpose of isolating the contribution made by improvements in the quality of labour. What we need to do is build a model in which the variable L is replaced by a variable M taking into account changes in the quality of labour. As in the vintage models, in which K was replaced by J, we can therefore write:

Inserire formula

«q» stands for the improvement in the productive efficiency of labour.

At this point, therefore, we can introduce a dichotomy: a) «q» may express an improvement in the average quality labour or b) «q» may express the improvement in the productive efficiency of the new workers who have had a higher level of vocational training.

In the first case it is necessary to draw attention to a fundamental difference as compared with improvements in the quality of capital. In the case of capital, technical progress is always embodied in new investments so that its effects are recognizable in two ways: through the flow of innovations and through the changes in the average age of the capital stock.

In formulas we have expressed this with the two following terms.

Federico Arcelli **Giorgio Di Giorgio**

Inserire formula

For labour, conversely, if «*q*» expresses an improvement in the average quality of labour, it is not necessary to make a distinction between old and new workers so that, in this case, changes in the age composition of labour will not have any effect on the growth of productivity.

If s = **q**/*q* iS taken to indicate the yearly rate of improvement in the efficiency of labour, the contribution made by the change in the quality of labour will then be equal to l *s*. If our model is of the Cobb-Douglas dynamic kind, we might even include this term in the disembodied progress component without the results of the analysis being distorted.

But when «*q*» expresses the improvement in the productive efficiency of the new workers, the analogy with the vintage model is complete. In this case we shall have to take into account also the change in the composition of the working classes by age.

Equation [1], which we have already had occasion to consider in connection with a study by Intriligator[22] – expressed in terms of rate of increase – will be written in two cases:

Inserire formule [23]

In case a) **Inserire formula** i.e. we have made a further breakdown of the disembodied component, which is above all of heuristic interest. But the result of the analysis would not have been distorted if *A'*(*t*) had not been broken down. In a certain sense improvements in the quality of all labour may also be defined as organizational progress.

The embodiment hypothesis, on the contrary, requires progress to be embodied in the new workers.

In case b), therefore, **Inserire formula** expresses changes in the average age of the working class[24]. In this case, as in the situation already analysed for the vintage models, if *A*'(*t*) had not been broken down, we should have overestimated the disembodied technical progress component to the detriment of the improvement in the productive efficiency embodied in the new workers.

Following this last breakdown of the technical progress term, *A*'(*t*) only includes the increases in productivity not directly embodied in the factors, i.e. productive improvements connected with the redistribution of factors among sectors of industry and areas, improvements in organization and management systems, etc. These are components we shall analyse farther on.

As regards improvements in the quality of labour with reference to case b), many economic analysis developments and related theoretical implications may be inferred by analogy from the elaborations expounded with reference to vintage models. I shall not, therefore, waste time discussing these aspects again as I feel it is sufficient to have pointed out the road to be followed.

What is more interesting, instead, is to pass in review the causes responsible for the improvements in the quality of labour and consider how an attempt may be made to quantify these terms. Denison's is definitely the most complete study in this connection[25]. He identifies some thirty factors that act upon the productive efficiency of labour; the factors that more directly increase the labour productivity are grouped, however, in three large variables. The first important factor that contributes to the improvement of labour is the level of education. According to Denison, 60% of the larger income correlated to better skills is ascribable precisely to education. Since statistical studies of the distribution of income show that an average increase of 10% in income corresponds to each additional year of education, 60% of this increase, namely 6%, is wholly due to education[26].

Denison calculates that in the period 1929-60 the increase in the average number of years of schooling in the United States caused a qualitative improvement of 1% a year in labour.

The second important variable as regards changes in the productivity of labour is the composition of the labour force by sex and age. Denison has estimated the effect of this variable on the improved efficiency of labour to be 0.1% a year.

Thirdly, the reduction of work hours determines an increase in productivity, but at a decreasing rate.

In the period 1929-47 half of the reduction in work hours is calculated to have made up for by the consequent increase in productivity, as compared with only one-third in the period 1947-60[26]. In the whole period 1929-60, however, the improvement in the productive efficiency of labour consequent upon the reduction of work hours may be assumed to have been about 0.3% a year.

Globally, therefore, *s* – which expresses the rate of increase in the average quality of labour – has been measured by Denison, for the period 1929- 60, at 1.4% a year. Positing **Inserire lambda**, the elasticity of the product with respect to labour, to be equal to 0.75, **Inserire lambda** *s*, namely the contribution made by the improvement in labour to the overall increase in productivity, is equal to 1% a year.

If we consider that the overall annual rate of increase in productivity in the United States, namely **Inserire formula** was estimated to amount to about 2% a year in the period in question, we have an immediate perception of the importance of improvements in the quality of labour[27]. Even if Denison's measurements may overestimate the contribution made by the improvement in the productive efficiency of labour, the results do in any case

redimension the importance of the technical progress component embodied in capital which had formerly been considered absolutely dominant.

However it must be added that, on the test bench of reality, the results do not appear to be indisputable. On the basis of tests of statistical significance, Intriligator's already quoted study, which also uses Denison's data for the labour factor, is inclined not to consider changes in the quality of labour important and calls our attention back, instead, to the disembodied and embodied technical progress components.

At the stage reached by research, doubts may once against legitimately be felt as to whether empirical evidence is sufficient as a test bench for our hypotheses. But the fact of having extended the radius of the discussion by stressing the importance of the human factor, while on the one hand it detracts from the dramatic tone of the disembodied versus embodied technical progress controversy, on the other hand opens out new and fecund perspectives to both theoretical and empirical research.

New Explicative Variables and Changes in Productivity. The Brown-Conrad Model

The attempts made to quantify the effects of qualitative changes in both factors by linking these magnitudes with other explicative variables are therefore particularly stimulating.

A recent study by Brown and Conrad[28] occupies a prominent position in this field; its object is to introduce technological research and education – as explicative variables – into a C.E.S. production function. In this study the parameters of the C.E.S. function also become variables and we move into the frame of a pluriequational model.

A labour productivity function having been derived from the C.E.S. function, the parameters are seen to be connected with all the fundamental variables of the model, among which are education and research. An empirical survey for the period 1950-60 has been carried out in ten sectors of American industry.

The variable «education» is essentially a measurement of the years of schooling of the labour forces employed. The improvement in the quality of labour brought about by better education is therefore directly proportionate to the years of schooling. The specification of the variable «research», however, gave rise to far from negligible problems as regards the identification of a suitable empirical correlate. It was a flow of knowledge, provided by research and influencing the productivity of labour, that had to be evaluated.

The increase in technological knowledge was assumed to be directly proportionate to the sums spent on research. However, for the purpose of taking into account know-how coming from outside, i.e. largely from firms with which business relations are entertained, a method was excogitated ad hoc to quantify the flow of know-how between industries.

The principal conclusions reached by Brown and Conrad's empirical survey may be summed up as follows: education and research have a comparatively more appreciable effect on firms manufacturing durable goods than on those manufacturing non durable goods. This can partly be accounted for by the fact that ties between firms belonging to the first group are stronger and the spread of the results of research and education therefore more marked. A rational policy regarding the allocation of funds to research should not therefore neglect these conclusions.

The introduction of the variable «research», however, has led Brown and Conrad to neglect the problems connected with the hypothesis of the embodiment of technical progress in capital: in this sense the contribution of the improvement in the quality of capital to the increase in productivity is no longer clearly individuated.

For Brown and Conrad, the variable «research» has, in fact, a generalized reductive effect on costs and does not therefore provide an adequate measurement of the qualitative changes in capital. The ties between research and embodied technical progress are not therefore unequivocally defined. Brown and Conrad nevertheless deserve praise for having opened up a new road to empirical research.

The breakdown of the technical progress term has therefore expanded the analysis, making the introduction of other explicative variables necessary.

Development problems are growing more clearly defined and a very promising line of investigation is already becoming recognizable, The wish to isolate the different components of technical progress, however, has temporarily obscured some ties of interdependence. It is these aspects that I intend to enlarge upon after completing the work of breaking down the residual disembodied technical progress. We have still to examine the effects of the redistribution of factors among the different sectors of industry for the purpose of distinguishing intra-industrial progress from inter-industrial progress.

Disaggregation of the Disembodied Component into Intra-industrial and Inter industrial Technical Progress

I took care to stress in the preamble the necessity of deepening – by means of successive disaggregations – the analysis of the production function built at the level of the whole economic system. Not only can the dynamic function with a trend term conglobing the effects of a series of unspecified variables not be considered stable, but it does not offer any instrument for the understanding of how the process of development takes place. Criticisms relating to the aggregation problem are therefore centred particularly on the technical progress factor. For the

purpose of dealing with these criticisms, we have preceded to make a number of breakdowns of the trend term, individuating the effects of a series of prevalently qualitative variables. So far, however, we are still within the limits of a macroeconomic function at the highest level of aggregation. The specification of the relation has improved as a result of the unexplained proportion of the disembodied component of technical progress having been reduced, but no proper disaggregation of the function has as yet been made. Proceeding in this direction, the disaggregation of the trend term allows us to trace a distinction between progress due to the re-allocation of resources among sectors of industry and areas and progress due to improved efficiency within industries. Both these factors contribute to the shift of the aggregate production function although, strictly speaking, only the second of these impacts should be termed technical progress[29].

Since a conspicuous part of technical progress is embodied in the new capital goods or is due to the improved quality of labour, however, it is obvious that the distinction between intra- industrial progress and progress resulting from the re-allocation of factors is not very clear and often difficult to identify empirically. The interdependence between technical progress and the new inputs introduced into the production process prevents the contribution made by the different components being clearly isolated. And this is also a subject to which I shall return.

Let us now follow the methodology elaborated by Massell[30] for the disaggregation of the dynamic production function. This will enable us to comment upon the results of de Meo's appliance of this method to Italian economic data[31].

Massell transforms the aggregate dynamic function so as to obtain intra-industrial technical progress as a weighted function of technical progress in each sector of industry. Inter- industrial progress is determined by subtracting the intra-industrial component from global technical progress. Inter-industrial progress will in turn be seen to consist of two components expressing the increases in productivity deriving respectively from the reallocation of capital and of labour.

Given the dynamic aggregate production function expressed in terms of rates of change[32]:

Inserire formula

we can write the corresponding equation for sector *i* as:

Inserire formula

Let us now define

Inserire formula

it ensues that

Inserire formula

The aggregate rate of growth of the product may be written:

Inserire formula

Substituting [2], [3] and [4] in [51 we get:

Inserire formula

Noting that

Inserire formula

we may write [6] as follows:

Inserire formula

Besides being carried out by sectors, as has been done, the same analysis may be carried out by areas. The left member of the last equation is equal to the rate of increase of global technical progress. Given the equality, the right member then expresses the rate of technical progress as the sum of three factors:

Inserire formula

The first term y_1 is the weighted average of the rates of technical progress in each sector and therefore expresses intra-industrial progress. On the other hand y_2 and Y_3 represent the increases in productivity due to a more rational allocation of factors: y_2 is the inter-industrial progress due to capital movements while y_3 measures the contribution made to efficiency by a better distribution of labour. We may ask ourselves what causes are responsible for the movements of factors among sectors and among areas.

The answer lies in the difference in the marginal productivity of factors employed in different sectors of production and in different areas of the same economic system. Dynamically considered, the problem must be viewed in relation to the different development of the sectors (or areas) and the different rates of sector (area) progress: the activities that proceed most briskly from the technical point of view tend to absorb the production factors of the less dynamic sectors (areas) owing to the differences in marginal productivity that come into being to the advantage of the more innovative sectors. It should be remarked that a higher level of vocational training encourages labour movements. In the same way the mobility of capital is generally associated with innovations that modify the development prospects of the different sectors (areas). Inter- industrial progress therefore appears to be frequently correlated with an improvement in the quality of factors. If this circumstance is overlooked, one is apt to overestimate the disembodied component of technical progress, here defined as intraindustrial progress.

The residual term must then be further broken down to individuate the components ascribable to technical progress embodied in the factors in compliance with the analysis developed in the preceding paragraphs.

De Meo's Calculations and a Critical Evaluation Thereof

We are now in a position to assess critically the results of some calculations made by de Meo. This assessment would have been full of gaps had it not been preceded by theoretical considerations. It must be stressed, however, that the purpose of de Meo's study is not so much to analyse the causes of technical progress as to calculate productivity. It is only the redistribution of factors effect that receives particular attention. In the introduction to his study de Meo pleads the difficulty of isolating the different components of technical progress as justification for this approach. Of course I fully realize this difficulty even if, in my opinion, the obstacles met with should not be allowed to stand in the way of further attempts to individuate the causes of technical progress empirically. In a certain sense therefore, the considerations I have made go beyond the targets set, which the research may be deemed to have attained.

Using a production function of the kind proposed by Solow, de Meo first calculates the changes in the per capita product in Italy in the period 1951- 63, distinguishing the contribution made by technical progress from that made by the deepening of capital. Then, applying the method suggested by Massell, he disaggregates technical progress into an intra-industrial component and an interindustrial component. Owing to the lack of data, the only sectors considered are agriculture, industry and services. Also considered, however, are three groups obtained from the binary combination of these sectors. At the same time the national territory is disaggregated into three areas[33]. It must be noted with regard to the first group of calculations – those relating to the distinction between technical progress and the deepening of capital – that:

a) the neutrality of technical progress is a hypothesis that simplifies the analysis and is for this reason frequently adopted by economists. It does not, however, appear to have been sufficiently tested empirically.

The absence of correlation between the rates of technical progress and the corresponding values for the deepening of capital K/L that results from de Meo's special dispersion graphs does not define the neutrality of technical progress beyond dispute.

As Resek[34] states «it is actually not too surprising for this test to show no relation in a non-neutral situation. For example technical advance could apply each year to the capital/labour ratio in use, but not to other ratios. Such a situation would not be neutral but would appear so to this test».

As an alternative test, Resek suggeste an examination of the correlation between the marginal substitution ratio and the capital labour ratio.

Neutral technical progress would require a very definite relation between these two variables as can easily be realized from the very definition of neutral progress in the Hicksian sense. But even this test, apart from the difficulty of obtaining the marginal substitution ratios, is not always decisive either. However, it does seem to be more reliable than the former test.

b) De Meo has not taken into consideration, empirically, the hypothesis of the embodiment of technical progress in capital goods. Yet, as has been seen in the preceding paragraph, it is only when the average age of the capital stock does not change that the complementary nature of technical progress and investments does not require a qualification of the results of the analysis made with a function of the $P = A(t) L^{l} K^{1-l}$ kind. In all other cases, the technical progress component is overestimated[35].

If, moreover, the factor substitution elasticity is other than the unit, or if technical progress is not neutral, the analysis made with a production function in which capital is assumed to be homogeneous leads, according to the case, to an over-or underestimation of the technical progress component.

It would therefore be necessary to integrate de Meo's analysis, proposing alternative functions in which – like in Solow's[36] and Intriligator's[37] studies – different hypotheses are assumed with regard to the absence or presence of embodied technical progress and its size. The tests of statistical significance, applied to the different functions, might widen the interpretative horizon of the data analysed, always of course supposing the measurement of capital with the permanent inventory method to be sufficiently reliable. Otherwise lack of precision in the data would render a more refined analysis superfluous, it does not seem out of place to mention in this connection Jorgenson's work on the possibility of interpreting systematic errors of investment measurement as rates of technical progress.

Probably, however, the hypothesis of the embodiment of technical progress in new capital goods is particularly important for Italy in the period 1951-63. In view of the fact that, as compared with technologically more advanced countries, our technology was on the whole considerably backward until the beginning of the fifties, the conspicuous investments made in that decade enabled the gap between average technology and vanguard technology to be narrowed, at the same time reducing the average age of the capital stock. But this actually means giving a particular significance to the embodiment hypothesis. It must not be forgotten that it is gross investments that are relevant with respect to the embodiment hypothesis. Now their value, in billions of 1954 lire, rose from 1,890 in 1951 to 5,342 in 1963[38]. These figures make it appear particularly important to break down the disembodied component of technical progress.

c) Lastly, it would be interesting to be able to carry out a research similar to Denison's for the purpose of individuating the contribution made by improvements in the quality of labour to the increase in productivity in the various sectors, suitably modifying the method of investigation.

As regards the disaggregation of technical progress into intrasectoral and inter-sectoral components, de Meo's calculations show that the intra-sectoral effect definitely prevails over the intersectoral effect. In the intersectoral effect, however, as was to be foreseen, it is the movement of labour that brings about the increase in productivity, whereas in Massell's survey of the American economy it was preponderantly capital movements that increased the efficiency of the system.

De Meo provides a particularly exhaustive explanation of the redistribution of factors and its consequent effect both on sectors of production and territorial areas; readers are therefore referred to this source for further information on the subject.

The almost simultaneous expansion of a great many sectors of the Italian economy and the relative slenderness of the intersectoral component as compared with the intra-sectoral component dispense us for the time being from examining more deeply the difficulty mentioned in the preceding paragraph of measuring the contribution made by each component correctly. As the interdependence between intra sectoral and inter-sectoral progress can only have a negligible weight for the moment, it does not adulterate the meaning of the calculations. In the next decade, however, the redistribution of factors is likely to assume more important dimensions with the result that the interaction between the various components will be more marked. It must not be forgotten that the redistribution of factors finds one of its principal incentives in the different rates of sectoral technical progress, while at the same time it accentuates the difference in the dynamics of the various sections of the national economy. It ensues that the various causes of increases in productivity (technical progress, improvements in the quality of factors, redistribution of resources) can only abstractly be considered independent. The part of de Meo's calculations dealing with the disaggregation of the trend term may therefore be held to interpret reality very efficiently as far as the past decade is concerned, but will probably have to be integrated with other consideration as far as the future is concerned.

A short reflection in more general terms seems advisable, therefore, regarding the interpendence of the causes of increases in productivity. These aspects deserve to be investigated for themselves as well as for their consequences on the parameter estimates obtained by means of linear regressions.

We have seen how the hypothesis of the embodiment of technical progress in new capital goods proposes to stand up to a certain type of criticism. We may now extend its heuristic fecundity by establishing, as Arrow has done for instance, a correlation between investments and technical experience, or increase its realism by avoiding the hypothesis of a constant yearly rate of technological improvement. But in the latter case the embodiment hypothesis is open to new objections, while old problems still remain: difficulty of measurement, failure to consider possible connections with the rate of improvement in the productive efficiency of labour, etc.

In other words, the more refined the decomposition of technical progress becomes, the thicker grows the network of interdependences between the various components and the greater the difficulty of measuring them by means of a simple linear regression. Nor should we forget the formation of multicollinearity phenomena, which has already given, rise to so many problems regarding the estimation of the classical Cobb-Douglas equation. No surprise should therefore be felt that hypotheses interpreting reality, which differ so widely, should, when applied to the same data, provide different relations, all of which have a high degree of statistical significance.

The formally complete solution of the problem, once the many components of technical progress have been individuated, would be the construction of an ampler development model explicitly recognizing the phenomena of interdependence.

However the progress already made in the endeavour to individuate empirically the different components of technical progress and the heuristic meaning of the preceding models should not be underestimated. In view of the difficulty of obtaining adequate empirical correlates to the theoretical concepts, also de Meo's study must, in this sense, be considered an appreciable step forward in the process of getting to know the working of our economic system.

Conclusions

Let me now recapitulate the results of this study of mine by briefly passing in review the principal conclusions reached. This will enable me to stress the appreciable progress made, not only in the field of theory, as a result of the decomposition of the technical progress term of an aggregate production function. The fact of having placed in evidence the links existing between changes in the quality of factors and technical progress is a step forward in the process of explaining the mechanisms of economic development and brings to light some valuable points for the solution of problems of economic policy. Not only have certain hypotheses for the interpretation of reality been better evaluated empirically, but some fundamental strategic variables speeding up the growth of the system's productivity have been individuated.

The analysis of the effects of changes in the quality of factors has been developed through successive approximations. First we considered the consequences of failure to include improvements in the quality of factors in a static production function. The conclusion reached was that if improvements in the quality of labour in general are neglected, the estimate of the partial elasticity of the product with respect to labour is biased downwards and if a positive correlation between capital and the quality of labour is accepted, the estimate of the elasticity of the product with respect to capital is affected, on the contrary, by a positive bias. As regards returns to scale, they are under-estimated in the hypothesis of the quality of labour changing proportionately less than the quantity of factors included in the production function. On the other hand, if the differences in the quality of capital goods are not taken into consideration, returns to scale are over-estimated if the change in quality has strong innovatory effects, while the bias of the partial elasticities of the product with respect to factors is generally under the same sign as in the preceding case relating to labour.

Passing from a static to a dynamic function, in which the sum of the elasticities is posited *a priori* as equal to the unit, the essential effects of the changes in the quality of factors merged in the trend term whose object was to express technical progress. In reality this expression is a disembodied component that varies in relation to time and includes all the unexplained residues of the production function. Not only is this term void of meaning, therefore, but the absence of decomposition makes the relation unstable and produces distortions of the analysis.

To deal with these fundamental objections, we followed three main lines of analysis in order to isolate what is really disembodied (for instance organizational progress) from what is dependent on other causes. First of all we improved the relation by substituting the family of vintage models for the class of production functions.

The hypothesis of the homogeneity of capital was thus discarded and stress was laid on the link existing between technical progress and investments: we therefore accepted the hypothesis of the embodiment of technical progress in the new capital goods reflecting vanguard techniques. The effect of the improvement in the quality of capital takes two forms: a) it steps up the rate of growth of productivity when the average age of the capital stock decreases as the gap between predominant technology and vanguard technology is narrowed; *b*) it steps up the same rate, following changes in the flow of innovations, measured by a rate of technological improvement embodied in the new investments. The flow of innovations is conditioned by the flow of inventions, which in turn depends on an appropriate scientific research policy: the prevalent effect is therefore a long run effect. Conversely the reduction of the average age of the capital stock resulting from a brisker rate of investments shows its effects in the short run.

The economic analysis showed that the interdependence recognized between technical progress and investments does not require a qualification of the results obtained with a $P = A(t)\, L^{l}\, K^{1-l}$ kind of function if no change occurs in the average age of the capital stock. This is, however, a particular case presupposing a unit factor substitution elasticity and neutral technical progress. In all other cases, that is to say: a) changes in the average age of the capital stock; b) non-neutral technical progress; c) factor substitution elasticity other than the unit, the analysis carried out with a production function in which capital is assumed to be homogeneous leads to over- or under- estimation of the technical progress term. This is the result we reached by examining more thoroughly the economic theory of vintage models.

This theoretical contribution led to the conclusion that the acceptance of the embodiment hypothesis makes it necessary to break the trend term into its components, not only for heuristic reasons, but above all to avoid distortions of the analysis. Thanks to this result we were able to assess critically the empirical significance of the hypothesis of the embodiment of technical progress in contradiction with a thesis upheld by Denison that tended to redimension the importance of this assumption at least as far as the technologically more advanced countries are concerned. I have also taken care to describe the most recent inquiries carried out by Solow, Intriligator and Mc Carthy on the American economy, which, with the aid of vintage models, estimate simultaneously the disembodied and the embodied component of technical progress. It must be pointed out, however, that the results,

despite the ingenious estimation procedures resorted, do not appear quite certain, although the relations that have been checked pass the statistical significance test brilliantly. The dispersion of the results corresponding to the different hypotheses for the interpretation of reality may in fact lead one to doubt whether empirical evidence is always sufficient as a test bench when making a definite choice between one hypothesis and another. Following this path of subtle epistemological scepticism, Jorgensen refutes, with an ad hoc model, the distinction between disembodied and embodied technical progress, attempting to demonstrate their equivalence on the empirical piane. In his radical criticism Jorgenson goes even farther, showing how any index of increased productivity can always be interpreted in terms of systematic capital measurement error. Without ignoring this last warning, I have been able to reject Jorgenson's destructive conclusions, showing that they appear to be tied to the particular assumptions of the model selected. The acceptance or rejection of the validity of the embodiment hypothesis is bound to have an effect on forecasts even if, ex post, reality can always be explained alternatively by one model or the other. Jorgenson's analysis, therefore, rather than being fecund from an operational point of view, allows above all interesting epistemological reflections to be made, Moreover the epistemological doubt cannot halt the process of seeking to acquire a better knowledge of the mechanisms of economic development.

In the same way as for changes in the quality of capital, the analysis of technical progress is undertaken by isolating the contribution made by the improvement in the quality of labour.

I have pointed out, however, that a fundamental difference exists as compared with the improvement in the quality of capital. In the case of capital, technical progress is always embodied in the new investments and therefore its effects are recognizable in two forms: through the flow of innovations and through changes in the average age of the capital stock. In the case of labour, on the contrary, what is generally considered is a change in the average quality of labour and consequently it is not necessary to make a distinction between old workers and new. It is in fact problematic to find a term corresponding exactly to the average age of the capital stock. In the case of capital the embodied technology may be considered unchangeable, in the case of labour, on the other hand, there is a process of learning related to productive activity that unequivocally alters the basic vocational training levels of the different labour levies.

Particular consideration has therefore been given to an inquiry of Denison's carried out to establish the causes of the improvement in the productive efficiency of labour and quantify their effects. According to Denison the contribution made by the improved productive efficiency of labour is particularly important and calls for a greater awareness of the share of the higher rate of productivity ascribable to the improvement in the human factor.

New research trends thus open out, whose purpose is to connect the changes in the quality of factors with other explicative variables. Important in this respect is a recent study by Brown and Conrad who suggest that technological research and education be included in a C.E.S. function as explicative variables. With Brown and Conrad the parameters of the function, which appear to be connected with the model's fundamental variables, become variables too and the analysis moves into the sphere of a pluriequational model. This approach appears particularly promising once the interdependence of the different components of technical progress has been recognized.

The third line of the analysis of technical progress consists in its disaggregation into intra-sectoral and inter-sectoral technical progress. The disaggregation of technical progress in compliance with a method proposed by Massell allows a distinction to be traced between the increase in productivity consequent up on the redistribution of factors by sectors and areas and the improved efficiency developing inside these sectors and areas.

It is thus possible to separate an intra-sectoral component, equal to the weighted average of the rates of technical progress in the separate sectors, from an inter-sectoral component which can in turn be divided into two parts: the first measuring inter-sectoral progress due to capital movements and the second measuring the inter-sectoral progress due to the redistribution of labour.

After having brought to completion the three fundamental lines for the analysis of technical progress, we were thus able to examine the results of some calculations of de Meo's relating to the Italian economy. A long paragraph contains all the observations made on this inquiry. I have also suggested some alternative research methods that allow recent hypotheses interpreting the causes of technical progress to be assessed. At this point it seemed advisable to reflect in general terms up on the interdependence of the causes responsible for the increase in productivity. After going more thoroughly into the process of decomposition of technical progress by analysing the contribution made by qualitative improvements in factors and their more rational allocation, a thick net of interdependences became recognizable. Once the many components of technical progress have been recognized, the formally complete solution that I point out is the construction of pluriequational development models explicitly recognizing the phenomena of interdependence.

However a substantial solution would require a thorough study of a series of exogenous variables influencing productivity that are scarcely considered at present. Steps would then have to be taken to link them up with the preceding models. But it must not be forgotten that these are often qualitative variables that are impossible to quantify strictly. Brown and Conrad's model provides a convincing proof of this. The theory of development cannot be completely summed up in a system of equations.

Federico Arcelli **Giorgio Di Giorgio**

Federico Arcelli **Giorgio Di Giorgio**

CHAPTER II
NEOCLASSICAL AND KEYNESIAN ECONOMICS

Federico Arcelli **Giorgio Di Giorgio**

FOREWORD

Mario Arcelli wrote the two following papers in the early Seventies before his experience at the M.I.T. of Boston, when he held the chair of economics in Padua. This first paper aimed at critically considering, in the perspective of the general systems theory, the essential features of the information structure in the Walrasian theory of General Economic Equilibrium and in the Keynesian model. The second one, "Keynesian Theory and Neo-Quantity Analysis: towards a Synthesis" gives an overview on the la test thoughts expressed by the Keynesian economists, with explanations of reality, including suggestions made by the followers of the Chicago school regarding a rational approach to economic policy, reveals many different positions today that cannot easily be reduced to a theoretical unit.

The comparison between the traditional Walrasian models and Keynesian theory is of particular interest in the context of the economic development in the early Seventies. In fact Mario Arcelli was writing before the oil crisis and when Italy was at the end of the long period recorded as the one of "lost occasions" of the late Sixties. Anyway he tries to understand the possible interpretations that modern economic theory could give about some situations, and, in particular, he stresses that "the essential features of Keynes' theory and its dynamic properties do not stem, however, merely from the recognition that transactors have not a perfect information in the sense of Walras, but also from the specification of assumptions concerning the structure of the economy that make it possible to falsify its propositions with reference to empirical reality. Keynesian theory in its recent interpretation, is therefore a systematic description of fundamental relationships among economic variables, whose control sphere rests on an information network that is well articulated and suitable to satisfy the standards proposed by Kornai." And also "Keynes' contribution to economic theory may be viewed as a great effort to extend the use of the tools of value theory beyond the scope of general economic equilibrium, into the area of macrodisequilibrium".

But the idea of the second paper is maybe more ambitious. The target is to compare and try to define the points of contact between Keynesian and Monetarist theories. Mario Arcelli notes that the "Keynesian doctrine, as time passes, has successively laid stress on different features of its analysis, so that it becomes increasingly difficult to give a full definition of what truly forms the corpus of this doctrine today. On the other hand quantity theory has absorbed some notions of Keynesian inspiration as it continued and still continues to be re-elaborated; as a result more softly shaded positions have recently been added to the original sharp contrapositions (maintained as a point of honour).". This could better explain how it is that the debate between Keynesians and Monetarists often involves different levels and perspectives of analysis. But this is not all: "a new approach to the original Keynesian thought states that many assertions concerning the presumed ineffectiveness and unimportance of monetary policy are completely extraneous to Keynes' teaching and should therefore be repudiated. This being without prejudice to some fundamental differences with respect to the monetarists".

The result of the study is very interesting, stating one could argue that the two schools, though starting from a common theoretical point of reference (general economic equilibrium), come in the end to two widely different philosophies. "That of the Monetarists reflects the natural order of the neo-classical construction, from which it deduces rules of economic policy designed not to disturb the spontaneous but ordered development of the economy. Money cannot alter the real order of facts in the long run, therefore it is only right that the supply of money, determined exogenously, should grow according to a uniform rate. The Keynesian philosophy, on the other hand, inspired by a theory of macroeconomic disequilibria caused by the absence of perfect information as well perhaps as by the presence of monopolistic elements in the system, reflects a less illuministic conception of economic activity and consequently assigns a more flexible role to the instruments of economic policy, monetary changes included. The quantity of money is therefore not only determined exogenously; it also reflects the needs of the economy and thus possesses an endogenous dimension too. The difference between the two schools, then, lies not so much in the mechanisms transmitting monetary impulses to real variables, which anyhow still appear controversial, as in the different rules of economic policy and the different philosophies supporting the logic of the interventions".

Federico Arcelli **Giorgio Di Giorgio**

INFORMATION STRUCTURE IN THE WALRASIAN THEORY OF GENERAL ECONOMIC EQUILIBRIUM AND IN THE KEYNESIAN MODEL [34]

This paper aims at critically considering, in the perspective of the general systems theory, the essential features of the information structure in the Walrasian theory of General Economic Equilibrium and in the Keynesian model. This theme has already been the object of separate or partial contributions in the recent literature, which can be fruitfully consulted [Sec: 1; 3; 4; 7; 8; 11; 12; 14; 17; 18; 19].

Our contribution to the argument consists chiefly of a scrutiny and a synthesis of most relevant conclusions. The considerations we shall formulate are in line with the rethinking of economic theory, inspired by the new scientific image of the world (a new «paradigm» T. Kuhn [15] would say) that discovers in the application of cybernetics and more generally of systems theory to the study of economic problems, a promising method to overcome some inadequacy of traditional research and theorizing. It is an approach, which has been marked with brilliant achievements also by Polish economists: first of all, by Lange and Greniewsky [17; 10]. Contemporary socio-economic systems and underlying technological and organizational structures are so complex that it becomes inevitable to employ global approaches to analyse their mechanics. The general systems theory, as defined by von Bertalanffy [29], offers a powerful set of instruments and methods of research apt to modify deeply our perspective and comprehension of the working of economic systems.

The economist becomes acquainted with concepts drawn from information theory (see Theil [27]) and from control systems analysis (see Milsum [20], Tustin [28]). He discovers in the working of economic systems the existence of a control sphere where information flows are recorded and control processes take place, dualistic to the real sphere where real processes are described. The advantage to reappraise the Walrasian scheme and the Keynesian model in the new perspective offered by systems theory has become more and more evident. A recent book by the Hungarian economist J. Kornai [14] supplies us with a new conceptual network and with a deep criticism of the general equilibrium model, developed according to the paradigm of the systems theory: his comments and observations which we do not always partake, are nonetheless sweeping and far-reaching and induce us to reconsider the weight of some contributions to the history of economic thought in a new light. It is however possible to proceed beyond this goal and to identify, by the same logic approach, the information structure of the Keynesian model, which better, describes the working reality of western economies. Some attention to this topic has been already devoted by recent reconsiderations of Keynes' doctrine [7; 8; 9; 11; 18; 19].

This new trend of analysis, which reappraises the authentic Keynesian thought, as opposed to income-expenditure vulgarisations exposed in textbooks, also relies heavily on the logic of modern systems theory, even if it often employs its concepts and instruments in a rather unsophisticated and non-technical way. We refer to Clower's, Leijonhufvud's and Hines' works whose heuristic fruitfulness lies particularly in an interpretation of Keynes' thought, rooted in the Walrasian general equilibrium framework, from which Keynes' theory departs assuming a different information structure. It is a major merit of Clower's and Leijonhufvud's analysis to have a clear connection between Keynes' macroeconomics and Walrasian microeconomics established: this contradicts Kornai's statement that «to this date there has been no organic integration between Keynesian macroeconomics and the traditional western microeconomics» [14, p. 371]. This achievement could also make us rediscover (but this is more controversial) a continuity in economic theory, which would grant to the Walrasian system the meaning of archetype in respect to subsequent developments of economic analysis.

Theory of tâtonnement (groping), stability and systems theory in the Walrasian model: a reinterpretation.

As it is well known, great was Walras' admiration for Newtonian physics: Walras' final object was the finding of a system of natural laws that is able to explain economic equilibrium in the same way as mechanics did for physics. After the end of the illusion of the existence of a natural pervasive order in the economy, the Walrasian system had to face very strong epistemological objections not easy to overcome.

A partial rescue of its explanatory power of reality or at least of its heuristic fruitfulness, may be granted however, by considering the Walrasian-model in the perspective of systems theory: both the formalization made by Walras and by the subsequent economists who achieved the formal perfection of the model and the comment exposed in the «Elements» [30] and revived by the long debate relative to the existence of equilibrium, its

---

[34] From: M. Arcelli, L´Industria, Milan (ed. L´Industria), fasc. 3 / 4, pagg. 3-17, 1972. Paper presented at the "Convegno Italo-Polacco" on «Modern application of mathematical systems and control theory in particular to economic and production systems», Gacow, 14-20 sets., 1972, organised by the Polish Science Academy.

stability and the adjustment mechanisms relevant to the clearing of the markets, are suitable to an interpretation in terms of systems theory.

Accepting the definition of a system as a set of interdependent elements or also as an integrated unity of structure and functions, we find reflected in such a general definition the Walrasian formalization of general economic equilibrium. The homogeneity of Walrasian model with that framework becomes evident when one insists on the whole rather than on the single elements, on interdependencies rather than on the individual and unconnected behaviour of the micro-unities. The systems theory, in fact, underlines the concept of organization and therefore explicit relationships among actors of the economy. If we follow that approach, it is then possible to understand better, the role of actors inside the economic system; the impact of communications on decisions; the notion of information structure; the concept of response function, of feed-back and so on. It also becomes clear that we are provided with an alternative conceptual network to interpret the mechanics of Walrasian system and, more in general, of economic models: such is the way followed by Kornai in his «*Antiequilibrium*» [14], although with a critical intent towards general economic equilibrium.

This turn, in the consideration of the Walrasian world, appears to be the natural conclusion of the long debate on the meaning and the role of relative prices in neo-classic economics and of the reappraisal of the Walrasian theory of *tâtonnement* in determining equilibrium and in proving its stability. In this connection, we should mention Patinkin's masterly work, «*Money, Interest and Prices*» [22], with special reference to Note B, analysing Walras' Theory of *tâtonnement*: the statement of the central role of this theory in Walras' argument and of the substantial equivalence between equilibrium stability and dynamic convergence of tâtonnement, offers a fundamental key to an interpretation of the Walrasian model within the framework of systems theory.

According to Patinkin, Walras through his *groping* «was primarily interested in describing, not the actual path to equilibrium, but the nature of automatic forces which propel the economy along this path – be what it may». The final aim of Walras' analysis was therefore the proof that «no matter what the initial position from which the system starts, the dynamic working of market forces brings it to its equilibrium position». This is equivalent, however, to acknowledge the existence of signals and of an information structure that, through proper automata, is apt to control the regular functioning of the economic system.

Although analytically inadequate, Walras' formulation represented, therefore, an attempt to solve the problem of equilibrium stability, even if the question had not been fully understood by Walras himself, according to Patinkin. Long neglected because of its inadequacies and of the poor comprehension of its meaning, with some remarkable exceptions (Lange, for instance, employed it to establish the feasibility of a socialist economy [16]), the Walrasian theory of *tâtonnement* meets its rehabilitation in the recent and the most rigorous treatments of the stability of competitive equilibrium. Samuelson's definition of stability in the small [26], but especially the finding of the conditions of global stability of the competitive equilibrium by Arrow, Block and Hurwicz [5; 6] mark the end of «the seventy-year analytical exile of Walras' theory of *tâtonnement* and constitute the first significant development of this theory beyond Walras' pioneering presentation» [22, p . 540]. Recently Allingham [3] has investigated «the working of a multisectoral *tâtonnement* stability process in an actual economy by simulating the adjustment path of an abstract disequilibrium model in conjunction with an econometric equilibrium model, indicating that such path is highly stable».

The information structure in the Walrasian system.

Underlying the discussion of the dynamic stability of the Walrasian equilibrium is an internal description of a dynamic system of which steady state solutions are looked for when there exist automata, like Walras' market forces, that are able to warrant the convergence of the system towards the equilibrium path, we say the system is stable. Whenever, on the contrary, the structure of the model does not show this property (what is excluded by Walras by means of a non - exhaustive argument); the remedy would consist in inserting in the structure feedback loops apt to dampen the path towards the steady-state equilibrium. This hypothesis, however, is not contemplated by the static analysis of general economic equilibrium. The debate on the equilibrium stability forces us therefore to investigate the nature of implicit stabilizers, the decision making processes and the information structure of the model; at the same time, the argument allows us to realize the convergence between the theory of dynamic stability and feed-back control systems theory: a chief goal of a control system actually being the achievement of stability.

Kornai's distinction of the processes which take place within the economic system between real processes which belong to the real sphere and control processes, which belong to the control sphere, actually appears particularly suitable to an inspection of the working of economic systems. «The real processes of the economic system are material, physical processes. These are production (including transportation, warehousing, material services etc.) consumption and trade. The control processes of the economic system are intellectual processes. These include observation, information transmission, information processing, decision preparation and decision-making. Real processes are described by real variables, control processes by control variables. The unequivocal separation of the real processes from the control processes is called the dualistic description of the economic system» [14, p. 39]. One can easily ascertain that developments of economics have been significant especially in

the real sphere; but less remarkable in the control sphere and very poor with regards to the description of information flows and the understanding of their impact upon the working of economic systems. Following Kornai we can divide information into three main classes: a) Money flows; b) Price-type information; c) Non-price-type information.

Money flows define a special class of information whose role is to transmit command over goods and services through the availability of purchasing power. The significance of this information is substantially determined by market structures (competitive, administered, monopolistic) and is also connected to institutional and structural characters of the economic system under inspection. The information of the third set (non-price-type information) includes all information (more and more important in the present day reality) not belonging to the other two classes. It includes, therefore, both qualitative and quantitative messages. The weight of different kinds of information of this group depends, of course, on institutional and structural characters of the economy and on spatial and temporal circumstances as well. The significance of the information of the second set (price-type information) depends chiefly on resources allocation criteria and on market structures and institutional characters of the economic system. In a totally decentralized economic system (a Walrasian free market economy), prices fulfil two tasks simultaneously, that is: they spread information required by the coordination of the plans of economic actors and provide incentives to an efficient resources allocation among different activities.

Plans coordination determines equilibrium, which is stable if the circulation of information is so pervasive as to generate forces dampening every shifting from the equilibrium position. The smooth working of the system requires, therefore, that every transactor must know, directly or indirectly, what is happening elsewhere, that is, must be able to receive messages, to evaluate them and reply to market signals in a foreseeable way. In the Walrasian system the information-communication problem among micro unities is actually solved fully and exclusively by means of price-type information. The rules governing the control of activities so as to make possible their coordination are, in fact, the following ones: [19]

1) Price incentives are effective; the transactors do react to changing relative prices by modifying produced and consumed quantities of goods and services in a foreseeable way.
2) Prices move «freely» in reply to excess demands and supplies in such a way that transactors are solicited to adjust their behaviour in the direction required by the achievement of full-employment of resources.

These rules, deductible from the theory of tâtonnement that is intended to simulate the working of the market, assure the equilibrium of the system instantaneously if false trading is avoided by means of Walras' implicit assumption of recontracting (Walrasian tickets). As we have already said however, the conditions of convergence of *tâtonnement* must be rigorously analysed within the framework of dynamic stability of the system. Stability conditions therefore must not contrast with the working of the market.

The acceptance of the existence of a stable equilibrium also implies the recognition of an implicit homeostatic system, that is of servomechanisms (in our case, the market forces) controlling relative prices: actually, this is what can warrant the clearing of the markets. Conversely, any interpretation of the Walrasian theory in the perspective of systems theory and therefore any deep insight into its information structure cannot leave out of consideration dynamic stability. Therefore the simple static formalization of general equilibrium system does not reflect adequately the complex interaction of decisions of micro-unities. Prices, in fact, are taken by the transactors as data. They choose the quantities to buy and to sell, to consume and to produce without worrying about prices. In other words, «no transactor would ever face any constraint on the quantities he can buy or sell at the market prices at which trading actually takes place. No transactor is ever required to final terms for himself [18, p. 76]». The transactors take prices simply as data, because «every morning», so to say, they are supposed to receive the market-clearing vector of prices tried out by the «auctioneer». This list of prices gives all information that is useful and necessary [19]. That is because «if the market clearing vector of relative prices (which is assumed to exist) is known by all transactors, and if each transactor adjusts to these parameters, the system simultaneously generates an optimal allocation of resources and ensures that they are fully employed [11, p. 11]».

Not only all information required by a consistent working of the economy is disseminated through prices, but also the degree of information of each transactor is optimal and excludes uncertainty. The anonymity and simplicity of information, which are by-products of a pure market, go together with the absence of costs for the transactors. Moreover it has been proved by Hurwicz [12] that compared with different adjustment processes, the one supported by the information structure in the Walrasian scheme shows the property of information efficiency, since it achieves the same results with less detailed information.

Kornai's criticism of the information structure in the Walrasian system.

Such simplicity and thriftiness of information is however the source of Kornai's strongest criticism, who opposes many serious objections to the Walrasian assumptions in order to prove the poor explanatory power of

that scheme. In brief, Kornai ascertains that the Walrasian information structure denies any role to information belonging to the first and third group of his classification; the information is direct, with exclusion therefore of organizations specialized in transmitting it; it is anonymous and comes from a single-channel: the market; the time of information is single-phase; its degree of fineness is constant and very small; there is no room for uncertainty.

Instead for every actually existing economic system, the information structure is complex. Historically, the complexity of the information structure increases with the expansion and development of the real processes. There is an ever greater amount of information connected with the same event which flows sequentially, over the course of time or simultaneously; in other words the information flow is expanding [14, p. 62]». «The complexity of decision problems, uncertainty, the comparative unreliability of information and the increasing risk involved in important decisions made under uncertain condition le ad to multi-channel information and the multiplication of information gathering. Concerning the same real event multiple information, partly of price-character, partly of non-price character will be received through several channels, in various phases, at several times, in various degrees of fineness [14, p. 66]».

The principle of information multiplication rejects therefore the simplicity of the Walrasian information structure. The refusal of Walrasian theory by Kornai is total also from the heuristic point of view, since he denies any connection between that system and the more operational Keynesian macroeconomic model. He states explicitly: to this date there has been no organic integration between Keynesian macroeconomics and the traditional western microeconomics (nor could there be).

Kornai excludes also that the Keynesian model can be interpreted in terms of systems theory: «It belongs largely outside the scope of economic systems theory» [14, p. 371]. If all this were true, the gap existing in western economic theory, would be irremediable and would deprive the general equilibrium theory of any heuristic significance.

However the recent stream of analysis including Clower's, Leijonhufvud's, Hines' and others' works that revaluate Keynes' original thought, pointing out the importance of relative prices within Keynes' framework and so establishing an integration with markets theory, allows us to reject Kornai's most destructive statements and to confirm, at least partially, the heuristic fruitfulness of the Walrasian system. It allows us also to proceed beyond Kornai's conclusions, by reconsidering Keynes' model in the perspective provided by cybernetics. By doing so, we should line out the connection between Keynesian macroeconomics and Walras' markets theory and remark how information structure in the former reflects a substantial change in respect to the latter.

The information structure in the Keynesian model: a more general scheme.

The «new image» of Keynes, supported by Clower's and Leijonhufvud's interpretation, suggests indeed an useful integration of macroeconomics with microeconomic analysis, while, at the same time, providing us with a theoretical explanation of unemployment which (contrary to traditional Keynesian literature) needs not rest on empirical assumptions like wage rigidity or on restrictive hypothesis like liquidity trap.

It is possible to say, after the reappraisal of the original Keynes' thought, that the rate of interest and relative prices perform within his theory a more important play than that described by post-Keynesians: it is therefore appropriate to distinguish Keynesian economics from economics of Keynes.

According to Leijonhufvud [18, p. 67] «The revolutionary impact of Keynesian economics on contemporary thought stemmed in the main from Keynes' reversal of the conventional ranking of price and quantity velocities. In the Keynesian models price velocities are not infinite; it is sometimes said that the implications of the model result from the assumption that money wages are rigid. This usage can be misleading. Income constrained processes result not only when price-velocity is zero, but whenever it is short of infinite».

What discriminates a Walrasian system from a Keynesian model is therefore the different role of variables in the adjustment process starting from disequilibrium.

In a Walrasian system, perfect flexibility of prices guarantees the fulfilment of full employment; in Keynes' model, on the contrary, the adjustment proceeds faster in terms of quantities rather than of prices (as a consequence also of inelastic price-expectations) and this may involve some resource - unemployment. From a different point of view, this character of a Keynesian economy is amenable to an information problem, which may be analysed by cybernetics [4]. In Walras' system perfect information is implied by the auctioneer mechanism and the groping search, which determine a market-clearing vector of prices.

Keynes' world instead, even if consistent with the neo-classical assumption of rationality of economic actors and with the scarcity logic of price-incentives, departs from that universe by accepting a less than perfect information. In other words, prices are still effective incentives in influencing resource-allocation, but they do not assure the consistency of plans of independent transactors any longer and thereby they cannot ensure full employment. An unemployment situation may therefore arise owing to a vector of relative prices departing from the market clearing vector of prices: the impossibility to operate the necessary corrections before erroneous transactions take place feeds the disequilibria. It is not necessary then to assume either wage rigidity or the

existence of a liquidity trap to verify unemployment conditions: it is sufficient that the vector of existing relative prices be different from the equilibrium vector as a consequence of inadequate information.

As Clower emphasizes, many transactions, which would have taken place at equilibrium prices, are not realized at the actually prevailing prices. This fact sets up a dichotomy between notional and effective demand, the latter being constrained by monetary income obtained through realized sales.

Price-inelasticity of expectations renders transparent the possible source of a gap between notional and effective demand as a consequence of hoarding money. This remark underlines the distinctive feature of a monetary economy like the Keynesian, where transactions are not constrained by the initial resources (the so called budget constraint) but by monetary availabilities. In full equilibrium situations, notional and effective demands coincide, but as soon as prevailing prices are not the same as market clearing prices, the working of the economy appears to hinge on effective demand.

Let us consider the specific example of an initial disequilibrium situation on the labour market: current income of workers supplying labour services is not determined by the quantity of services they would like to sell at existing market prices, but by the services they actually succeed in selling. It follows that their effective demand on other markets is constrained by income actually realized. Notional demand becomes unrelated to the trading process. Entrepreneurs therefore will curtail their purchases and possibly, prices on related markets and also on the labour market, since they link their decisions with realized receipts, which are smaller than, expected. This is the start of a dynamic process generating a cumulative contraction that brings the system into a Keynesian unemployment equilibrium.

The lack of price-flexibility, which is the mirror of sub optimal information, involves disequilibria and generates unemployment. The unemployment-deflation cycle, which is a vicious one, may be well represented by a typical positive feedback, described by cybernetics [20, p. 7]. To sum it up, the full meaning of Keynesian theory may be grasped only within a dynamic framework where the multiplier is described as a feedback loop magnifying the original impulse: the absence of homeostatic mechanism controlling relative prices, detaches the economy from full-employment conditions.

Negishi [21] remarks that this departure from Walrasian adjustment process *par tâtonnement*, besides generating the described chain reaction (spill over effect), determines also redistributional effects because of false trading. But this is not the central point. What Keynes has tried to show is that full coordination of decisions is impossible because of an intrinsic lack of information. In fact, Keynes' fundamental problem was to analyse «the economic behaviour of the present under the influence of changing ideas about the future»[13]: now the dissemination of adequate information about intertemporal price vector suitable to coordinate savings and investments at full employment is not achieved except casually.

Instead of a known vector of intertemporal prices, the transactor is provided with expectations about marginal efficiency of capital and has a liquidity preference guiding him in the choice of his portfolio composition. If plans of the transactors are to be consistent in the aggregate, «both the marginal efficiency of capital and the structure of asset prices must correctly anticipate and reflect future prices and quantities in commodity and financial markets» [11, p. 15]. But this proves to be absurd, owing to continuous revisions of expectations and in the absence of appropriate forward markets that can substitute the Walrasian auctioneer or the adjustment process *par tâtonnement*.

It is true that Keynes mentions the Stock Exchange in his *General Theory*; but immediately he reminds us of the distortions generated by speculation and so discards it as a means of perfect information about future yields. Unemployment is then not necessarily the result of price-rigidities, but more often of the poor information, connatural to the logic structure of the model. The transition from the atemporal Walrasian world, where the knowledge of equilibrium price vector is perfect, to the Keynesian world where expectations rule, marks the leave of a class of invariant states of the model and the attribution of a special *status* to money.

Only in the Walrasian system, prices are the one channel of information for individual decisions: the lack of uncertainty makes all goods perfectly liquid and simplifies the information sphere. Instead, Clower's and Leijonhufvud's analysis describing the behaviour of transactors within a Keynesian framework, reveals the formal complexity of a model where both market prices and expected prices, money flows and realized and anticipated quantities have an impact on transactors' decisions.

The control sphere of a Keynesian model is characterised by a complex information structure, which expands over all three classes of information described by Kornai. Prices still play a major role in the decision-making: they fulfil the role of incentives, even if they do not warrant consistency of plans in the aggregate. But by the side of this type of information, both money flows and non-price-type information (including excess demands or supplies) perform a prominent role.

Money flows, in an income-constrained approach, identify the expenditure limits and determine in the aggregate, the level of total effective demand.

Information of the third group (non-price-type information) is important especially in evaluating the marginal efficiency of capital and the liquidity preference, influencing thereby saving-investment and hoarding decisions.

When one leaves the timeless logic of neoclassic economics and cancels out the Walrasian auctioneer leading to equilibrium prices, the system appears to be ruled by uncertainty.

It is then logic to keep account of information costs, as Alchian [1] did, and to assume a complex and multi-channel information structure, as in Clower's and Leijonhufvud's analysis.

The essential features of Keynes' theory and its dynamic properties do not stem, however, merely from the recognition that transactors have not a perfect information in the sense of Walras, but also from the specification of assumptions concerning the structure of the economy that make it possible to falsify its propositions with reference to empirical reality. Keynesian theory in its recent interpretation is therefore a systematic description of fundamental relationships among economic variables, whose control sphere rests on an information network that is well articulated and suitable to satisfy the standards proposed by Kornai.

Keynes' contribution to economic theory may be viewed as a great effort to extend the use of the tools of value theory beyond the scope of general economic equilibrium, into the area of macrodisequilibrium [18, p. 333].

MANCA TUTTA LA PAGINA 15 DEL TESTO ORIGINALE!!! INSERIRLA!

Federico Arcelli **Giorgio Di Giorgio**

Bibliography of Works cited.

[1] Alchian A., Information Costs, Pricing and Resource Unemployment, «Western Economic Journal», 1969, June.
[2] Allen R G. D., Mathematical Economics, London, Macmillan, 1960.
[3] Allingham M., Tâtonnement Stability: an Econometric Approach, «Econometrica», 1972, n. 1.
[4] Arcelli M., Teoria Keynesiana e analisi neo- quantitativa: verso una sintesi, «Rivista di Politica Economica, 1972, n. 4.
[5] Arrow K., Hurwicz L., On the Stability of the Competitive Equilibrium I, «Econometrica», 1958.
[6] Arrow K., Block H., Hurwicz L., On the Stability of the Competitive Equilihrium II, «Econometrica», 1959.
[7] Clower R, The Keynesian Counter-revolution: a Theoretical Appraisal, in Brechling F. e Hahn F., «The Theory of Interest Rates», London, Macmillan, 1965.
[8] Clower R., Reconsideration of the Microfoundations of Monetary Theory, «Western Econornic Journal», 1967, n. 6.
[9] Clower R., Theoretical Foundations of Monetary Policy, in CLAYT0N G., Gilbert J., Sedwick R., «Monetary Theory end Monetary Policy in the 1970 s», Oxford University Press, 1971.
[10] Greniewski H., Kybernetisch - oeknomische Modelle, in «Matematik und Kybernetik in der Oekonomie», Berlin, Akademie Verlag, 19.65.
[11] Hines H. G., On the Reappraisal of the Keynesian Economics, Londra, Martin-Robertson Ltd., 1971.
[12] Hurwicz L., Optimality and Information Efficiency in Resource Allocation, in Arrow K., Karlin S., Suppes P., «Mathematical Methods in the Social Sciences», Stanford University Press, 1960.
[13] Keynes J. M., The General Theory of Employment, Interest and Money, Macmillan, 1961.
[14] Kornai J., Antiequilibrium, North Holland, 1971.
[15] Kuhn T., La struttura delle Rivoluzioni Scientifiche, Torino, Einaudi, 1969.
[16] Lange O., On the Economic Theory of Socialism, Minneapolis, Lippincott, 1938.
[17] Lange O., Introduction to Economic Cybernetics, Oxford, Pergamon Press, 1970.
[18] Leijonhufvud A., On Keynesian Economics and the Economics of Keynes, Oxford, University Press, 1968.
[19] Leijonhufvud A., Keynes and the Classics, London, I.E.A., Occasional Paper, 1969.
[20] Milsum J., ed, Positive Feedbacks, Oxford Pergamon Press, 1968.
[21] Negishi T., General Equdibrium Theory and International Trade, North Holland, 1972.
[22] Patinkin D., Money, Interest and Prices, 2nd ed., New York, Harper and Row, 1965.
[23] Phillips A. W., Stabilization Policy in a Closed Economy, «Economic Journal», 1954.
[24] Phillips A. W., Stabilisation Policy and the Time-form of Lagged Responses, «Economic Journal», 1957.
[25] Phillips A. W., La Cybernetique et le Controle des Systemes Economiques, Cahiers de l'ISEA, Paris, 1958.
[26] Samuelson P., Foundations of Economic Analysis, Cambridge Mass., 1947.
[27] Theil H., Economics and Information Theory, North Holland, 1967.
[28] Tustin A., The Mechanism of Economic System, London, Heinemann, 1953.
[29] Von Bertalanffy L., Teoria Generale dei Sistemi, Milano, ILI, 1971.
[30] Walras L., Eléments d'Economie Politique pure, Paris, Pinchon, 1926.
[31] Wiener L., The Human Use of Human Beings, 1953.

Federico Arcelli **Giorgio Di Giorgio**

KEYNESIAN THEORY AND NEO-QUANTITY ANALYSIS: TOWARDS A SYNTHESIS [35]

Foreword

A comparison of the tenets expressed by the Keynesian economists, with explanations of reality, including suggestions made by the followers of the Chicago school regarding a rational approach to economic policy, reveals many different positions today that cannot easily be reduced to a theoretical unit. Normal economic science, that expounded synthetically in handbooks, "which the scientific community recognizes in a certain period as able to provide the fundamentals of its praxis"[36], appears to have reached a critical point. Perhaps, as Hines[37] remarks, time has come for a new synthesis. This premise makes it advisable to re-examine the meaning, today and in historical perspective, of a comparison between Keynesian theory and «neo-quantity» analysis.

On the one hand Keynesian doctrine, as time passes, has successively laid stress on different features of its analysis, so that it becomes increasingly difficult to give a full definition of what truly forms the corpus of this doctrine today. On the other hand quantity theory has absorbed some notions of Keynesian inspiration[38] as it continued and still continues to be re-elaborated; as a result more softly shaded positions have recently been added to the original sharp contrapositions (maintained as a point of honour).

All this explains how it is that the debate between Keynesians and monetarists often involves different levels and perspectives of analysis. But this is not all: a new approach to the original Keynesian thought states that many assertions concerning the presumed ineffectiveness and unimportance of monetary policy are completely extraneous to Keynes' teaching and should therefore be repudiated. This being without prejudice to some fundamental differences with respect to the monetarists.

The object of this paper, therefore, is on the one hand to focus the comparative survey of the two schools positions correctly, both in historical perspective and as regards the different levels of analysis, and on the other hand to point out, in addition to the lines of convergence, the reasons why it is deemed necessary to step beyond both positions owing to the extension of the disaggregate aspects of analysis. Briefly there are at least three different perspectives in accordance with which a comparative survey can be made.

A first approach consists of the traditional comparison between Keynesian and quantity theory through the contraposition of monetary to fiscal policy. This is the kind of comparison historically underlying the identification of the two schools and allowing some empirical tests of the different contentions to be made.

A second approach, at a more sophisticated and topical level, shifts the antithesis between the followers of the two schools to the different value attached to the effectiveness of monetary policy, in the sense of the control of the quantity of money, as compared with a policy pursued to control the pattern of interest rates. The acceptance of the portfolio approach by both sides has swollen the original terms of the controversy so that today the contraposition of fiscal to monetary policy is accompanied by the controversy as to which is the best indicator of monetary policy in the broadest sense: the rate at which the quantity of money changes, or the level of interest rates. As it is known, the need to control the structure of interest rates has been strongly supported by the Radcliffe Report, the conclusions reached by which met above all with the approval of the Keynesian economists. But the same philosophy, expressed with greater theoretical strictness, may also be found in the writings of some well-known Yale economists, foremost among whom is Tobin. The dispute has thus expanded considerably parallelly with the progress of economic theory and particularly with the development of the notion of liquidity and the consideration of the complex system of credit and financial instruments and securities negotiated on the money and financial markets.

Lastly, the third perspective of comparison between the two doctrines, refers to the present reappraisal of the original Keynesian thought, which often differs from the later elaboration of the doctrine, as it is pointed out in

---

[35] From: Rivista di Politica Economica, Selected Papers, No. 6, 1972.

[36] See T. KUHN: La struttura delle rivolazioni scientifiche, Turin, 1969, p. 29.

[37] See A.G. HINES: On the Reappraisal of Keynesian Economics, London, Martin Robertson Co. Ltd., 1971, p. 62.

[38] See M. FRIEDMAN: A Monetary Theory of Nominal Incomes in "Journal of Political Economy", March 1971, pp. 323-37.

particular by Clower, Leijonhufvud and Hines[39] A meaningful comparison can therefore be made between the true model (or models) implicit in Keynes' analysis and what is asserted by the monetarists.

This last approach is particularly fruitful both because, following Leijonhufvud, we shall consider an interpretation of Keynes' thought, rooted in the Walrasian general equilibrium framework, from which Keynes departs rejecting the hypothesis of full information, and because it is possible to take up again the theory of prices, neglected by the Keynesians and not gone into throughly by the monetarists. And lastly because the analysis of the models implicit in Keynes will allow an interesting comparison of the positions of the two schools with reference to the present problems of cost-push inflation. We shall therefore deal successively with these different perspectives, concentrating our attention above all on the last approach, which provides matter for new considerations, which should enable the traditional positions of both schools to be overstepped.

The Keynesian Revolution in Historical Perspective

As it is known the «Keynesian revolution» contributed to demolishing the conceptual scheme of the old quantity theory, revealing its tautological character or alternately pointing to the lack of realism of the hypothesis of full employment, which generally accompanied its formulation, and stressing that the velocity of money circulation is marked by instability[40]. The Great Depression seemed to make whatever judgment regarding its credibility superfluous.

However considerable energy was spent, in the elaboration of the later Keynesian doctrine, on demonstrating the irrelevance of money and the inefficiency of monetary policy interventions designed to reanimate an economy afflicted by unemployment. From the theoretical point of view the existence of possible «liquidity traps» was stressed to explain the uselessness of expanding the supply of money. The unforeseeable and unstable nature of the demand for money was claimed as a reason that made it irrational to place any trust in monetary intervention. From the empirical point of view a famous survey carried out by Oxford economists seemed to demonstrate the rigidity of investments with respect to changes in the interest rate; this deprived the utility of monetary action of all practical meaning.

Since consumption too seemed to depend on income rather than on the level of the interest rate, fiscal policy was identified as the only possibility of intervening in the economy. This position was the outcome partly of an incorrect interpretation of-Keynesian thought, as will be seen farther on, and partly of objective historical requirements. In fact the post-war period, with a huge public debt to manage and the fear of serious depressions as a result of the difficulty of reconverting a wartime industry, made it quite natural for the United States and Britain to accept a cheap-money policy, which meant confining the possibility of manipulating and regulating the economy to the fiscal instrument.

Whatever the reason was (theoretical or empirical), the period immediately following the Second World War witnessed the affirmation of the Keynesian doctrine supporting fiscal policy while the quantity theory was completely discredited. In a recent paper on monetary theory, Johnson says: «The revival of a quantity theory that could claim to rival with the Keynesian theory required a restatement of it that would free it from the Keynesians' objections and give it an empirical content».

From the historical point of view, circumstances favourable to the revival of monetarism occurred when generalized inflationary pressures became evident instead of the deflation that had been feared, while fiscal policy was found inadequate as an instrument for stabilizing the economy.

The debate on the Pigou effect and the so-called neoclassical synthesis helped to further redimension the Keynesian revolution in a more traditional frame and thus prepared, also at the theoretical level, circumstances favourable to the advent of the so-called monetarist counterrevolution. Soon it will be high tide again for quantity theory and Friedman will be the leader of new monetarism. Friedman proposes a restatement of the quantity theory in terms of a theory of money demand, placing at the centre of his construction the existence of a stable functional relationship between the real demand for money and a limited number of variables, among which permanent income must be considered of particular importance.

Friedman's Money Demand and the Portfolio Approach

---

39 See R. CLOWER: *The Keynesian Counter- Revolution: a Theoretical Appraisal* in F. BRECHLING and F. Hahn, "The Theory of Interest Rates", London, Macmillan, 1965; R. CLOWER: *Reconsideration of the Micro Foundations of Monetary Theory* in "Western Economic Journal", 1967 N° 6; A, LEIJONHUPVUD: *On Keynesian Economics and the Economics of Keynes*, Oxford University Press, 1968; A.G. Hines: *On the Reappraisal of Keynesian Economics*, op. Cit.

40 In a recent essay Patinkin (see D. PATINKIN: Sulla non neutralità a breve termine della moneta nella teoria quantitativa in "Moneta e Credito", March 1972) contests the foundedness of the criticism as far as the short term is concerned, demonstrating that the current interpretation of the quantity theory differs from the original doctrine, which is free from the objections made.

Friedman explains[41] that his quantity theory is primarily a theory of the demand for money, not a theory of money income or of the level of prices. For the economic transactor money is an asset comparable with other assets, one of the many ways of holding wealth. The quantity of money demanded results from the theory of choice based on a scheme similar to that underlying the theory of the consumer: it depends in fact on the preferences of the transactors, on the alternative prices and returns of the different forms of investment as well as on the total wealth to be held in the various forms (the equivalent of the budget constraint for the consumer). In his analysis, Friedman states that money that appears as a distinct variable in the utility function provides, like other assets, productive services and non-pecuniary services identifiable in the feeling of security and the pride of possession. But, as Hahn[42] pointed out recently, Friedman does not specify the parameters of the utility function, but confines himself to introducing the notion of the marginal utility of money in very vague terms.

However, leaving aside the imprecise definition of the neo-classical scheme underlying the neo-quantity theory, we can say that «what Friedman has really presented is an elegant exposition of the modern portfolio approach to the demand for money which... can only be seen as a continuation of the Keynesian theory of liquidity preference»[43]. In effect, «contemporary monetary theorists, whether avowedly «Keynesian» or «quantity», approach the demand for money essentially in the same way, as an application of the general theory of choice, though the former tend to formulate their analysis in terms of the demand for money as an asset alternative to other assets, and the latter in terms of demand for the services of money as a good»[44]. The kernel of the contraposition between modern Keynesians and monetarists lies therefore with Friedman's hypothesis that the demand for money is a stable relation and reflects the particular definition he gives of money and its substitution range with regard to other monetary and financial instruments and real assets.

Empirical Experience as Test Bench for the Two Theories

The assumption of a stable relationship provides a very definite picture, subject to empirical testing and allows a series of tests comparable with the tests supporting the Keynesian theory. The effort made by Friedman and his school is recognizable in the statement of simple uniformities incorporating the nucleus of the two theories and easily empirically tested. Friedman obtains from the demand for money a functional relationship between quantity of money and income which allows him to identify a kind of monetary multiplier, specified by the function of velocity, to oppose to the Keynesian multiplier linking autonomous expenditure to increases in income. This makes a comparison between the effectiveness of monetary and fiscal policy possible.

On the other hand, by evaluating the sensibility of the demand for money to changes in the interest rate, it is possible to ascertain whether or not money can be easily substituted by other financial assets. It thus becomes possible to obtain a clear indication of the importance of controlling the quantity of money as compared with controlling the pattern of interest rates. These are the essential points of the dispute between monetarists and Keynesians, which has witnessed the shifting of the debate from the field of theory to that of econometric analysis in the attempt to affirm the superiority of one interpretation of economic reality over the other. At this point it must be mentioned that the tests suggested and carried out are based on some methodological principles, an understanding of which is essential in order to appreciate the intellectual energies exerted and to appraise the results achieved, accepting or rejecting them as the case may be[45].

The crucial point accepted by Friedman is that the validity of a good theory is measured by its capacity of foreseeing very important phenomena through the medium of a simple and stable theoretical relationship. The essence of the "Quantity" theory is thus specified in the relation linking money to income, whereas the Keynesian theory is individualised by the uniformity of the multiplier. An accurate estimate of the parameters, achieved by identifying a complex model of general equilibrium, is therefore renounced in favour of an easy and synthetic empirical test[46].

---

41 See M. FRIEDMAN: *The Quantity Theory of Money: a Restatement* in the "Optimum Quantity of Money", London, Macmillan, 1969, p. 52.

42 See F.M. HAHN: P*rofessor Friedman's Views on Money* in "Economics" Feb. 1971, p. 63 ff.

43 See D. PATINKIN: *The Chicago Tradition. The Quantity Theory and Friedman* in "Journal of Money, Credit and Banking", Feb. 1969, pp. 46-70.

44 See H.G. JOHNSON: *Monetary Theory and Policy* in "Essays in Monetary Economics", London, Allen & Unwin, 1967, p. 26.

45 See H.G. JOHNSON: *Recent Developments in Monetary Theory* in "Money in Britain 1959-1969", Oxford University Press, 1970, p. 86-87.

46 Recently Laidler starting his argument from a macroeconomic model of the Hicks-Hansen kind from which he derives both the equation of the Keynesian multiplier and the money velocity function according to Friedman's theory, contests the validity of the proposed econometric tests aiming at affirming the superiority of one theory over the other. See D. LAIDLER: *The Influence of Money on Economic Activity* in Clayton, Gilbert, Sedgwick: "Monetary Theory and Monetary Policy in the 1970's" Oxford University Press, 1971.

The second principle implicitly accepted by Friedman is that behavioural relations are unaffected by historical and institutional changes: this explains the interpolation of the relations on very lengthy time-series. It is clear that both these principles are somewhat debatable at the methodological level. There is in fact some foundation for the objection that there is no justification for the transposition of rules acceptable for the physical world, where the protopostulate of the invariance of reality holds, to a social reality characterized by continuous changes and complex interdependences. A full appraisal of the significance of the tests supplied by Friedman in support of this thesis calls in any case for a more thorough examination of the concepts and of the relations among variables implicit in the two theories compared with each other. In particular it is important to give a definition of money in order to define the mechanisms transmitting monetary impulses to real variables. Among other things this calls for the consideration of relations among the several financial assets and between the latter and real assets.

Substitution Between Financial and Real Assets in the Keynesian Theory and in the "Quantity" View

Whereas recent Keynesian literature stresses the close relations of substitution generally existing among financial assets so that there is an intrinsic difficulty of defining a particular set of assets which possesses exclusive features allowing the stock of money to be identified; for the monetarists on the contrary money is a «unicum» not easily substituted by other financial instruments and therefore clearly identifiable. Friedman considers the stock of money to include currency and demand and time deposits of commercial banks. For the Chicago school, however, money – which is therefore not a close substitute for other financial assets – includes also real assets in its spectrum of substitution.

The Keynesians, on the contrary, maintain that whereas financial assets are close substitutes for money (and consequently give no precise definition of the stock of money, replacing it with the notion of liquidity), only a very tenuous and remote relation exists between money and real assets. For the Keynesians, therefore, the crucial distinction is between the financial and real sectors of the economy and not between banking and the rest of the economy as it is for the monetarists.

However they too, though for different reasons, share the opinion that compared with the analysis of «*A Treatise on Money*», the definition of liquidity preference given in the «*General Theory*» oversimplifies the problems connected with the determination of the equilibrium of money and financial markets. In modern economics the alternative is not only between money and long-term bonds, however they are defined. The development of financial intermediaries and the correlative range of financial liabilities has lent greater scope to the money and stock markets. Keynes' schematism has thus been superseded by the introduction of a complex network of relations between the financial and the real system[47].

Both currents thus converge in the modern portfolio approach, but owing to the different hypotheses adopted concerning the relations of substitutability existing among the various financial assets and between the latter and real assets, the conclusions they reach regarding the kind of monetary policy to be pursued, the effectiveness of the alternative measures and the mechanisms transmitting the impulses are different. «The transmission mechanism whereby monetary influences affect decisions to spend generally, will be determined by the way in which people adjust their equilibrium portfolio of assets in response to a disturbance initiated, for example, by the intervention of the authorities in financial markets. These reactions and therefore the transmission mechanism will depend on which assets people view as particularly close substitutes for money balances»[48].

What has been said so far makes it possible to affirm that for the Keynesians an increase in the quantity of money following an open market transaction will produce a new portfolio equilibrium chiefly through the purchase of financial assets substituting money. As a secondary effect there will be a reduction in the yield of these liquid assets and this will lead transactors to extend their purchases to other less liquid securities, with a tendency to spread them over the whole spectrum of financial assets while the reduction of interest rates diffuses. Should the yield of long-term securities change too, the difference between the rate of interest on these securities and the marginal efficiency of capital might bring about a change with regard to decisions to invest in real assets.

For the Keynesians, therefore, the mechanism transmitting a monetary impulse to the real variables is identified chiefly with the change in the rate of interest on long-term securities, the implication being that monetary policy could be pursued more effectively by acting on interest rates rather than by changing the quantity of money[49]. The Keynesians hold the effects of monetary policy to be very slight in any case, firstly because the close relation of substitution existing between money and financial assets allows the portfolio to be adjusted with slight changes in interest rates, also in the presence of conspicuous changes in the quantity of

---

47 See M. ARCELLI: *Il controllo diretto degli intermediari finanziari: una controversia scientifica in via di superamento* in "Giornale degli economisti", Sept.-Oct. 1969, p. 7.

48 See C.A.E. GOODHART and A.D. CROCKETT: *The Irnpartance of Money* in the "Quarterly Bulletin of the Bank of England", 1970, p. 160.

49 See C.A.E. GOODHART and A.D. CROCKETT, *op. cit.*, p. 161.

money or of any other financial assets secondly because spending decisions are held to be comparatively insensitive to changes in interest rates.

For the monetarists, on the contrary, the portfolio is adjusted by buying either financial assets or goods, as real assets are also included in the spectrum of the direct substitution of money. The impact of changes in the stock of money therefore has a more general effect than is assumed by the Keynesians, operating diffusely and not only through the change in a particular category of interest rates.

Whereas for the Keynesians the effect of monetary policy on spending decisions in real assets acts indirectly and reflects almost exclusively on investments (very weakly moreover, it is presumed), for the monetarists the effect is distributed directly over the whole range of assets, both financial and real. For the followers of the Chicago school, in fact, the open market transaction alters the ratio between the stock of money and the other assets of which wealth is composed. The new equilibrium will therefore be established through a redistribution of the surplus (or deficit) of monetary balances over all assets, both financial and real, so as to re-establish a satisfactory ratio between the stock of money and other assets. But this is not all: «according to this logic, which views money as an alternative to any asset in the whole spectrum of forms in which wealth can be held, money itself must be considered a component of total wealth.

Therefore changes in the real supply of money do not necessarily entail portfolio adjustments only, but may be directly reflected by changes in consumption, thus influencing the system real demand»[50]. The effectiveness of monetary policy therefore assumes an immediacy unknown to the Keynesians. This does not however exclude the presence of indirect mechanisms transmitting the effects. In fact, since for the monetarists financial assets are not close substitutes of money, changes in the yield rates are substantial and through them the effectiveness of monetary policy is observed again. It is consequently unnecessary to act on the structure of interest rates, as the Keynesians sustain, since the effects of changes in the stock of money on spending decisions are reflected also by changes in the relative yields of financial assets.

It must be remarked, however, that for the monetarists (in contrast with the Keynesians) no single interest rate exists that clearly indicates the intensity and direction of an intervention of monetary policy, because the effects of the monetary impulse are pervasive and may affect the various financial assets differently as a consequence of particular preferences for one security or another[51]; moreover the neo-quantity theory explicitly introduces and stresses the importance of expectations concerning the level of prices as an element of the cost of holding money. It ensues that it may be very misleading to interpret changes in the interest rates on financial assets as indications of changes in the trend of monetary policy, without taking into proper consideration the changes in the relation between money rates and real rates produced by different expectations regarding the level of prices[52].

The monetarists therefore attribute an essential role to changes in the stock of money in determining monetary policy trends, while they deny the existence of a rational foundation for the need to control the structure of interest rates. This thesis, based indirectly on the assumption that the demand for money can be foreseen, is linked to the monetarists' particular conception of the substitutability of the various financial assets and real assets. To sum up, the monetarists assume that interrelations in the financial sector are steady enough to allow changes in the quantity of money to be used to forecast effects on the real sector. The stability of the demand for money does not however imply the elasticity of money with respect to the interest rate to be zero, though it points out that the possibility of using the monetary multiplier (i. e. velocity function) operationally implies very low values of this elasticity.

In order to dissipate any misinterpretation, Friedman makes the following statement: «I know no empirical student of the demand for money who denies that interest rates affect the real quantity of money demanded – though other have misinterpreted me as so asserting»[53]. The hypothesis of the constancy of the velocity of money is actually typical of the current interpretation of the original version of the quantity theory. What is instead crucial for Friedman is that the demand for money is a stable function of a limited number of variables, not that no correlation exists between this magnitude and the interest rate.

Once again, therefore, the validity of the monetarists' contentions is brought back to the comparison with experience. The Keynesian function of liquidity preference should it prove stable, would not be incompatible with the assertions of the monetarists, ruling out the tract referring to the liquidity trap. Should empirical experience support the claim that the elasticity of the demand for money is low (less than unity) everywhere, the neoquantity theory would as a result be substantially strengthened. This is what Friedman has attempted to demonstrate through a number of empirical researches.

---

50 See R. MASERA: Recenti sviluppi nella politica dei saggi di interesse in Italia analizzati in una prospettiva internazionale in "La politica economica a breve termine", Rome, Dec. 1969, p. 63 (mimeograph edition of the National Council for Scientific Research).

51 See C.A.E. GOODHART and A.D. CROCKETT: *The Importance of Money, op. cit.*, P. 165.

52 See H.G. JOHNSON: *Recent Developrnents in Monetary Theory, op. cit.*, p. 89.

53 See M. FRIEDMAN: *Interest Rates and the Demand for Money* in "The Optimum Quantity of Money", *op. cit.*, p. 142.

The comparison between monetarists and Keynesians thus comes to be based on two tests: on the one hand the monetary multiplier (velocity function) is compared with the Keynesian multiplier in the explanation of the income dynamics; on the other hand an attempt is made to discover whether money is really a «unicum» or whether it is easily substituted by other financial assets, through the determination of the elasticity of demand to the interest rate or rates. A high elasticity would bear up the Keynesian assertions as to the necessity of controlling interest rates rather than the quantity of money, whereas the discovery of a substantial inelasticity of the demand for money would support the neo-quantity thesis.

This second empirical inquiry should logically precede the comparison between the multipliers. In fact there would not be much sense in postulating the existence of a monetary multiplier if the elasticity of money to interest rate were substantial. It may also be observed that neither of the tests is independent of the definition of the variables included in the relationships. Thus Friedman's particular definition of money generally reduces the effect of interest rates on the demand for money and confers stability on the monetary multiplier. Similarly the explanatory power of the Keynesian multiplier is influenced by the definition of autonomous expenditure assumed by the different researchers.

It is not however the purpose of this paper to go into a detailed discussion of the innumerable empirical tests used in favour of or against the theories in question. Rather, what may be interesting are some general considerations regarding the areas of agreement and the doubts given rise to by the tests. There is no lack of people in fact, who point out substantial weaknesses in the specification of the equations proposed by the monetarists for the test. As it has been seen, the monetarists obtain from the equation of the demand for money a relationship between stock of money and level of income: this second equation is assumed to identify the so-called monetary multiplier. In the first equation income and interest rate are considered exogenous variables, whereas the real demand for money individuates the endogenous variable. In the second equation, on the contrary, income is considered a dependent variable, while money appears as an exogenous variable.

Now it is rather strange, Hines[54] points out, that the stock of money should identify the demand for money in the first equation whereas in the second equation, derived from the first one, it represents the supply of money. In effect the equations proposed do not appear to be adequately identified. To suppose, for example, that income depends on money, as implied by the monetary multiplier, and not also conversely, is bound to give rise to a number of doubts. Extending the analysis it will be seen that the income equation does not provide a satisfactory explanation of the level of money income. «For we are not given a model which is distinct from that of Keynes or the Keynesians and which includes an explicit mechanism whereby the effects of a change in the stock of money are transmitted to the level of income. We are simply told that, contrary to what was assumed by Keynes the empirical evidence shows that the elasticity of substitution between bonds and money is low whereas that between goods and money is high...

The results do not prove that changes in the money stock rather than changes in autonomous expenditure are the cause of changes in income. In a monetary economy, in order to affect an increase in planned expenditure, there must usually be a prior increase in the demand for money... If the supply of money is increased to meet demand, we would then observe that an increase in money balances precedes the increase in income, even if the latter is the result of an increase in autonomous expenditure»[55].

In short, if the proposition affirming that the stock of money allows income to be foreseen is to have any analytical meaning it is necessary not only for the supply of money to be truly exogenous and unambiguously definable, but also for this relation to emerge clearly as an equation in reduced form of some underlying model, the structure of which remains unchanged through time[56]. Friedman seeks to justify his results with the consideration of time lags among the variables, but his arguments do not remove the perplexities.

A similar but symmetrical discourse could be made concerning the equation of the demand for money, with very serious logical consequences for the rules proposed by the monetarists concerning the expansion of money in time. On the other hand the charge often brought against Friedman of having reintroduced the dichotomy between the real and the monetary sector of the economy must be rejected. This is so both because the quantity of money acts also on the real variables and also because the velocity of money is not assumed independent of

---

54 See A.G. HINES: *On the Reappraisal of Keynesian Economics, op. cit.*, p. 32.

55 Ibidem, pp. 33-34. Hines appears to restate the doubts raised by Laidler at the Sheffield meeting regarding the effectiveness of empirical tests founded on such econometric models. Compare with D. LAIDLER: T*he Influence on Money on Economic* Activity, *op. cit.*.

56 FRIEDMAN'S recent paper, *A Monetary Theory of Nominal Income, op. cit.*, which presents a model combining part of Fisher's analysis with an equation of Keynesian inspiration, deriving there from the quantity expression linking money to nominal income, can only be considered a promising attempt to overcome the objections made.

interest rates and contributes together with them to the determination of the equilibrium of the whole system[57]. In effect the underlying model, which starts to take shape as a result of the debate and above all of the empirical researches, appears to accept points from both doctrines (Keynesian and monetarist), thus individuating a possible area of agreement.

Exposing these first conclusions does not of course mean ignoring the fundamental differences between the two lines of thought: on these aspects, however, I shall speak in the last part of this study when referring to the present problems of cost-push inflation. Here I shall merely point out that the many empirical tests made, appear to reject the extreme positions of both theories, both that of the elasticity of the demand for money to the rate of interest being zero and that of its being very high in some parts of the function.

The relative inelasticity of the demand for money, recognizable from experience, undoubtedly reinstates the use of monetary policy. However there is an important point that the empirical inquiries have not yet made sufficiently clear: within exactly what limits the ascertained inelasticity of the demand for money depends on a limited substitutability with other financial assets rather than on expectations. If, for instance, quotations were to fall steadily on the stock market, then despite of increasing rates of yield pushed up by sales, part of the supply of money – rather than flowing to securities – might be channelled to satisfy the growing liquidity preference, irrespective of the degree of substitutability between money and financial assets[58]. Similarly the instability of the demand for money caused by sharply fluctuating expectations could determine a portfolio equilibrium similar to that which would be obtained in the presence of a demand for money insensible to the rates of yield of financial assets. But, whereas in the latter case the action taken to modify the size of the stock of money would prove fully effective, the instability of expectations might reassert the need for the control of interest rates as an alternative to acting on the quantity of money.

This doubt, together with the consideration of the many channels through which changes in the quantity of money take place[59] and the insufficient knowledge of the relationship between money rates and real rates of interest with respect to expectations, make it difficult to observe a clearly identifiable mechanism transmitting the impacts of monetary variables to real variables. However the two doctrines agree that monetary policy pursued through open market transactions will generally have an appreciable initial effect on conditions in the financial markets. Decisions to spend money on real assets will tend to re-establish portfolio equilibrium by levelling marginal yield rates, or better, the weighted marginal utilities of financial and real assets. Only in this sense is there agreement concerning the claim that monetary impulses are transmitted to real variables by interest rates. Otherwise, for the Keynesians the monetary impulse will act on real variables only if it gives rise to an autonomous expenditure, which sets the multiplier mechanism in motion.

The monetarists, on the contrary, are inclined to redimension this mechanism, setting changes in consumption in relation to the real supply of money and stressing once again the relevance of portfolio adjustments. As will be seen farther on, the recent reappraisal of Keynes' thought, which allows the consumption function to be linked to Walras' market theory, makes the two theses compatible[60]. Actually both models provide fruitful explicative elements of the reactions that occur in economic systems. This recognition is strengthened by the awareness that all attempts to measure the effects on spending taking into account exclusively changes in interest rates, even were they to prove significant, would not fully reflect the effectiveness of monetary policy as they would not adequately consider price expectations and the effects on consumption.

On the other hand, if changes in the quantity of money were to be directly connected with changes in the level of income, there would be a tendency to overrate the force of monetary policy, as the supply of money would be assumed to be completely independent[61]. What emerges from the above considerations is an interpretation of reality which, avoiding dogmatic positions regarding the effectiveness of changes in the quantity of money, stresses the complexity and possible variety of the mechanisms transmitting monetary impulses to real variables. This confrontation between monetarists and Keynesians would be incomplete, however, if it failed to include a critical consideration of the ultimate meaning of the theoretical schemes identifying the logic of the two currents of thought. This reflection is facilitated by the efforts at present being made to interpret Keynes' original

---

57 See in this connection FRIEDMAN'S exhaustive reply in his essays *Interest Rates and the Demand for Money* in "The Optimum Quantity of Money", *op. cit.*, pp. 141-155. On p. 152, in particular, Friedman explicitly affirms the operation of the Pigou effect.

58 A recent study by the Bank of England points to the introduction of suitable time lags among the variables as a possible way of solving the doubt. Further the demand for money should be disaggregated, at least making a distinction between the household and the firm as operators. See L.D. PRICE: *The Demand for Money in the U.K: a Further Investigation* in "Quarterly Bulletin", March 1972, pp. 434.

59 Among other things the opening of the economy makes it necessary to take into account the financing of any balance of payments deficit as the latter affect the changes in the quantity of money.

60 See A.G. HINES: *On the Reappraisal of Keynesian Economics, op. cit.*, p. 28. See also Leijonhufvud's penetrating analysis (A. LEIJONHUFVUD: *Keynes and the Classics* I.E.A., Occasional Papers, 1969) who explicitly reconciles Keynes' multiplier mechanism with modern theory of consumption.

61 See C.A.E. GOODHART and A.D. CROCKETT: *The Importance of Money, op. cit.*, p. 179.

thought. With the help of the contributions made by various authors, it is thus possible to underline the fundamental differences between the two schools, and to get into focus their different philosophies regarding one of the foremost problems of contemporary economies: inflation.

Reappraisal of Keynes' Thought and Comparison with Friedman's Analysis

Whereas Friedman's approach has been labelled Marshallian by Laidler from a methodological point of view, in the sense that the empirical tests he proposes and his suggestions regarding economic policy are founded on partial equilibrium models, the theoretical reference of his analysis, on the other hand, is clearly identifiable in a Walrasian general economic equilibrium model capable, however, of containing the classifications of a portfolio approach. Actually Friedman refers to a Walrasian scheme in his essay «The Optimum Quantity of Money», with which the collection of his old monetary writings opens. For the purpose of outlining the theoretical structure underlying his analysis, in a more recent paper, Friedman explicitly expounds an aggregate model (not unlike the Keynesian one) with a link however to the Walrasian general economic equilibrium equations[62].

This reference framework is not without logical difficulties as far as the role of money is concerned, as Hahn acutely observes[63], and this quite apart from the consideration of the fundamental problems of a monetary economy brought to light by a modern appraisal of Keynes' original thought. For these considerations I refer to the contributions of Clower, Leijonhufvud and Hines[64] who, re-interpreting Keynes' work along the lines of Walras' general theory of markets (though discarding the hypothesis of full information), have made possible a more meditated comparison with the monetarists' analysis. The theoretical fruitfulness of this reappraisal consists in having connected the Keynesian theory with microeconomic analysis, in having supplied a theoretical explanation of unemployment without needing to resort to empirical assumptions such as wage rigidity of to restrictive assumptions such as the liquidity trap, as well as in having defined some fundamental distinctive characteristics of a monetary economy as compared with a barter economy; Moreover this new perspective allows the extension of Keynes' theory in the direction required by the portfolio approach.

Keynes' position, as it emerges from this revision, allows the affirmation to be made that not only the interest rate, but also relative prices play a more important role in his theory than is assigned to them by the post-Keynesian tradition. In this sense it is advisable to distinguish Keynes' theory from economics handed down by the Keynesians. According to Leijonhufvud «the revolutionary impact of Keynesian economics on contemporary thought stemmed in the main from Keynes' reversal of the conventional ranking of price and quantity velocities. In the Keynesian models price velocities are not infinite. It is sometimes said that the implications of the model result from the assumption that money wages are rigid. This usage can be misleading. Income-constrained processes result not only when price-level velocity is zero, but whenever it is short of infinite»[65]. What distinguishes a Walrasian model from a Keynesian scheme is therefore the different role of the variables in the adjustment process of an economy out of equilibrium.

In a Walrasian system the perfect flexibility (instantaneous velocity of adjustment) of prices ensures the achievement of full employment equilibrium. In Keynes' model, on the contrary, adjustment proceeds faster in terms of quantity rather than of prices (as a consequence also of inelastic price expectations) and this may lead to resources not being fully employed. According to another perspective, the difference in the functioning of the Keynesian economy is amenable to a different information structure, which may be analysed by modern cybernetics.

If we consider that, in an economy with decentralized decisions, prices perform two functions simultaneously, that is diffusing the information necessary for the coordination of the plans of economic actors and providing incentives for the efficient allocation of resources among the different activities, we are able to trace the distinction between Keynesian and Walrasian economics underlining the fact that the former does not assume perfect information such as is assumed by the mechanism of the auctioneer and the groping search (*tâtonnement*) postulated by Walras.

---

62 See M. FRIEDMAN: *A Theoretical Framework for Monetary Analysis* in «Journal of Political Economy», March 1970, pp. 217-20 and p. 223. The last part of the paper is devoted to illustrate the transition from the short run to equilibrium in the long run.

63 See F.H. HAHN: *Professor Friedman's Views on Money*, op. cit., pp. 69-72.

64 Works quoted in the foreword.

65 See A. LEUONHUFVUD: *On Keynesian Economics and the Economics of Keynes, op. cit.*, p. 67. This conclusion, namely that involuntary unemployment does not necessary depend on wage rigidity, had in any case already been reached by Patinkin. (See D. Patinkin: *Money, Interes and Prices*, New York, Harper and Row, 1965, p. 340). The original aspects of Clower's and Leijonhufvud`s approach consists in their having established a link between Keynes' analysis and Walras' theory of markets in more general terms.

If the market clearing vector of relative prices (which is assumed to exist) is known by all transactors and if each transactor adjusts to these parameters, the system simultaneously generates an optimal allocation of resources and ensures that they are fully employed»[66], but if this is not the case a state of unemployment of the Keynesian kind may ensue. Keynes' world is therefore consistent with the neo-classical assumptions of rational behaviour of economic actors and with the logic of scarcity linked to the system of price incentives. However the assumption of imperfect information is a departure from this universe.

All states of unemployment are therefore due to a vector of relative prices departing from the market-clearing vector and to the impossibility of making the necessary corrections before erroneous transactions take place. So, as Clower points out, a whole series of transactions, which would have been performed at equilibrium prices, are no longer carried out. The chief consequence is the dichotomy established in a monetary economy between potential and effective demand, the latter being constrained by monetary income obtained through realized sales. In equilibrium the two magnitudes coincide, but out of this condition the functioning of the economy is constrained by effective demand.

Clower, in close harmony with Keynes' theory, considers the specific example of a situation of disequilibrium on the labour market. The current income of the families supplying labour services is determined not by the quantity of services they would like to supply at market prices, but by the quantities they actually succeed in selling. It follows that their effective demand on other markets is limited by income actually realized. Entrepreneurs will in turn curtail their purchases on related markets and on the labour market since they link their decisions with realized receipts, which are smaller than expected. A process of cumulative contraction is thus set in motion, which brings the system into a situation of temporary equilibrium with unemployment[67]. The lack of perfect price flexibility, which is the mirror of sub optimal information, brings out disequilibrium situations and consequently unemployment of resources. This appears quite natural in a monetary economy like the Keynesian in which the amount of information is necessarily sub optimal.

Keynes' central question regards in fact the analysis of the economic behaviour of the present under the influence of changing ideas about the future: now the diffusion of adequate information concerning the inter-temporal price vector appropriate to dimension savings and investments correctly appears to be impossible. Instead of a vector of this kind, the transactor conjectures about the marginal efficiency of capital and possesses a liquidity preference, which guides him in his portfolio choice. If plans of the economic actors are to be consistent, the marginal efficiency of capital and the vector of relative prices must correctly anticipate future prices and quantities in commodities and financial markets[68]. But this proves to be absurd owing to continuous revisions of expectations and in the absence of adequate forward markets able to substitute the function of the Walrasian auctioneer and the groping search.

It is true that Keynes mentions the stock exchange market in his General Theory, but he recognizes immediately the distortions that occur in it owing to speculation.

Unemployment is therefore not necessarily the result of wage rigidity or liquidity traps, which affect empirical assumption but rather of the lack of information inherent in the logical structure of the model. The transition from the timeless Walrasian world, in which knowledge of the vector of clearing market prices is perfect, to the Keynesian world, where expectations rule, requires a number of invariant properties of the model to be discarded and the attribution of a special status to money. The reappraisal of Keynesian thought then allots a prominent role to the theory of prices, contrary to the neglect of this subject shown by post-Keynesian analysis.

A double analytical line is thus recognizable in Keynesian thought. On the one hand, the familiar macroeconomic model (income-expenditure model) popularised by post-Keynesian tradition, which neglects the theory of prices, considers essentially real magnitudes and accepts the assumption of wage rigidity and the presence of monopolistic elements in the system. On the other hand, a model founded theoretically on microeconomics, which attributes a prominent role to relative prices and individuates the cause of unemployment in an erroneous structure of the relations between prices of investment goods, prices of consumer goods and wages. In this second model the effectiveness of the interest rate is also reaffirmed, in contrast with the Keynesian tradition, both because investments are held not to be insensible to its changes (this thesis is consistent

---

66 See A.G. HINES: *On the Reappraisal of Keynesian Economics, ap. cit.*, p. 11. Kornai (see J. KORNAJ: *Anti-equilibrium*, North Holland, 1971), a critic of Walrasian general equilibrium, recently made a deeper study of the information structure of the models, developing new analysis perspectives in the frame of a more general theory of economic systems.

67 Negishi (see T. NEGISHI: *General Equilibrium Theory and lnternational Trade*, North Holland, 1912, p. 208) points out that the discarding of the Walrasian mechanism of adjustment par *tâtonnements,* (groping search), besides inducing a spill over effect, determines redistribution effects originating from disequilibrium transactions. K. J. ARROW AND F.H. HAHN, in *General Competitive Analysis* Holden Day, S. Francisco, 1971, have analysed the two effects jointly. Further progress along the same line of analysis is to be found in an essay by F.H. FISHER *On Price Adjustment Without an Auctioneer* in "The Review of Economic Studies", Jan. 1972.

68 See A.G. HINES: *On the Reappraisal of Keynesian Economics, op. cit.*, p. 15.

with the aggregations of the variables in the General Theory)[69] and because of the existence of the so-called «windfall effect» which is believed to connect the level of consumption inversely with the interest rate through a wealth effect.

If doubts remain concerning the effectiveness of monetary policy, they are due not so much to the fact that the real variables might not be particularly sensible to monetary impulses, as to the fact that, according to Keynes, the long-term interest rate could not easily be changed. While it is thus possible to reject the thesis that money doesn't matter, it should be remarked that the logical structure and the effort of theoretical clarification provided by this reappraisal of Keynes' thought, allow a more far-reaching comparison with the monetarists' propositions. Keynes' contribution to economic theory may therefore be viewed as an effort to extend the use of the tools of the theory of value beyond the area of general economic equilibrium into the field of macroeconomic disequilibria[70]. In situations of disequilibrium, potential demands are not consistent with one another on the whole; but we know from the theory of general economic equilibrium that an appropriate vector of relative prices exists which would allow equilibrium to be re-established. It then matters to find out where market prices depart from the equilibrium vector.

Keynes' opinion is that although disequilibria occur above all on the labour market, the prime cause of the distortion can generally be individuated in too low a level of the prices of financial assets. Therefore it is not usually high wages that cause unemployment, but rather the high long-term interest rate that depresses the prices of assets and consequently investments. This being the case, the remedy for unemployment would not be deflation, as such action would not necessarily alter relative prices, but on the contrary a series of provisions designed to re-establish a physiological relation between wages and the prices of non liquid assets.

Wage Push Inflation in Keynesian Theory and in Monetarists' Analysis

While underlining the fundamental differences with monetarists' views, the reappraisal of Keynes also allows a better understanding of the kind of inflation process caused by wage pushes. Following Keynes' logic, once the trade unions have succeeded in imposing wage increases, the supply of money should be adequately expanded to prevent the unemployment caused by a higher demand for money and a rise in interest rates set in motion by the wage push. Monetary action taken to preserve the level of economic activity would thus play a sustaining role. It is under this aspect «that Keynes' analysis deviates from the monetarist position. The money increment is not a causal factor of inflation; it is an effect of business condition»[71].

The monetary authorities are engaged in tackling the problem of unemployment, but they cannot be held guilty of generating inflation or sharing in its generation[72]. For the Chicago school, on the contrary, inflation is a phenomenon that can be clearly traced back to changes in the quantity of money and is therefore the direct responsibility of the monetary authorities. But the explanation of how this process takes place is not well defined. As a matter of fact Friedman himself explicitly admits that, in the short run, changes in nominal income due to increases in the quantity of money are not easily split up into changes in price and changes in output[73] so that monetarist macro theory appears as separate from microeconomic analysis as from the post-Keynesian income-expenditure model[74]. In the long run however, the theoretical reference of Friedman's arguments relies on a

---

69 Keynes, in fact, aggregates bonds and real assets in the category of non-money assets, separating them from consumer goods, whereas the post-Keynesians separate the financial from the real sector (investment goods and consumer goods). See in this connection the exhaustive treatment in Chapter 3 of A. LEIJOHUFVUD: *0n Keynesian Economics and the Economics of Keynes, op. cit.*

70 See A. LEIJONHUFVUD: *On Keynesian Economics, op. cit.*, p. 33. In this context Keynesian theory must be interpreted dynamically (see *ibidem*, p. 161).

71 See S. WEINTRAUB: *Keynes and the Monetarists* in "The Canadian Journal of Economics", 1971, n° 1, p. 45.

72 The monetarists object, however, that if the effect of expectations on the level of prices is neglected, the monetary authorities' supporting action may become one of the factors accelerating inflationary dynamics.

73 An outline for this analysis is sketched by Friedman in his paper "A Theoretical Frarnework for Monetary Analysis", *op. cit.*, pp. 223-24; but later he gave up dividing the nominal income into real income and prices, elaborating the quantity theory as a theory of nominal income (see M. FRIEDMAN: *A Monetary Theory of Nominal Income* in "Journal of Political Economy", *op. cit.*, March 1971, pp. 323-37).

74 It is no mere chance that Weintraub (see S. WEINTRAUB: *Keynes and the Monetarists, op. cit.*, pp. 41-3), introducing some simplifications, finds a substantial identity between the implications of the model accepted by post-Keynesians, when investments are sensitive to the interest rate, and the monetarists' claims concerning the effects of changes in the quantity of money. Like Leijonhufvud, Weintraub recognizes however in Keynes the existence of another model giving ample relief to the theory of relative prices.

neoclassical model of general economic equilibrium of the Walrasian kind, but which takes into account Patinkin's doctrinal contribution. It is therefore in this context that we must examine his claims concerning inflation and his economic policy proposals for controlling the trend of the economy.

According to this approach, full employment follows automatically on the action of the market forces, which establish appropriate relative prices of goods and services, so that the effective real structural relations will emerge irrespective of the size of the stock of money. Actually it is the invisible hand of competition that determines the equilibrium vector of relative prices that ensures full employment and an efficient allocation of resources. Starting from this situation, the increase in the quantity of money is reflected entirely by the level of prices. In his *The Optimum Quantity of Money* Friedman imagines a helicopter successively dropping a multiple quantity of money on the economy in equilibrium. As every individual wishes to maintain the structure of his portfolio unaltered in real terms, the result of these operations will be an increase in nominal incomes and prices: i.e. inflation. We may note however, that whereas this analysis is appropriate for a demand-induced inflation, the disequilibria caused by wage pushes do not on the other hand call for the same attention. What is lacking in fact in the monetarists' usual version of inflation is any particular concern with nominal wages.

This is no accident as there are several passages in which the price wage spiral is rejected or denounced as meaningless. In the monetarist view, once M (quantity of money) has been introduced and the real forces defined, the consequent equilibrium will be determined. «Suppose that in this tranquil Walrasian image there is an increase in nominal wages sponsored by a new sentiment that permits this to happen. Clearly with M (and V) rigid, than prices must rise, while output and employment must drop. Unemployment will occur as the demand for nominal money balances increases because of the price rise and interest rates rise depressing investment»[75]. In this framework, therefore, the wage rate should adjust to the real forces of the economy. If the workers were to persist in resisting the automatic mechanisms, the result would be a temporary state of inflation with the contemporary presence of unemployment; in other words the well-known stagflation of recent years. What must be stressed is that in this conception the monetary variable always appears independent.

The stability of prices and the economy is linked therefore to an orderly monetary policy and the respect of market mechanisms. This is the premise underlying Friedman's well-known proposal that the quantity of money be increased in time at a constant rate[76]. A suitable rate of growth having been chosen for money, the result should be not only a steady price level but also a time path of economic development where the wage rate is synchronized with the increases in productivity. Should the workers fail to adjust voluntarily to this wage pattern and base their behaviour on disorderly claims for wage increases, they would bear the consequences in the form of unemployment and inflation. The maintenance of the norm proposed by Friedman might however cause unbearable levels of unemployment, should the trade unions insist on wage claims appreciably exceeding the growth of productivity. Experience has taught that it is very hard to contain cost-push inflation with measures designed to contain global demand.

Not unlike what happens in the Keynesian model, therefore, an orderly development of the economy depends on a certain amount of cooperation among the social forces also in the monetarists' model. However the monetarists are unwilling to admit this openly, trusting in the regulatory power of a uniform and controlled change in the quantity of money. For them an erratic trend of the quantity of money is likely to cause more evil than good, as the effects of the monetary manipulation require variable and partially unidentified time lags. Those of the Keynesians who are most inspired by the *General Theory* are willing on the contrary, though of course within certain limits, to adjust the stock of money to the changing needs of the economy. They hold in fact that this is likely to attenuate cyclical fluctuations. In view of these opposite conceptions, Weintraub's judgment does not therefore seem a paradox. According to him «Keynes may be better suited for the honorific monetarist title than those who insist that by implementing a single rule, money, thereafter, does not matter»[77].

---

75 See S. WEINTRAUB: *Keynes and the Monetarists, op. cit.*, p. 40. Weintraub points out that this is the same as saying that unemployment is caused by wage rigidity: a conclusion reached more than once in the debate between Keynes and the classics.

76 This rule however, appears criticisable when considering the functioning of the economy at a lower aggregation level. Hines (see A.G. HINES: *A Reappraisal, op. cit.*, p. 50) points out that the money demand for transactions should at least be broken up into demand for consumption and demand for investment. For this reason the overall money demand for transactions is not independent of the composition of the national product. A sectoral consideration of the economy along the lines indicated by Perroux in a recent report (see F. PERROUX: *La Componente Monetaria dell'Inflazione negli Anni Sessanta, con particolare riguardo agli Influssi Internazionali*, in Proceedings of the Meeting "L'inflazione negli Anni Sessanta", Cassa di Risparmio di Roma, Roma 1972) lends further strength to the criticism. Lastly, it must not be forgotten that in an income-constrained approach the effective overall demand depends not only on the quantity of money, but also on the distribution of money among the economic actors. All this leads to the conclusion that the time has come to supersede the overall macroeconomic approach, whether it is of the neo-quantity or of the Keynesian type.

77 See S. WEINTRAUB: *Keynes and the Monetarists, op. cit.*, p. 49.

Federico Arcelli Giorgio Di Giorgio

Conclusions

The survey of the doctrine of the two schools has thus been concluded with somewhat surprising results. The debate between monetarists and Keynesians generally dwells at length on the different empirical assumptions supporting the two theories (the different elasticity of substitution between money and financial assets and between money and real assets; the greater or lesser importance attached to the hypothesis that the demand for money is stable and the consequent different meaning of expectations in the two models), deducing there from that fundamental and apparently incompatible differences exist between their respective positions with regard to economic policy interventions. This discussion usually acts as background for empirical comparison between monetary and fiscal policy. We have noticed, however, that what has emerged from the empirical research does not in any way prevent the formation of possible areas of agreement concerning effectiveness and the mechanisms transmitting the monetary impulses.

The logic of the portfolio approach, which is common to the more sophisticated versions of both models, lends strenght to the convergence of the analysis of the two schools from the formal point of view. At several points, moreover, we have observed that the effects anticipated by the monetarists do not appear incompatible with the implications and structure of the Keynesian model, once the importance of the interest rate has been acknowledged. The widespread opinion that money does not matter for Keynes has thus been definitely rejected. Weintraub paradoxically turns this accusation against the monetarists. However these results might obscure the real fundamental differences between the two lines of thought. To make their positions clear, attention must therefore be centred on other levels of analysis. The substantial differences between the two schools have in fact to be sought in the different theoretical frameworks and in the different philosophies underlying the operative tools. Both the quantity model, expressed by the so-called monetary multiplier (or better by the function of money velocity), and the Keynesian model of effective demand define macroeconomic alternatives apparently unconnected with micro-economic analysis.

On closer inspection, however, we find that Friedman's theoretical reference is identifiable in a model of general economic equilibrium of the Walrasian kind, though more generalized so as to include the complex categories of a portfolio approach; while as far as Keynes' model is concerned, a recent reappraisal of the General Theory also allows a connection to be made with the general theory of markets discarding, however, the hypothesis of full information.

Keynes' introduction into the analysis of expectations concerning an uncertain future calls, however, for a radical change of perspective. Keynes' contribution to economic theory can in fact be identified in the effort made to extend the use of the tools of the general theory of value into the area of macro economic disequilibria.

Though starting from a common theoretical point of reference (general economic equilibrium), the positions of the two schools thus belong in the end to two widely different philosophies. That of the monetarists reflects the natural order of the neo-classical construction, from which it deduces rules of economic policy designed not to disturb the spontaneous but ordered development of the economy. Money cannot alter the real order of facts in the long run; therefore it is only right that the supply of money, determined exogenously, should grow according to a uniform rate.

The Keynesian philosophy, on the other hand, inspired by a theory of macroeconomic disequilibria caused by the absence of perfect information as well perhaps as by the presence of monopolistic elements in the system, reflects a less illuministic conception of economic activity and consequently assigns a more flexible role to the instruments of economic policy, monetary changes included. The quantity of money is therefore not only determined exogenously; it also reflects the needs of the economy and thus possesses an endogenous dimension too. The difference between the two schools, then, lies not so much in the mechanisms transmitting monetary impulses to real variables, which anyhow still appear controversial, as in the different rules of economic policy and the different philosophies supporting the logic of the interventions.

The preceding discussion thus seems to have led to some important conclusions: the recognition of the different philosophies has not prevented us from individuating lines of convergence and contact between the two types of analysis. The study of the models has shifted from a macro-to a micro-economic level and this allows a deeper understanding of their meaning and a clearer explication of the hypotheses. Lastly, it seems to me that, seen in this perspective, the reappraisal of Keynes' thought has not only clearly defined the characteristics of a monetary economy, but also points out a very promising road along which economic theory should advance towards a synthesis already foreseen by some *avant-garde* economists. «The synthesis, if it comes, will re-establish the claim of economics to be the fundamental theory of society»[78].

---

78 See A.G. HINES: *On the Reappraisal of Keynesian Economics, op. cit.,* p. 62. On the same page Hines notes: "The frontiers of economic theory are now really very exciting. The work of Keynes which has now been re-interpreted in a theoretically relevant fashion opens the way to a serious analysis of disequilibrium phenomena whether of the deflationary or inflationary variety. The reswitching debate and the Sraffa revolution hold out the possibility that questions of relative price can be divorced from questions of distribution...One now feels that we are probably at the beginning of what Kuhn calls a period of revolutionary science ...One's faith for the future is in a synthesis which will include the best of Ricardo,

Marx, Keynes and the Neo-Classics". As an example of appliance of the new Keynesian perspective to inflationary phenomena and to the study of the labour market, it is worth mentioning the book by E.S. PHELPS et al.: *Microeconomic Foundations of Employment and Inflation Theory*, W.W. Norton, Philadelphia, 1970, which contains studies by Alchian, Holt, Phelps Mortensen, Hynes and others all developing the analysis of disequilibrium situations, connected with sub optimal information. This approach has already opened new fields both of theoretical and empirical research We can also mention Hines paper: *The Determinants of the Rate of Change of Money Wage Rates and the Effectiveness of Income Policy* in "The Current Inflation" ed. H. Johnson and A.R. Nobay, London, Macmillan, 1971.

Federico Arcelli **Giorgio Di Giorgio**

CHAPTER III
MONEY AND THE BALANCE OF PAYMENTS

Federico Arcelli **Giorgio Di Giorgio**

FOREWORD

Mario Arcelli wrote the two following papers in the mid-Seventies after the year of study spent at the M.I.T. of Boston. He tried to resume and discuss some thoughts about the state of economic theory on the role of money and its significance according to the latest development.

If by one hand he developed some reasoning about the theory of balance of payments adjustment (known as the Tinbergen-Mundell Approach), at that time considered dominant but also under severe criticism, considered, after the oil crisis, with its tremendous impact on balance of payments equilibrium of many industrialized countries, that it appears questionable and lacking even for the capitalistic world. On the other one, Mario Arcelli focused more on the economic concept of money, and on its foundations.

The outcome of his overview on economic theory about money was a bit dissatisfactory by the perspective of a researcher, given the uncertainty of the interpretation on the table referred to a so important and key point of the theory. He notes, in fact, that it "has not arisen merely from political or empirical facts, but it is deeply rooted in the analysis of its own foundations". Mario Arcelli noted that the accepted theory which "relies on a targets-instruments approach in order to achieve both "internal and external balance" of an economy", suffered for major criticism when it was not maximizing any definite social welfare function and it may resulting even destabilizing for the economy, at the limit even when instruments are appropriately paired with the objectives on which they have the most influence.

He acknowledged the three main approaches followed by economists during the previous years in order to lay the foundations of monetary theory, and in particular the *monetary constraint to exchanges* (equivalent to deny the appropriateness of budget constraints of traditional theory); the line that *tries to define a role for money inside the general equilibrium theory of exchange, without assuming it a priori* (in this case the focus is on the role of money as a means of transactions); and, a third line of reasoning, where monetary exchange and the role for money may instead be derived as a reflection of imperfect information about prices. But, by describing the current state of the debate in 1975, Mario Arcelli also notes that there are several points in favour of each interpretation, but not a clear dominant path.

Notwithstanding this, he concludes that "some important results have been achieved. Not only have we succeeded in clearly defining the impact of money on the economy by a comparison with barter conditions, before considering price-uncertainty: a) money allows specialization; b) money bypasses ineffectiveness of barter transactions; and, c) money helps reduce explicit or implicit transactions costs; but we have come to grasp that when transactions are costly, money may play a constraining role in the working of the economy in the sense that a Pareto-efficient equilibrium cannot be achieved if some agents run out of stocks of it at any step of a sequential trading process."

This allows to consider in a new light the importance of price uncertainty and imperfect information. He recalls that "we have also learnt that costs of transactions and costs of information gathering are the main reasons for an imperfect knowledge of prices in the short run. The introduction of price uncertainty in the economy has enlarged and modified our notion of the constraining power of money, making clear the distinction between notional and effective demand. We have then realized that the introduction of imperfect information about prices confers to money the additional distinctive power of source of effective demand. Money flows, therefore, become relevant providers of market signals together with price conditions.". He then concludes that "one could argue that only when the above situation is considered is there opportunity to perceive the full implications of a monetary economy. So finally one is turned back to the microeconomic foundations of Keynes' theory that tentatively provides us with a description of the working of the economy, consistent with the previous position: progress in this field will be helpful not only in precising the behaviour of money both at an individual level and for the working of the whole economy, but also will provide a fruitful framework for the theoretical foundation of monetary policy". Keynesian theory was, then, far from being only history.

Federico Arcelli **Giorgio Di Giorgio**

SOME THOUGHTS ON THE FOUNDATIONS OF MONEY [79]

1.1 Three main approaches have been followed by economists during the last few years in order to lay the foundations of monetary theory.

a) *A first line of reasoning is presented by imposing a monetary constraint to exchanges*, what is equivalent to deny the appropriateness of budget constraints of traditional theory. The role for money as a means of transactions is assumed, without explaining it inside the structure of the models under consideration. It is, however, possible to distinguish within this approach quite different kinds of models. On one hand, we have the work done by Clower [7; 8] and followed by Barro and Grossman [4], related to the Keynesian theory that he reinterpreta in a microeconomic framework. But Clower, different from Keynes, insists on the role of money as a means of payment, neglecting the store of value aspect of money. As he states [8]: «If money is to be distinguished by the functions it performs, it is to the medium of exchange function that we must address our attention». By so doing, what he actually proposes is «are formulation of established microeconomic analysis suitable as a foundation for explicit analysis of the working of a money economy» [8] without treating in an articulated way the problems of uncertainty. Besides, the allocation of resources at different times as a consequence of an intertemporal utility maximization process is not taken into consideration[80].

   On the other hand, we find some papers by Grandmont and Younes [12; 13], who present a sequence economy similar to the one described by Friedman in his «Optimum Quantity of Money». In their model, however, «money has no direct utility but is the only store of value; in addition, the role of money as exchange intermediary is introduced (in a very crude way) by assuming that money is the necessary counterpart of any purchase or sale of a consumption good and by imposing a priori restrictions on transactions». The two kinds of models included in this approach are deeply different in nature: their only common character being the imposition of monetary constraints to exchange. Grandmont and Younes are actually concerned with some of the basic neoclassic issues in monetary economics such as the «Optimum Quantity of Money» and the «Classical Dichotomy» and assume that every trader seeks to maximize an intertemporal utility function whereas Clower's work is somewhat related to Keynesian theory and does not explicitly consider an intertemporal dimension. Separate comment on them is, therefore, necessary[81].

---

79 From: Metroeconomica, vol. XXVI, Gennaio-Aprile 1975, Fasc. I – Cappelli Ed. Bologna. The first version of this study was prepared for presentation as a discussion paper at the Money Workshop of the Department of Economics of Massachusetts Institute of Technology on December 1973. Prof. M. Arcelli have profited of helpful comments by the participants in the seminar and particularly by Professors Franco Modigliani and Stanley Fischer who aided him greatly. Acknowledgments of the author Mario Arcelli: a) The C.N.R. (Consiglio Nazionale delle Ricerche) has supported this research which is a part of a more comprehensive work; b) M. Arcelli also gained support by the Committee on International Exchange of Persons (Senior Fulbright-Hays Program); c) M. Arcelli discussed this matter also with Professor James Tobin (Yale University) with whom he had a fruitful conversation on the subject of the research and Professor Herschel Grossman (Brown University) that in a later correspondence sent me unpublished papers on related topics.

80 As Barro and Grossman acknowledge [5] both Clower's papers [7; 8] and their own [4] «explicitly considered only the momentary flow budget constraint and dealt only implicitly with the essential intertemporal nature of the household choice problem». It is only in [5] that Barro and Grossman consider explicity the intertemporal nature of the household utility maximization problem. Their new analytical framework provides us with a deeper and more articulated insight on the role of money as a constraint on exchanges and on effective demand. See farther on my comment on this fundamental issue.

81 Within the first line of reasoning, one could also mention Shubik's new approach to the theory of money [33]. Shubik's contribution to the theory of money is related to the solution of non-cooperative games that are intended to represent a trading economy. By the formulation of a model of an exchange economy in terms of a strategic game, which can be played non-cooperatively, Shubik is led to beat one of the commodities in a nonsymmetric way. Such a commodity could be regarded as commodity money used for transactions. The requirement that a special commodity be used in payment imposes a constraint to exchanges. If this constraint is binding, the limiting non-cooperative equilibrium points are not the same as the set of competitive equilibrium. For the individual that means that the cash constraint has prevailed over the

b) *A second and more problematic line of reasoning tries to define a role for money inside the general equilibrium theory of exchange, without assuming it a priori. Here it is also possible to identify two main streams of thought.* On one hand, I would set some recent papers by Starr [34] Ostroy [29] and Niehans [27] departing from the standard Walrasian approach with the assumption that exchange is a do-it-yourself affair: in other words individuals do not exchange with the «market» but rather they exchange with each other. The focus is on the role of money as a means of transactions. It is argued that «if we cannot find a role for money when equilibrium prices are known, we shall not find one when they are unknown. There can hardly be a speculative demand for the medium of exchange without a transactions demand and this transactions demand does not depend on price uncertainty» [29]. In these models, the ineffectiveness of barter and informational requirements for trade appear as the main reasons for the introduction of money.
A particular feature of Niehans' analysis is the explicit consideration of transactions costs: Niehans, however, assumes that the ultimate flows of commodities in the economy are not affected by transactions costs. Therefore the problem of payments arrangements on which he dwells is theoretically separated from the determination of the ultimate flows in general equilibrium analysis. In a later paper, however, Niehans [28] overcomes this separation. Niehans' new approach marks a major departure from his previous analysis and relates him to Hahn's position [15; 16] which I would set on the other side of the second line of reasoning. Hahn too is trying to define a role for money inside the theory of exchanges, but he differentiates sharply from the former authors both assuming the existence of a central agency (the marketeer), to coordinate transactions which are not costless, and extending the analysis with the addition of a transaction technology and the formulation of an explicit intertemporal structure. A similar scheme with some simplification may be found also in Kurz [22; 23] and Starrett [35] and with no convexity in the transaction technology, more recently in Heller [18] and Heller-Starr [19].
The features of the model considered by Hahn induce him to insist on the fundamental importance of the intrinsic sequential structure of a monetary economy and seem finally to reconcile him with the third approach to the foundations of monetary economy, which I will describe below. If one should select a common character of the models included in the second line of reasoning, one could easily point out that both Hahn and the former authors have tentatively tried to provide us with an equilibrium theory of money where transactions costs (explicitly or implicitly considered) and ineffectiveness of barter seem to play a crucial role.

c) According to the third line of reasoning, monetary exchange and the role for money may instead be derived as a reflection of imperfect information[82] about prices. This is a product of an uncertain future as well of lack of knowledge about the present. The former condition (uncertain future) specifically helps to enlighten the role of money as a store of value and well known propositions in the theory of choice under uncertainty can be invoked to explain why transactors should include money in their portfolio of assets. These propositions, however, assume the existence of money as given, without inquiring about its origin. We must refer to Radner's [30] fundamental distinction between uncertainty about the environment and uncertainty about the behavior of the agents to get an insight of some of the circumstances that induce the use of money. Radner, through the analysis of sequence economies, has suggested that the demand for money might arise from the unpredictability of future prices, produced by uncertainty about the behaviour of the other transactors or by computational limitations. Radner's suggestion agrees with some implications drawn from the microeconomic reinterpretation of Keynes' General Theory. In fact, it has been observed that the transition from the Walrasian world, where the knowledge of intertemporal price vector is perfect, to the Keynesian world where expectations rule, marks the end of a class of invariant states of the model and the attribution of a special status to money as a liquid asset.
On the other hand, the lack of knowledge about present prices and qualities of commodities strengthens specifically the role of money as a means of transactions. Alchian [1] and Brunner and Meltzer [6] have underlined that collating information about potential exchange opportunities is not without cost: there is a variety of bias and offers and the search for the best opportunities requires a costly acquisition of knowledge about the terms of trade one can expect to prevail. The role of money as a means of transactions can then be fully appreciated: money enters «into almost every exchange because it provides

---

budget constraint. The way followed by Shubik to relax the former constraint is the modelling of a loan market with an explicit bankruptcy law. Extensions of the basic model are in progress.

82 The word «information» will be used both as a synonym for «knowledge» and, at a more sophisticated level, as an equivalent for «signal related to the state of the environment which determine a response according to the decision maker's function».

the most economical vehicle of exchange» [1] and this is due to the fact that transactors can acquire information about it at comparatively low marginal cost [6]. So at the end of this sketch of a relevant part of contemporary literature on the argument at issue, one can state that costs of transactions (implicitly or explicitly considered) and imperfect information about prices, are the two crucial assumptions upon which the foundations of monetary theory may be laid[83]: very frequently the two conditions are strictly interwoven.

1.2 From the preceding considerations, it is by now clear that money cannot be integrated into the standard Walras-Debreu theory of value. If we follow Clower's proposal [8] to introduce money into the standard theory by superimposing on the existing budget constraints the injunction that current purchases be financed by sale of money only, then we fall into the paradox pointed out by Ostroy: «This modification contradicts the belief that the introduction of money improves the allocation of resources. Because it is an additional constraint, it is at best not binding; and if binding it will narrow the set of permissible exchanges compared to its barter counterpart. Clower's constraint makes no sense in the Walrasian model of exchange, but for a perfectly sensible reason. It does no good to append what is a trading constraint to a model which ignores trade» [29]. Moreover, in a classical Walras-Debreu economy the opportunities open to individual households do not depend on transaction dates and this cancels out the store of value function of money.

We might actually consider money in a sequence economy, which is consumption equivalent[84] to the Debreu model, as Hahn did. Money is introduced there to sustain the intertemporal connectedness of a sequence economy of spot markets. Then it is shown that the same result could have been reached by creating a contingent futures market, one for each state, in unit of account and for each pair of adjacent dates. So money appears to be inessential in the model in the sense that no monetary variable need enter into the description or determination of the equilibrium of that economy. Two chief partial conclusions can, therefore, be drawn at this point:

a) Monetary theorists are becoming more and more aware that trade arrangements in a money economy «*correspond more closely to the completely decentralized process envisioned in traditional treatment of barter exchange than to the completely centralized process implicit in Walrasian models*» [9]. Monetary exchange actuallv appears «as a phenomenon involving the existence of a set of organized markets in each of which units of one particular commodity can be traded directly only for units of one or more other commodities that custom or law has specifically designated as means of payment» [9].
b) On the other hand money *becomes essential only in the context of an intrinsic sequential structure* of a monetary economy, where transactions cannot be concentrated at one time and where the future does affect present decisions: that is outside the Arrow-Debreu economy.

1.3 Barter economy can be defined in at least two main different ways and accordingly the introduction of money into the economy needs different explanations. Either we may require a double coincidence of wants in order that barter can take place, or more simply we may define barter as any exchange of a commodity for a commodity in view of an immediate or a future advantage of both traders. The double coincidence of wants requirements is a severe restriction on the feasible trades: when this condition is respected it is possible to generate examples of efficient allocations which could be reached, should trade take place, but that actually are not achieved because the lack of double coincidence of wants prevents trade. In those models money, when introduced, is defined as the only commodity to which the standard restrictions on desirability of commodities traded do not apply and which permits to achieve an efficient equilibrium. Money then is accepted also when there is no excess demand for it and may be given away even in the absence of excess supply. Starr [34] has proved that in a monetary

---

83 It is certainly useful to remind at this point that the store of value function of money is connected with the notion of liquidity, which generally defines a broader set of assets than the one of the means of transactions. The identification of money with the medium of exchange is not without criticism as Brunner and Meltzer recognize [6] «A definition embracing a larger collection of assets is appropriate if there are close substitutes for the medium of exchanges on the supply side. In this case, slight changes in relative prices reallocate output between the medium of exchange and other assets so the collection of assets most useful for explaining changes in the general price level differs from the assets that serve as medium of exchange. However, even if evidence suggests that a broader collection is justified empirically and the term money is used to refer to the broader collection the significance of the medium of exchange function and its importance for explaining the productivity of monetary assets remains».

84 A sequence economy is said to be consumption equivalent to a Debreu economy if in equilibrium each individual enjoys an equivalent consumption level in the two economies. That implies that for each allocation through the market in the sequence economy, there is a corresponding allocation in the Debreu-economv, so that each individual is equally well off in the two economies. A necessary and sufficient condition for the sequence economy to be consumption equivalent to a Debreu-economy is that the price system of the sequence economy has the Debreu-property [16]: paragraph 1.4 provides a simple illustration of that property.

economy so defined, given an equilibrium price vector, there is always a monetary exchange satisfying the following three conditions:

a) price consistency (that is, exchanges between each pair of traders are equal in value);
b) monotone excess demand diminution (that is, trade proceeds monotonically towards equilibrium);
c) excess demand fulfilment (that is, trade satisfies all excess demand);

Whereas in the corresponding barter economy only two out of these three conditions might be met at a time if the double coincidence of wants requirement is not satisfied. So it is understandable that the *introduction of money allows the achievement of allocations, which are unfeasible under strict barter conditions*. Such definition of barter economy and of the corresponding monetary one is nevertheless unsatisfactory from a logical point of view. Why not eliminate the double coincidence of wants for all commodities instead of introducing peculiar good called money? The release of the double coincidence of wants requirement may allow the fulfilment of efficient equilibrium also in the barter economy but then, apparently, there is no room left for money. Therefore a new role for it must be defined. That is what Ostroy has tried to do while assuming more realistic conditions for trade [29]. His results will be the object of my subsequent comment.

Here I wish to point out that a role for money could be defined if one dismissed the assumption that trade takes place at equilibrium prices. In fact, in a barter economy without the double coincidence of wants requirement, there is room for speculative transactions when trading takes place out of equilibrium. In that case the introduction of money provides the economy with a means for diminishing the number of transactions and reducing the speculative element[85]. The structure of the economy is, however, deeply changed and one is led to discuss the role of money in a sequence economy. So one is back to the second partial conclusion, that is, proposition b) of paragraph 1.2, and it may be useful in that connection to sketch some fundamental distinctions.

1.4 Let me state precisely, following Hahn, the distinction between a sequence economy where transactions dates are essential and a Debreu transactions economy: this distinction will appear critical in order to define when the role of money is binding. The latter economy is distinguished from the former by two postulates [16]: the relative prices of goods faced by agents are independent of the transaction date and households do not have to balance their books at each transaction date but only over all dates taken together. It follows that in a Debreu economy varying transaction dates may be possible. But these transaction dates are inessential as the same equilibrium is obtained independently of these dates and all transactions may be concentrated at the first period. This proposition does not generally hold in a sequence economy. It is, however, possible to build sequence economies consumption equivalent to a Debreu economy that are, therefore, inessential.

Consider an Arrow-Debreu economy where transactions are concentrated in one period and assume there is a unique equilibrium with price vector p^*. Alternatively allow transactions at every date in every good with availability date not preceding the transaction date. There will be a full vector of Arrow-Debreu prices at every date and let us name q_t the vector at time *t*. If in this enlarged economy equilibrium is attained when each q_t is proportional to p^* as viewed from *t*, the price system is said to possess the Debreu property and nothing will have been altered by allowing transactions at every date [17]. In fact, in a Debreu economy households may always be interpreted as facing a present value budget constraint, all of them using the same discount factor to convert prices at *t* into present prices. The described sequence economy is not only equivalent to a Debreu economy; it is also inessential[86]: actually the sequence makes no difference to what we have already deduced about the equilibrium of the Debreu economy. In such an economy, money too is inessential, because the determination of equilibrium can be achieved without the presence of any monetary variable. It is now possible to state the conditions of a sequence economy that is inessential [17] and by exclusion to identify the requirements of a true sequence economy:

a) there must be an absence of sequential learning in the Radner sense [30], that is, the sequence of the markets and the extra prices should not make possible more information on the environment than was

---

85 Arrow and Hahn [3] state: «Of course, one of the advantages of a monetary economy is that it makes it possible to have anonymous markets such that every agent need make, at most, two separate transactions if he desires to exchange a given non-money good for another. We have seen (Chapter 13) that this does not entirely remove the speculative element from transactions, but it greatly economizes in the number of such transactions. In particular, only one transaction is needed in the exchange of money for any good».

86 The introduction of transactions costs which implies a distinction between buying and selling price of goods does not modify the conclusion previously reached concerning the inessentiality of the sequence, provided that the Debreu-property is still holding. The formalization of that property is more elaborated [16] but always amounts to saying that each agent uses the same discount factor to convert prices at t into present prices. The introduction of a transactions technology, however, implies that one has to define Pareto- efficiency relatively both to transactions technology and distribution of endowments among households.

available when transactions were concentrated in the first period. More precisely, the pair information function – decision function must be unaltered;

b) the prices at all dates must be known to all agents at any date.

We then recognize that all consumption equivalent sequence economies derived from an Arrow-Debreu model are inessential. More than that: to say that a sequence monetary economy is inessential is equivalent to saying that it is Pareto-efficient and that the same equilibrium is attainable in a Debreu economy without money [16]. However, *to say money is inessential does not imply that money has no role: it merely means that the role of money is not binding*. I shall discuss this problem in the next paragraph. So to say money is inessential is something different from saying money is neutral, even if the two concepts are strictly related: when the classical dichotomy is holding money is inessential but it is also neutral.

1.5 A sequence monetary economy is said to be inessential according to Hahn's definition, when the equilibrium achieved by the economy is Pareto efficient. However the inessentiality of the economy might well be the outcome of the introduction of money. Hahn actually recognizes that the introduction of money can improve on sequence economies whose equilibrium is inefficient, in the sense that a corresponding sequence monetary economy may result efficient [16].

If one of the main differences between a sequence economy and a Debreu-economy is that in the former all agents must balance their books each time *t* while in the latter they need not, and this appears to be a fortuitous institutional requirement which has nothing to do with the real part of the economy [16], then the introduction of money may actually release the stringent set of restrictions imposed by the sequence of budgets and allow the achievement of an efficient equilibrium[87]. Hence a role for money is defined; however this is neither a general nor an essential one.

Indeed, in general, there is no reason for a sequence economy with money to be inessential and it is even possible that the sequence economy is inessential when the corresponding economy with money is not. In this latter case, the role of money appears really binding[88]. To be explicit: money is given an exclusive role to play as a medium of exchange, as it was assumed in Grandmont and Younes' model [13], that imposes a constraint on the feasibility of exchanges such that a Pareto-efficient equilibrium cannot be achieved. So by imposing a monetary constraint, the equilibrium is worsened. Then if we want to avoid the paradox pointed out by Ostroy, we ought to assume in this instance that the transactions technology was improved by the introduction of money.

---

87 Kurz [23] like Hahn underlines that it is generally possible that the sequence of budgets associated with the sequence of markets of a barter economy will constitute a sufficiently stringent set of restrictions that will lead to an inefficient equilibrium. This is the reason for the introduction of money. He also provides an intuitive understanding of the dependence of the inefficiency involved in the transaction technology. «Consider a case in which you and I found ourselves in the following situation: my endowments today and tomorrow are:

today = Inserire formula
tomorrow=Inserire formula

Your endowments are:

today = Inserire formula
tomorrow= Inserire formula

where (Inserire formula) are common to both of us. Now I want part of your «today bread » and you want part of my «apples tomorrow». Such an exchange is possible except that it calls for a futures contract where I shall buy spot bread and sell futures apples while you do the opposite. If the cost of carrying out a futures transaction is higher than the cost of spot transaction, then the trade involving futures markets is inefficient: you can give me spot bread today and I shall give you spot apples tomorrow without a futures contract but rather using spot transactions only. This leaves me with a debt to you until tomorrow. Such an act of creating a debt today will violate our budget constraints and is not feasible in the barter economy above without the introduction of money as a store of value. If we allow in our economy the appearance of a medium of exchange and the emergence of credit, then the intertemporal transfer of obligations may release the stringent set of restrictions imposed by the sequence of budgets. Thus we may tentatively conjecture that in a world with transaction costs and a sequence of markets a medium of exchange and credit system is needed in order to achieve efficiency». Different from Kunz and Starrett, Hahn, however, rather than arguing extensively on the possibility of re-establishing an efficient equilibrium, appears to be concerned with situations in which the monetary constraint is really binding. In those cases money plays an essential role.

88 A deeper comprehension of the binding role of money is obtained when transactions take place out of equilibrium. See paragraph 1.9.

1.6 A binding role for money generally arises in connection with a sequence economy when the sequential budget constraints cannot be fully avoided by granting each household a positive initial endowment of money. The resulting monetary equilibria are not Pareto-efficient: spot transactions in money that are costless, are not enough to fulfil Pareto-efficient equilibrium.

When the Debreu property is not holding and transactions of goods are not costless (therefore the buying price of the goods is different from the selling price) «the most important consequence is that in general, households face a sequence of budget constraints and there may be no unique set of discount rates applicable to all households which allows one to amalgamate all those constraints into a single present value budget constraint. The intertemporal transfers open to a household depends on whether at each of the two dates he is a seller or a buyer of the good in terms of which such transfer is most advantageous» [15]. To avoid such budget constraints, the household is expected to undertake spot and also forward transactions in money, should the former not be enough. But when the latter transactions are actually undertaken, being costly, as it is plausible to assume, the outcome is necessarily equilibrium not Pareto efficient. Since the household could not fully avoid budget constraints by means of spot transactions in money that are costless, the economy suffers a loss, as part of its resources must be allocated to face the costs of forward transactions. In this framework money plays a constraining role and is, therefore, essential.

Hahn is actually interested in the result of the interplay between payment arrangements with money and flows of goods in a true sequence economy. Not too much importance, therefore, should be attached to the word «efficient» as a goal to be achieved. To keep the same vocabulary to define concepts belonging to very different structures may be misleading. We can, therefore, restate that there is only a limited role for money inside a Debreu economy, but that it is meaningless to make a comparison between it and a true monetary economy. Hahn's approach to money is certainly clarifying in that it claims the essential monetary phenomena to be strictly associated with true sequence economies, but it is far from being a satisfactory foundation of monetary theory. Actually he has lately acknowledged that to demand, as he does in his models, that all prices be announced and that there is no uncertainty may well lead to uninteresting conceptualisations. Indeed nothing is known about disequilibria situations and the constraining role of money does not reveal all its implications.

Hahn's work is, nevertheless, helpful as far as it gives us a standard whereby to judge some recent efforts aiming at the formulation of monetary theory. The distinction between essential and inessential economies allows us to define theoretically when the role of money is binding and, therefore, essential. It has been employed by Hahn himself to show that Grandmont and Younes' economics are either isomorphic to an inessential economy or, if not, this is «because monetary management has not neutralized the transaction constraint or because money is the only means of storage» [17]. Moreover, in Grandmont-Younes' models there is no uncertainty and the role of money as a means of payment is assumed by imposing a priori a monetary constraint on transactions. Hahn can, therefore, conclude that the foundations of monetary theory have not been laid by Grandmont and Younes' analysis.

1.7 I will use the same standards in order to comment on Ostroy's attempt to explain the role of money inside a general equilibrium model of exchange. Ostroy's scheme departs from the Walrasian model in that he assumes that trade occurs between pairs of individuals and not with the market, so that the advantages of multilateral exchange must be obtained through a sequence of bilateral trades. Exchanges leading to equilibrium are then viewed as a sequence of simultaneous bilateral trades.

Following Ostroy, we have, however, to distinguish between equilibrium (*W*), which is fulfilled, when the sum of individual excess demands is zero for each commodity and equilibrium (*A*) where not only the aggregate but also individual excess demands are zero. Ostroy is mainly interested in the possible paths from (W) to (A) following which prices are unchanged and equal to competitive equilibrium prices. In comparing money and barter arrangements, Ostroy considers the number of periods in takes to go from (*W*) to (A) as the single criterion whereby to judge the improvement determined by the introduction of money. The trading sequences are limited *a*) by technical feasibility *b*) by informational feasibility; and *c*) by behavioural feasibility.

The first limit defines the restrictions on a sequence of exchanges imposed by the fact that trade occurs between pairs. By superimposing the bilateral balance on every trade, one restricts further the set of trades beyond the demands of technical feasibility. It is, however, not required in case of barter that there be a double coincidence of wants. Ostroy actually considers also an indirect exchange model. The informational feasibility adds a further restriction to trade establishing that a pair cannot base its trading decision on information available only to other pairs. Each member of a pair knows only the equilibrium prices, his current position and his excess demands given the prices and the budget constraint. «Each member of a pair of trading partners tells what he knows to the broker who then decides what trades they should make. The different trading pairs have different brokers who cannot communicate so the situation is much the same as if the pair themselves decided what to trade» [29].

The third property that is c) requires the previous limits and the additional restriction that each individual has no incentive to depart from the sequence. Starting from given endowments each individual aims at maximizing his utility function subject to the constraints described above. It is also assumed that there is no commodity which

is a significant fraction of the value of the individual's planned purchases so that it is possible to bring out more clearly «that the essential feature of monetary exchange has its origin in the trading arrangement and not in the nature of the money commodity» [29].

Given the assumptions, Ostroy first shows that the competitive equilibrium allocation is not information ally feasible in the indirect exchange model. In fact, traders «do not know each other's excess demands and trading decisions are bound by this ignorance. But to take advantage of indirect exchange, individuals must act as middleman passing excess supplies in just the right sequence of intermediary trading to the final demands. What is the right sequence depends on the entire configuration of excess demands as well as on the order in which all pairs will meet» [29]. So a first relevant conclusion is the denial of the informative sufficiency of prices in the framework of indirect exchange.

But Ostroy's central thesis is that the competitive equilibrium allocation is not informationally feasible also in the direct exchange model, when trades must satisfy the bilateral balance condition. The argument relies on the explicit assumption that no commodity is in sufficient supply to serve merely as a balancing item and it will not do to pay one's debt in just any commodity. So individuals are involved in indirect trade whose informational requirements they cannot meet. However, to have a budget balance over the course of trade is a necessary condition for equilibrium. Hence the need for money as a record-keeping device that helps achieve individual budget balances without imposing bilateral trade balances.

In accordance with that, Ostroy defines the monetary version of the model of a trading economy by introducing a monetary authority whose function is «to collect the bits of information individuals have about each others' trading histories. Each will require his trading partner to write a statement indicating the amount by which the partner's purchases exceed his sales. This record is forwarded to the monetary authority who revises individual accounts on the basis of this new information. Sellers, by requiring payments in money, are guaranteeing a steady flow of information such that the monetary authority is able to monitor trading behaviour» [29]. Then money as a record-keeping device is not only a unit of account, but it possesses an informative content which allows the achievement of equilibrium. So Ostroy has correctly defined a role for money inside the traditional theory of exchange.

1.8 I shall argue, however, that the work by Ostroy, although theoretically fruitful, is missing some important points in the definition of the central features of a monetary economy and, in particular, the structure of his model does not satisfy the requirements proposed by Hahn in order that the role of money may be judged essential[89]. In fact, we realize:

a) that *the set of equilibria attainable in Ostroy's economy is independent of the dates of the transactions*. The same competitive equilibria can actually be achieved at different times [29]. Ostroy's proposition that the competitive equilibrium allocation is not informationally feasible in the direct exchange when trades must satisfy the bilateral balance is true only when one imposes a limit to the number of periods allowed for transactions. Actually as Ostroy himself recognizes [29]: provided that individuals are willing and able to allow more time for trade, the opposite conclusion may be drawn;

b) *there is no sequential learning* because traders are not getting more information on the environment by trading. If you were to ask one of them what the knows about the state of the economy at any time he would always reply according to his informationally feasible trading rule: «I have no idea. All I know is that it belongs to the set of possible states, and I can, therefore, tell you what competitive equilibrium prices are, and what I would want if I had to balance my budget at those prices, and that I have made certain trades leading to my current position» [29];

c) *prices are known all the time and are the competitive equilibrium prices* (initial assumption);

d) *the sequence rnonetary economy considered by Ostroy is inessential* because the same equilibrium is attainable in a Debreu economy without money.

Alternatively one could point out that to attain the same equilibrium, additional time can be substituted for the loss of information incurred by avoiding money. Therefore either with money or without it a Pareto equilibrium may be achieved. So according to Hahn's standards, *money in Ostroy's model is inessential. It is also neutral, at least in the sense defined by Starr* [34]: «this does not mean that transactions are unaffected by the introduction of money. Rather the classical dichotomy means that the introduction of money does not affect the total net trade (i.e. final consumption) achieved by any trader».

---

[89] I, therefore, discard the objection that since Ostroy's economy does not represent a true intertemporal structure, to apply Hahn's criterion to judge Ostroy's achievements is meaningless. Hahn, in fact, distinguishes between sequence economies that are essential and sequence economies where transaction dates are not relevant: we will realize that Ostroy's scheme clearly belongs to this second set of economies. It is not, therefore, inconsistent, in spite of other substantial differences in the features of the model, to apply Hahn's main distinction to evaluate Ostroy's issue.

In spite of its inessentiality and neutrality, *money has, nevertheless, an informative content*, that is, it is an input of the information function of the traders. Ostroy is then able to define a role for money: which is a heuristically fruitful one, but not sufficient, in my view, to the foundations of a well grounded monetary theory. However, Ostroy's findings should not be underrated. Money, in Ostroy's model, is a means to disseminate information among traders. So one is led to reflect that to limit the information channels only to the price system, like in Walras, it makes the analysis poorer and unrealistic. Ostroy's model is not yet a good mirror of modern economy systems, as I will argue later, but it is already a step forward.

It is also possible to uncover some analogies between it and Hahn's work: in both systems money saves transaction costs, once one agrees that losses of time arising from differences in information are transactions costs. In both cases the role of money appears to be activated by the informative insufficiencies (although of a different type) of the price system. Furthermore, in both models uncertainty and the production side of the economy are excluded, thereby neglecting some important features of monetary economics. So one is induced to examine more carefully the relations between money and the information structures (described in a loose way) of the models taken into consideration and to extend the analysis to the Keynesian framework.

1.9 While in the Walrasian system the information-communication problem among micro unities is actually solved fully and exclusively by means of price-type information, with the help of the auctioneer and the recontracting assumption [2]; in Ostroy's more sophisticated world where exchange is a do-it yourself business, prices are no longer sufficient to provide full information about the quantities of different goods to be exchanged between different pairs of traders. In Radner's language, there is uncertainty about the behaviour of other transactors but that does not lead in this instance to unpredictability of relative prices, since they are known from the beginning, but only to imperfect information of quantities to be exchanged between different traders. The problem of getting more information may then be solved either by direct search between traders which is a very lengthy and implicitly costly procedure[90], or by the introduction of money as a record-keeping device which allows the monetary authority to monitor trading behavior.

However, when, as in Ostroy, transactions costs are not explicitly taken into account and prices are Walrasian equilibrium prices, all goods in principle and aside from restrictive assumptions are equally liquid and are indistinguishable from money as a source of effective demand[91].

But *money possesses an additional informative power when, dismissing the fiction of the auctioneer, we move farther into the world of imperfect information* which is marked with the existence of many bid and offer prices for the same commodity.

Then money plays a distinctive role in economic activity as a means of transactions. It is our experience that: «Transactors can acquire information most readily and at lowest cost about commodities that are most widely used and best known; so the prices of these commodities have the least dispersion.

Thus the role of money as a medium of exchange, as a transaction

dominating asset, results from the opportunities offered by the distribution of incomplete information and the search by potential transactors to develop transaction chains that save resources» [6]. Whereas money can be traded costlessly or almost so (at least under normal conditions) for all other commodities in the economy, this property is ineffective for other commodities because of their lack of liquidity.

The introduction of money therefore simplifies the transactions technology as traders may economise on gathering information. Providing the most economical vehicle of exchange, *money not only enters into almost every exchange, but also may become a constraint on effective demand.*

Money flows thus define a special class of information whose role is to transmit command over goods and services through the availability of purchasing power [21].

Under this perspective money flows may be regarded as relevant providers of market signals, together with price conditions. These signals are interrelated. In a monetary economy, positive, null or negative excess demands are revealed to economic agents by money flows and accordingly prices are confirmed or revised.

Moreover, in a monetary economy, following Clower, total net value of goods demanded cannot, in any circumstances, exceed the amount of money held by the transactor at the outset of the period and therefore money

---

[90] Saving [32] in order to derive the demand for money, has explicitly introduced transactions time as a choice variable for consumer in the utility function.

[91] When transaction costs are introduced in the way proposed by Hahn, money, which can be traded costlessly, may improve the equilibrium of the economy but, if one remains in the framework of a Debreu transactions economy or of an equivalent sequence economy, the role of money is still not essential. On the other hand, Hahn's model provides us with a definition of a binding role for money in the framework of a sequence economy, but the analysis is concerned with equilibrium properties of an exchange economy rather than focusing on effective demand.

constitutes a constraint on effective demand[92]. Clower's expenditure constraint [8], mentioned above, has been reformulated recently by Barro and Grossman [5] as a liquidity constraint with an intertemporal dimension.

This kind of analysis, like Tobin's model [36] which superimposes a liquidity constraint on a life cycle consumption function, provides a clue to the explanation of the relationship between consumption, income and liquidity and offers a view of the multiplier as an illiquidity phenomenon [25]. But money (liquidity) may also become a constraint on investment. As Davidson has underlined [10]: a The investment market can become congested through a shortage of cash... thus if the economy is at less than full employment, financial constraints rather than lack of real resources can be an effective barrier to expansion of economic activity». So money matters not necessarily because there is a stable demand of it, but because it is a necessary source of effective demand [20].

The preceding argument implies a denial of the appropriateness of budget constraints of traditional theory and makes clear the source of possible discrepancies between notional demand which is consistent with a situation of Walrasian equilibrium and effective demand determined by realized incomes and actual monetary constraints. In full equilibrium situations, notional and effective demand coincide, but as so on as information is costly and prices are dispersed or prevailing prices are not the same as market clearing prices[93], the working of the economy appears to hinge on effective demand.

1.10 Disequilibrium transactions may actually arise even when prices are announced by an auctioneer, if we drop the recontracting assumption. Then the lack of perfect price flexibility, which is the mirror of sub optimal information network, involves disequilibria and when one considers a production economy, may generate unemployment. This outcome is consistent with Clower and Leijonhufvud's reinterpretation of Keynesian theory. According to Leijonhufvud [24] «The revolutionary impact of Keynesian economics on contemporary thought stemmed in the main from Keynes' reversal of the conventional ranking of price and quantity velocities. In the Keynesian models price velocities are not infinite; it is sometimes said that the implications of the model result from the assumption that money wages are rigid. This usage can be misleading. Income constrained processes result not only when price-velocity is zero, but whenever is short of infinite».

Whereas in a Walrasian system perfect flexibility of prices is implied by the auctioneer mechanism and the groping search, which determine a market-clearing vector of prices; in Keynes' model, the adjustment proceeds faster in terms of quantities rather than of prices and this may involve some resource-unemployment. The difference in the adjustment mechanism which confers a strategic role to money in the Keynesian system is, therefore, amenable to differences in the information structures of the two systems. When we explicitly introduce the impact of an uncertain future over the present, as Keynes did, we get a very plausible explanation of why the economy should behave like this.

On one hand recognition of the capricious element in the course of future events, feeds the desire to avoid commitment of claims on resources and thereby provides a sensible motive for hoarding money. On the other hand the existence of transaction costs and of setting costs for the markets involve that *relevant futures markets may not be active and that implies by its own imperfect information about future prices*[94]. That, in turn, *owing to inelastic expectations implies slow current responses of prices to changes in demand and this may lead to some resources unemployment*.

When the uncertain future has an influence over present economic behaviour, the role of money as a store of value becomes intertwined with its transaction function and appears crucial for the comprehension of the working of the economic system. Indeed the transition from the Walrasian world where the knowledge of the intertemporal equilibrium price vector is perfect, to the Keynesian world where expectations rule, marks the end of a class of invariant states of the model and the attribution of a special status to money.

Instead of a known vector of intertemporal prices, the transactor is provided with expectations about marginal efficiency of capital and has a liquidity preference guiding him in the choice of his portfolio composition: money

---

92 Clower, however, seems no longer to be assuming that monetary income represents a constraint for transaction balances. According to Clower's latest view [9] the definition of what money actually is, must be derived theoretically in terms of the role and form of payments media in organized market: that is, in terms of explicitly postulated restrictions on trading alternatives that assign a special role to certain commodities as payment media. This approach could lead to include overdraft facilities and trade credit in the money supply. This was a major topic of discussion emerging from a paper presented by Clower at a Sheffield Money Seminar in 1970.

93 Foley and Hellwig call these two situations «trading uncertainty».

94 One could point out that even spot markets might not give appropriate signals, because of lack of organization or because of monetary constraints forbidding the transformation of potential demands into effective demands. The former reason is generally imputable to the existence of set-up costs and, therefore, does not alter substantially the argument of the text; while the latter is a consequence of the role of money, which we are trying to explain by price uncertainty.

then besides satisfying his transactions needs, becomes one among many alternative assets to keep, following the suggestions of the theory of portfolio choice.

But equilibrium might actually become informationally unfeasible owing to continuous revisions of expectations and in the absence of the required organized forward markets which can give appropriate signals for the transactor[95]. The behaviour of transactors within a Keynesian framework does reveal the formal complexity of a model where both market prices and expected prices, money flows and realized and anticipated quantities have an impact on transactors' decisions. A Keynesian model that keeps in due account the impact of expectations over the present state of the economy is, therefore, characterized by a complex information structure which expands beyond price and money flows signals. Moreover, this structure increases historically with the expansion and development of real processes. This may explain why an adequate formalization of the Keynesian system is far from having been achieved.

1.11 Some progress has, however, been made. The formalization of the connection between trading uncertainty and the holding of liquid assets has been the object of a recent paper by Foley and Hellwig [11] suggesting some revision of Clower's treatment of Keynes' theory. With a particular concern for the relationship between consumption income and money, and more in line with Clower's analysis, this issue is also debated by Barro and Grossman [5] who introduce many simplifying assumptions in order to clarify the impact of the liquidity constraint on household behaviour in an intertemporal framework. Clower's insistence on the role of money as a medium of exchange has led him to neglect the impact of changing trading uncertainty over the holding of liquid assets and to assume that current expenditure is constrained in a mechanical one to-one fashion by realized monetary income.

Foley and Hellwig reject this view, since according to their scheme «the agent sees its budget constraint not in terms of current expenditure equalling current income, nor as the present value of lifetime expenditures equalling the present value of lifetime income, but as never being able to incur negative money balances». [11] This is because the agent faces a residual and unhedgeable uncertainty as to its opportunities to buy and sell and confronts these opportunities in a sequence of decisions, not all at once at some initial trading session. Therefore the risk of inability to sell affects the agent's consumption spending even in periods when he is lucky enough to achieve a sale at market price. Foley and Hellwig's model provides a better motivation than the one advanced by Clower for the consumption behaviour of the individual, while at the level of aggregates the scheme may be shown to be consistent with Keynes' suggestions concerning the existence of disequilibria positions.

Moreover, the conventional theory may be recognized as an asymptotic case of the model for which trading uncertainty is ignored. The remarkable results of the analysis should not, however, make us forget its main limitations. The model has only one asset and it is, therefore, a model of saving rather than of preference for liquidity assets. The authors' qualification that the motive of saving is precautionary does not fully avoid the objection. Any consideration of decentralized price formation in a sequential trading framework is missing and, therefore, no solution is given to the problem of explaining the connected phenomena of price formation, money holding and unemployment of resources inside a general equilibrium structure.

1.12 As Hahn reminds us great is our intellectual debt towards Radner's sequential approach under uncertainty [31]. At the same time, one may seriously doubt that attempts to further improve the Radner's approach by introducing explicit transactions costs and looking for the existence of an equilibrium, even if successful, will give us a much better picture of the working of a real economy. So one might feel a bit discouraged with the present state of monetary theory. I do think, on the contrary, that some important results have been achieved. Not only have we succeeded in clearly defining the impact of money on the economy by a comparison with barter conditions, before considering price-uncertainty:

b) money allows specialization;
c) money bypasses ineffectiveness of barter transactions;
c) money helps reduce explicit or implicit transactions costs; but we have come to grasp that when transactions are costly, money may play a constraining role in the working of the economy in the sense that a Pareto-efficient equilibrium cannot be achieved if some agents run out of stocks of it at any step of a sequential trading process.

We have also learnt that costs of transactions and costs of information gathering are the main reasons for an imperfect knowledge of prices in the short run. The introduction of price uncertainty in the economy has enlarged and modified our notion of the constraining power of money, making clear the distinction between notional and effective demand. We have then realized that the introduction of imperfect information about prices confers to money the additional distinctive power of source of effective demand. Money flows, therefore, become relevant providers of market signals together with price conditions.

---

95 Compare with Arrow and Hahn [3], p. 369.

At this point, reversing Ostroy's statement that a role for money must be defined before the introduction of price uncertainty, one could argue that only when the above situation is considered is there opportunity to perceive the full implications of a monetary economy. So finally one is turned back to the microeconomic foundations of Keynes' theory that tentatively provides us with a description of the working of the economy, consistent with the previous position: progress in this field will be helpful not only in precising the behaviour of money both at an individual level and for the working of the whole economy, but also will provide a fruitful framework for the theoretical foundation of monetary policy[96].

[96] I wish to recall in this connection that a major topic of discussion at the 1970 Sheffield Money Seminar, concerned Clower's view that the ways in which new supplies of money were injected into the economy led to different initial results and different final results, a point overlooked by the monetarists. Certainly present microeconomic interpretation of Keynesian theory and future progress in the field can become very helpful in providing a consistent rigorous background for economic policy decisions.

REFERENCES

[1] ALCHIAN A.: *Information Costs, Pricing, and Resource Unemployment*, in «Microeconomic Foundations of Employment and Inflation Theory», Macmillan, London 1971.
[2] ARCELLI M.: *Information Structure in the Walrasian Theory of General Economic Equilibrium and in the Keynesian Model*», «L'industria», 1972, n. 3-4.
[3] ARROW-HAHN: «General Competitive Analysis», Holden-Day, 1971.
[4] BARRO R., GROSSMAN H.: *A General Disequilibrium Model of Income and Employment*, «American, Economic Review», Maich 1971.
[5] *Consumption, Income and Liquidity*; discussion paper, presented at the conference, «Equilibrium and Disequilibrium in Economic Theory» Institute for Advanced Studies, Vienna, July 1974.
[6] Brunner-Meltzer: *The Uses of Money: Money in the Theory of an Exchange Economy*, «American Economic Review», December 1971.
[7] CLOWER R. W.: *«The Keynesian Counter-Revolution: A Theoretical Appraisal*, in «Monetary Theory», ed. Clower, Penguin Books, 1970.
[8] *Foundations of Monetary Theory*, in «Monetary Theory», ed. Clower, Penguin Books, 1970.
[9] *Theoretical Foundations of Monetary Policy*, in «Monetary Theory and Monetary Policy and the 1970», ed. Clayton, Gilbert, Sedgwick, Oxford University Press, 1971.
[10] DAVIDSON P.: *Money and the Real World*, Macmillan 1972.
[11] FOLEIY-HELLWIG: «Assets Management with Trading Uncertainty», Working Paper 108, M.I.T., April 1973.
[12] GRANDMONT-YOUNES: *On the Role of Money and the Existence of a Monetary Equilibrium*, «Review of Economic Studies», Vol. 39, July 1972.
[13] *On the Efficiency of a Monetary Equilibrium*, «Review of Economic Studies», Vol. 40, April 1973.
[14] GROSSMAN H.: *Effective Demand Failures: a Comment*, Forthcoming in «Swedish Journal of Economics».
[15] HAHN F.: *Equilibrium with Transaction Costs*, «Econometrica», Vol. 39, n. 3, 1971.
[16] *On Transaction Costs, Inessential Sequence Economies and Money*, Technical Report No. 64, August 1972, Encina Hall, Stanford University, «Review of Economic Studies», October 1973.
[17] *On the Foundations of Monetary Theory*, in «Essays in Modern Economics», ed. M. Parkin, Longman 1973.
[18] HELLER: *Transactions with Set-up Cost*, «Journal of Economic Theory», Vol. 4, 1972.
[19] HELLFR-STARR: «Equilibrium in a Monetary Economy with Non-Convex Transactions Costs», Technical Report, September 1973, Encina Hall, Stanford University.
[20] HOWITT P.: «On Revising Keynesian Economics» discussion paper, Department of Economics, University of Western Ontario, Nov. 1973
[21] KORNAI J.: «Antiequilibrium», North-Holland, 1971.
[22] KURZ M., «Equilibrium with Transaction Cost and Money in a Single-Market Exchange Economy», Technical Report, No. 51, January 1972, Encina Hall, Stanford Univers.
[23] «Equilibrium in a Finite Sequence of Markets with Transaction Cost», Technical Report, No. 52, February 1972, Encina Hall, Stanford University. «Econometrica», January 1974.
[24] LEIJONHUFVUD A.: «On Keynesian Economics and the Economics of Keynes», Oxford University Press, 1968.
[25] *Effective Demand Failures*, «Swedish Journal of Economics», March 1973.
[26] MCGUIRE-RADNER: Ed., «Decision and Organization», North-Holland, 1972.
[27] NIEHANS J.: *Money in a Static Theory of Optimal Payments Arrangement* «Journal of Money Credit & Banking», November 1969.
[28] *Money and Barter in General Equilibrium with Transactions Costs*, «American Economic Review», Dec. 1971.
[29] OSTROY J.: *The Informational Efficiency of Monetary Exchange*, «American Economic Review»x, September 1973.
[30] RADNER R: *Competitive Equilibrium under Uncertainty*, «Econometrica», Vol. 38, No. 1, January 1968.
[31] *Existence of Equilibrium Plans. Prices and Prices Expectation in a Sequence of Markets*, Technical Report, No. 5, University of California, Berkeley, 1970, «Econometrica», March 72.
[32] SAVING T.: *Transactions Costs and the Demand for Money*, «American Economic Review», June 1971.
[33] SHUBIK M.: C*ommodity Money, Oligopoly, Creit and Bankruptcy in a General Equilibrium Model*, «Western Economic Journal», n. 4, 1972.
[34] STARR R.: *The structure of Exchange in Barter and Monetary Economies, «Quarterly Journal of* Economies», May 1972.
[35] STARRETT D.: «Inefficiency and the Demand for Money in a Sequence Economy», 1972. «Review of Economic Studies», October 1973.
[36] Tobin J.: Wealth, Liquidity and the Propensity to Consume, in «Human Behaviour in Economic Affairs», Essays in Honour of George Katona, Elsevier, 1972.

Federico Arcelli **Giorgio Di Giorgio**

SOME ISSUES ON APPLICATION OF CONTROL THEORY TO BALANCE OF PAYMENTS ADJUSTMENT: A NEW TREATMENT OPPOSED TO THE ACCEPTED THEORY [97]

The dominant theory of balance of payments adjustment (known as the Tinbergen-Mundell Approach) has recently become the object of serious criticism. Not only, as it was observed since long, it is inadequate to face the problems of less developed countries; but, after the oil crisis with its tremendous impact on balance of payments equilibrium of many industrialized countries, it appears questionable and lacking even for the capitalistic world.

The dissatisfaction with the present state of the theory, however, has not arisen merely from political or empirical facts, but it is deeply rooted in the analysis of its own foundations. Briefly, the accepted theory which relies on a targets-instruments approach in order to achieve both "internal and external balance" of an economy, suffers from important limitations: it does not maximize any definite social welfare function and it may result destabilizing for the economy even when instruments are appropriately paired with the objectives on which they have the most influence.

The present paper after a summary of Mundell's strategy to obtain external and internal balance, with a comment on its inadequacies, deals with an alternative approach based on an application of control theory to problems of balance of payments adjustment. While the former approach follows a logic, which is essentially static, the latter consists in a process of intertemporal maximization of a social welfare function, from which prescriptions for balance of payments are derived. A mathematical model illustrates the problems and outlines the main results obtained from this new approach.

In a famous paper: "The appropriate use of monetary and fiscal policy for internal and external stability", Mundell stated that it was possible to attain both internal stability and balance of payments equilibrium by an appropriate choice of the mix of fiscal and monetary policy. Internal balance is defined as equilibrium between aggregate demand and supply at full employment. The demand is assumed to be inversely related both with fiscal policy (measured by high-employment budget surplus *g)* and monetary policy (measured by the interest rate *r*).

It is then possible to trace a locus of points of internal balance in the g/r piane, downward sloping (IB). Also the balance of payments equilibrium depends on interest rate and budget surplus: capital flows are assumed to be responsive to interest rate differentials and the balance of trade is inversely related with the level of domestic expenditure and therefore depends on *g*. The locus of points of external balance in the g/r piane (EB) is steeper than IB, as the increase of interest rate requires a decrease in the budget surplus to achieve the external balance more consistent than the one, which is necessary for internal balance (capital flows being responsive to interest rate differentials). Hence the curves IB and EB must intersect at some point in the g/r piane where it is possible to attain both internal and external balance.

Mundell stresses that when the exchange rate is fixed and monetary and fiscal policy can be used as independent instruments to attain the two targets, monetary policy ought to be aimed at external objectives and fiscal policy at internal goals. The failure to follow such prescription might worsen the disequilibria situation preceding the policy changes. In other words: "policies should be paired with the objectives on which they have the most influence. If this principle is not followed, there will develop a tendency either for a cyclical approach to equilibrium or for instability". Thus Tinbergen's principle that in order to attain a given number of independent targets it is necessary to dispose at least of an equal number of instruments, refers only to the problem of existence of a solution to the system. It does not guarantee that any given set of policy will lead to that solution. Mundell's strategy is an integration to Tinbergen's principle, investigating the stability properties of a dynamic system.

The so-called Tinbergen-Mundell approach has largely dominated the theoretical debate of economic policy of western economies during the sixties. The prescriptions to follow were that "a surplus country experiencing inflationary pressure should ease monetary conditions and raise taxes (or reduce government spending), whereas a deficit country suffering from unemployment should tighten interest rates and lower taxes (or increase government spending)". To a certain extent this policy has been also implemented. A main implication of this

---

97 From: *Èconomie Appliquèe*, Archives de l´ISMEA, Droz, Gènève, Tome XXX, 1977/4. See also M. Arcelli, Sviluppo Economico, scambi internazionah e crisi monetaria, Atti della 15ª Riunione Scientifica della Societa Italiana degli Economisti, Giuffrè, 1975.

approach is the neglect to consider as a separate target the balance of current accounts, since attention is devoted to global external equilibrium. Consequently current accounts deficits may be countered by capital inflows, thereby progressively raising foreign indebtedness. The possibility to attain external balance without restoring trade balance, also encourages a slack domestic policy, which endangers price-stability. Briefly to attain the two targets (internal and external balance) in a given moment does not imply that a welfare function is maximized; neither the approach is necessarily consistent with long run objectives.

The inadequacy of Mundell's strategy for less developed countries had always been recognized. For these countries the trade deficit is a structural characteristic and interest rate differentials are not sufficient to attract the needed foreign investments. More recently, however, the impact of oil crisis has raised many questions about the feasibility of Mundell's approach also for many capitalistic countries. Not only the so-called dominant theory appears concerned with a global external balance, neglecting current accounts situations and long run objectives; but it also shows powerless when confronted with the huge deficits induced by sudden crude oil prices growth. Stagflation is a farther complication in the working of modern economies, which makes out of date the simple rules prescribed by Mundell.

So it is not surprising that in the present situation Mundell's theory is under severe scrutiny, while alternative approaches to balance of payments adjustments are carefully considered. The criticism of Mundell's approach however reached its zenith well before the oil crisis. At the Sheffield Seminar on Monetary Theory and Monetary Policy in the 1970s, held in September 1970, Williamson presented a paper "On the normative theory of balance of payments adjustment" that on one hand pushed farther the dissatisfaction with the targets instruments approach on theoretical grounds; while providing on the other hand an alternative approach based on an application of control theory to balance of payments adjustment.

Williamson argued that international capital movements are better described by a stock adjustment theory than by a flow theory. But even if long-term capital movements were sufficiently well described by a flow theory to save the model as a positive model, Mundell's strategy would deserve no status as normative economics. Should one consider a multiperiod analysis, then one would realize that in many cases the trajectories traced out by Mundell's prescriptions from differing initial conditions necessarily diverge.

A less competitive economy will have a less favourable current account at the designated unemployment level. The internal balance will therefore require a more relaxed fiscal-monetary policy mix; whereas on the contrary external balance will require tighter interest rates to attract more capital. Fiscal policy will be looser, monetary policy tighter and foreign debt greater than in the case of a more competitive economy. IB would be pushed inwards and EB outwards in the g/r piane. This trend would be reinforced in the following periods because the interest burden of foreign debt and the progressively falling competitiveness of the economy owing to a lower investment-product ratio. So IB would be pushed more and more inwards and EB outwards. Instability would be the final outcome of Mundell's strategy[98].

Indeed Professor Johnson had already remarked that a policy of manipulating the capital account to counter current accounts deficits was scarcely likely to lead to an efficient pattern of international investment. Two partial conclusions therefore emerge at this point: first the targets instruments approach proposed by Mundell does not maximize any welfare function; second an efficient mechanism of balance of payment adjustments must explicitly refer to the current account.

Williamson in his paper shows that the optimal trajectories that result from maximizing intertemporal welfare starting from different initial conditions converge asymptotically in a wide range of problems. Hence the limitations of Mundell's approach can be overcome according to Williamson, by constructing a social welfare function that is a reasonable reflection of the ultimate ends of the economic system and maximizing this, subject to the constraints imposed by the positive economy. Appropriate values of proximate objectives are then obtained as a by-product of intertemporal maximization.

Whereas a targets-instruments approach does not necessarily reconcile and may even bring about conflicts between short run and long run objectives in national economy, Williamson's approach renders consistent to-day decisions with to-morrow measures. Moreover it gives due attention to current account balances and hence to the change in foreign indebtedness which are neglected in the simplified Mundellian approach. Following this new stream, in this paper it is described a model, which may be considered a variant of Williamson's approach.

---

98 This kind of instability is different from the short run in stability discussed by Mundell and related to the fact the instruments are not rationally paired with objectives.

Federico Arcelli **Giorgio Di Giorgio**

I propose to maximize intertemporally a social welfare function where the utility depends on the level of consumption and on income distribution, subject to income determination equations, which give substantial weight to the constraints of an open economy. The process of maximization consists in an application of the maximum principle to the objective function where consumption and labour share are treated like control variables, whereas foreign indebtedness, capital stock and index of competitiveness are state variables. The model and its asymptotic solutions are exposed in details in the following pages. The adjustment process is to be interpreted as one where-by the policy makers return the economy to its asymptotically optimal trajectory from initial conditions off this path.

Given the complexity of the model I have limited the analysis to the steady state solutions: the presence in the model of three state variables excluded the customary solution by phase diagram. The results obtained are in line with, but not the very same Williamson's conclusions; comments on them follow the model. Here I want to point out that simplifications and improvements may be found in the formulations of the objective function and in the differential equations. In particular the price equation has been made consistent both with cost-push inflation situations and demand-pull inflations. Besides, and I believe with more adherence to reality (at least for Italy), the model considers changes in foreign debt rather than in foreign assets accumulation. The model, however, ought to be considered only as an instance of this new kind of approach. It lacks in realism in many respects and further improvements should be introduced with due consideration for monetary and financial aspects, for expectations and by suitable disaggregation.

Notation

INSERIRE TABELLA

Inserire Formula

X (exchange rate): additional control variable (not necessarily used).

Inserire Formula

Subject to:
Income determination equations:

Inserire Formula

Hence

Inserire Formula

Differential equations:

Inserire Formula

From (8) and (9):

Inserire Formula

From (11) and (7):

Inserire Formula

Hamiltonian

Inserire Formule

Optimum conditions (assuming interior solutions) and derived steady state solutions are

Inserire Formule

The marginal utility of consumption must be equal to the imputed price (shadow price) of foreign debt, which accrues from a worsened trade balance. I is supposed to be independent from, C: the consistency of the optimum conditions of the model implies in this case that the accumulation of capital is non-optimum. If, on the other hand, at the limit any change in C is compensated by an equal change of opposite sign in I, equation (14) is replaced by:

Inserire Formule

The marginal utility of consumption must be equal to the imputed value (shadow price) of marginal capital accumulation.

Inserire Formule

The marginal utility of a preferred income distribution must be equal to the marginal disutility (shadow price) of foreign debt, which accrues from a worsened trade balance due to a loss of competitiveness, plus the imputed value of a lower competitiveness determined by increased internal prices. In steady state the last term of the right side of equation (15) vanishes. The canonical equations for the costate variables:

Inserire Formule

In equilibrium (steady state) equation (16) implies f'(K) > r + d provided $q_3 > 0$: that is the marginal productivity of capital is greater than the rate of time preference plus the depreciation rate (modified golden rule). Hence the accumulation is sub optimal. If however the objective functional has become insensitive to small changes of the value of the state variable D, then q3* = 0 and the modified golden rule is verified

Inserire Formule

In equilibrium (steady state) the third term of the right side of equation (17) varishes. Equation (17) then reduces to:

Inserire Formule

that implies that the shadow price of foreign asset times Inserire formula must be equal to the shadow price of competitiveness times the rate of time preference. Notice that in steady state $q_2 > 0$ implies that competitiveness is undesirably low.

Inserire Formule

In steady state equation (19) implies that the level of foreign debt is such that the marginal cost is equal to the rate of time preference. Additional control variable

Inserire Formule

that gives

Inserire Formule

hence

Inserire Formule

In steady state the rate of change of competitiveness and the rate of change of exchange rate are null. Equation (20), to be consistent, implies therefore that $q_3 = 0$. With reference to equation (16) that implies that the modified golden rule is verified and accumulation of capital becomes optimal. When the internal rate of inflation, during the transitory period before asymptotic solution, is not consistent with the international rate of price growth, then the accumulation of capital remains sub-optimal; but with adequate exchange rate manipulation optimal conditions are re-established. The main results of the model may be summarized as follows:

1. In optimum conditions (equilibrium), it is necessary to attain a trade surplus sufficient to pay for interests on foreign debt. According to equation (19) optimal foreign indebtedness depends on domestic rate of time preference and on marginal cost of foreign debt. A rise in the marginal cost of foreign debt would decrease the level of foreign indebtedness in steady-state: by equation (14) that would imply larger steady state consumption and by equation (15) a more preferred income distribution. An increase in the rate of time preference would on the other hand raise the steady-state foreign indebtedness and conversely decrease steady-state consumption and worsen income distribution. The trade surplus in steady-state would be lesser in the former case and greater in the latter. It must be underlined that a change in the initial level of foreign debt D would not have any influence on steady state solution. At time zero foreign payments deficits would be larger for a rate of time preference and consumption too would be larger.

A rise in the marginal cost of foreign debt would instead require during transition less payments deficits or larger surplus if the initial $D > D^*$ (steady-state indebtedness).

2. In equilibrium the marginal utility of consumption is equalled either to: a) the imputed value of foreign debt, which accrues from a worsened trade balance (equation (14)); or b) to the shadow price of marginal capital accumulation (equation (14bis)). This latter condition is active when accumulation is optimal. In that case equation (16) states that the net marginal productivity of capital is equal to the rate of time preference. In accordance to equation (19) also the marginal cost of foreign debt must be equal to the rate of time preference: thus in steady-state the level of capital accumulation and consumption are strictly related to foreign indebtedness. A change in the rate of time preference or in the marginal cost of foreign debt will therefore vary both capital accumulation and consumption of steady state inversely as foreign indebtedness.

3. Often however, the accumulation is sub-optimal: the marginal productivity of capital is greater than the rate of time preference plus the depreciation rate (modified golden rule). That happens when the desired internal rate of inflation is superior to the international one. If fixed exchange rate is maintained the country will suffer from welfare loss: it cannot achieve the mix of absorption and inflation that it would prefer, given the balance of payments constraint, and will indulge in sub-optimal accumulation since it is forced to accept a lower level of real income. With adequate exchange rate manipulation, however, the required degree of competitiveness is attained and thereby conditions for optimal accumulation are re-established. Notice that low competitiveness and sub-optimal accumulation are two faces of the same problem (equation (18)).

4. Foreign currency reserves have not been explicitly mentioned in the working of the model: in steady state their role is null, whereas their function may be relevant during transition. They enable a country to avoid adjusting to transitory disturbances or help to optimise the rate of adjustment. Under fixed exchange rate, they allow to avoid income deflation under the internally optimal level while a gradual improvement in competitiveness is achieved. Managed floating, an active exchange rate policy, fulfils a similar task: it reconciles external and internal objectives (like attainment of a preferred income distribution) and, as we have seen, it also creates conditions for optimal capital accumulation, allowing for a higher welfare level.

5. The intertemporal maximization of a social welfare function excludes the indiscriminate recourse to capital inflows as a permanent means to offset current accounts deficits. Indeed manipulation of capital accounts has no influence on desirable steady state solution. This conclusion does not prevent the use of interest rate policy to attract capitals during transition as a form of quasi-adjustment to solve temporary scarcity of foreign currency. This approach however must carefully take into account the distortions it causes to the desirable consumption-investment mix.

6. Among all possible exercises in comparative dynamics, it is interesting to analyse the effects of a change in the parameter n (which depends on marginal propensity to import). It is possible to show that an increase in the marginal propensity to import would reduce steady-state consumption. See Appendix for proof.

7. The whole story—optimal trajectories of the variables and steady state solution—hinges on how the welfare function has been specified, that is on which objectives have been included in the function and the weight attributed to them. The difficulty to specify a welfare function, which reflects the ultimate ends of the society accurately and realistically, may induce some scepticism on the validity of the previous analysis. The constraining functions too, are too schematic in order to provide an adequate

framework for decision-making. I think nevertheless that the previous approach allows at least a critical assessment of the existing doctrine on desirable adjustment policies of balance of payments. In this perspective I consider a major result to have pointed out the fallacies of Mundell's prescriptions, by showing that capital inflows have no influence on steady state solutions: therefore they cannot be employed as permanent measures to offset current account deficits. Mundell's strategy may be on the contrary seriously distorting and can contribute to bring the economy farther off the optimal path, thereby increasing disequilibria. The welfare maximization approach is distinctly superior in this respect as it makes short run decisions consistent with long run objectives: it outlines a method which ought to be followed in any decision making

Appendix

A proof may be given showing that by equation (19) in steady state the level of foreign debt depends only on the rate of time preference and on the marginal cost of indebtedness. Steady state consumption is defined by equation (14). It follows that the level of consumption depends inversely on the value of the parameter *n*. An increase in *n* (determined by an increase in the propensity to import) will therefore produce a decrease in the steady-state consumption.

Inserire Formule

deriving with respect to *n* gives

Inserire Formule