TOWARDS A THEORY OF 'RATIONAL' DISTORTIONS
TOWARDS A THEORY OF 'RATIONAL' DISTORTIONS

by

GEOFFREY NEWMAN *

Many traditional models of economic behavior generate their results by way of a rather simple methodological strategy. This is to take a prototype Walrasian competitive equilibrium structure — built on the explanatory power of tastes and technology alone — and then to impose on it a set of exogenous institutional constraints which "distort" the conditions for Walrasian maximization. Thus, traditional versions of Keynesian macro-theory circumscribe an otherwise Walrasian framework by assuming the distortion of fixed wage or price contracts. Traditional monopoly theory does likewise by positing an exogenous barrier to entry. Furthermore, a wide variety of micro/policy models follow in the same tradition insofar as they assume some distortion in the private sector — the persistence of externalities, an imperfection in the capital market, the absence of certain relevant markets altogether — and/or one which emanates from the public sector — a distortionary tax, tariff, or rate of return regulation.

Many theorists have been prepared to endorse models of the above form simply on the grounds that real-world economies are observed to contain the institutional distortion in question, so that any economic model which makes a pretense to "descriptive realism" should reflect this fact in its posited constraint set. As typified by the methodological structures of the new "classical" macroeconomics (see Lucas, 1980, 1981; Lucas and Sargent, 1979; Barro, 1977, 1979; and Friedman, 1970), it has now become commonplace however to criticize such formulations (and especially those which assume any fixed "price") as explanatorily incomplete, that they take as given constraints which should be properly explained as endogenous.

* Drew University, Department of Economics, Madison, N.J.

I thank my former colleagues at the University of Toronto and the University of California, Santa Barbara, and seminar participants at Brussels, Leuven and Simon Fraser Universities for helpful but non-culpable comments on earlier versions of this paper.
outcomes of the rational choices of all individuals. By this argument, any general equilibrium model which posits distortionary institutional rigidities but fails to explain them to an appropriate microfoundational can be seen only as "behaviorally arbitrary", or "not consistent with rationality", since it fails to illuminate the issue of why maximizing individuals would sustain a distortionary institutional regime if they could gain by altering it.

This paper accepts the above critique and thus defends the view that institutional distortions can be introduced legitimately into general equilibrium analysis if and only if their existence can be theoretically explained. It also defends the position that the achievement of such explanations constitutes one of the two central prongs of a basic research programme to produce theories of general equilibrium which are "complete in individualistic terms", the other being to account for the formation of expectations as rational. The problem addressed here concerns exactly how the above objectives might be achieved together; namely, by what methodological strategy is it possible to successfully explain any institutional phenomena as distortionary in the context of models which are complete in individualistic terms?

Little headway has been made on this problem. No matter whether one examines the literatures on non-competitive market structure (Demsetz 1968; Baumol, Panzer and Willig 1982), barriers to entry (von Wizackerley, 1980), the presence or absence of markets as such (Demsetz, 1968; Williamson, 1975; Gould, 1980), externalities (Coase, 1960; Buchanan and Stubblebine, 1962; Arrow, 1969; Dahlman 1979), business cycles (Lucas, 1975), the existence of long-term contracts (D.F. Gordon, 1974; Azariadis, 1975; Townsend, 1982), or the theory of second-best (McKee and West, 1981), the analytical results are remarkably similar. Attempts to rationalize the sustainability of distortional institutional constraints have succeeded only by making these constraints an endogenous outcome of a more general/sophisticated Walrasian maximum problem, where the sophistication comes from adding various transactions, adjustment or information limitations, and individual attitudes towards these new "costs", to an otherwise standard scenario. Of course, since Pareto optimality still prevails and agents are still seen to maximize gains from trade, relative to the enlarged specification of tastes and technology, it follows by definition that the (now-explained) institutional distortions cannot in fact be distortionary. Here all of the exogenous distortions of traditional theory are rendered only apparent, not real, when explained by Walrasian gives.

That typical programmes to render institutional constraints endogenous do not lead to a theory of rational distortions but to the forfeiture of distortionary phenomena altogether is not a very satisfying result; it can hardly improve on the traditional strategy of positing distortions on grounds of descriptive realism but not explaining them. The reason for this outcome, however, is evident: that the concept of "rational choice" postulated to complete these traditional structures entails that the only sustainable institutional setting is one in which agents maximize gains from trade, relative to Walrasian gives alone. Under this concept, complete Walrasian rationality, if any programme to render all institutional constraints endogenous is successful, then it can leave no constraints as distortionary; hence, institutional distortions can appear only in models which are logically incomplete.

This paper advocates the only line of attack to deal with this methodological dilemma: to specify a Non-Walrasian concept of complete rationality which is logically consistent with those distortionary institutional situations where gains from trade are not at a Walrasian maximum. It is only through such a concept that the objective of illuminating distortionary institutional phenomena can be wedded to the desire to produce general equilibrium models that are complete with respect to their microfoundations. As will be shown, the ultimate analytical objective here is to demonstrate the possibility of explaining as (completely) rational a situation where an individual or group conjecture that an institutional constraint could be gainfully changed, have the resources to effect the change, but still refrain from making the attempt.

Since immediately it might be thought that a sophisticated application of game-theoretic techniques would be all that is necessary for this task, it is important to understand why this is not the case. Many game theory equilibria do imply unexploited gains from trade (e.g., non-repeated Nash) but these "rational" equilibria are actually achieved only by a subtle compromise of institutional rationality — by failing to explain the determination of the "rules of the game", by exogenously imposing a constraint on agents which forces them to continue to play the game rather than change it, etc. The
approach put forth here must dig deeper insofar as it makes it a methodological imperative not to accept models with any unexplained institutions as complete explanations. The thrust of this approach is to start from the process by which agents potentially can change institutional constraints (under limited knowledge) and then to specify a “reform technology” which can explain the failure to effect a potentially-gainful change in any one institution as rational relative to the problem of gainfully adjusting all sub-optimal institutions together.

Little need be said on the practical importance of this undertaking. For many years, economists’ interest in institutional distortions has been matched by an equal preoccupation with institutional (governmental) policies to correct them. Largely due to Keynesian influences, the principle underlying this traditional orientation is that, if institutional distortions are self-evident by inspection, then an active role for government policy must be self-evident as well. Moreover, the basic output of this research tradition is a comparative static analysis showing how the real variables of a general equilibrium structure can be affected by different exogenous settings of government policy instruments, given at least one private sector institutional distortion. It is possibly unfair to remark that the only reason that these analyses can rationalize the role for active policy, and generate sustainable effects on real variables, is because they set up their relevant general equilibrium structures in such a way that they must be incomplete. The fact however remains that if all institutional constraints associated with these structures were to satisfy Walrasian rationality, and expectations were also rational, then in principle there could be no active role for government policy in affecting a Walrasian real equilibrium (Lucas, 1972; Sargent and Wallace, 1976). Here not only would the private sector distortions cited be rendered only “apparent” – which would undermine the role for active policy by itself – but also all institutional variables constituting government policy already would be predetermined by tastes and technology (as “optimal”).

Once again, it becomes necessary to argue for a Non-Walrasian concept of complete rationality, which can account for the existence of sustainable private sector distortions in the first instance, and which thus allows the logical possibility of rational, active policy behavior to be explained alongside. In lieu of this, there are no interesting methodological options open for policy questions. More generally, without a Non-Walrasian concept of complete rationality, there is no logical way to break out of the entrenched dichotomy between market-clearing (classical, policy passive) approaches to economic theory on the one hand and non-market clearing (Keynesian, policy activist) approaches on the other.

In Part I of this paper, the arguments introduced here are analyzed in a more general methodological framework. In Part II, it is shown that there exists no obvious way in which to generate “rational” institutional distortions from private sector informational distortions. In Part III, the elements of the solution to this problem are outlined.

I. The “Completeness” of General Equilibrium Explanations and the Role for Institutional Distortions

1.1. Basic Definitions. It is taken as axiomatic that the objective of any general equilibrium (GE) theory is to provide a complete explanation of the set of observed quantitative actions of all individuals in an economy over all production/consumption activities over time (denote by $\chi$). To accomplish this, any GE structure must posit a set of (dated) exogenous variables which is deemed to be just sufficient to explain $\chi$. By the definition of GE itself, it follows that $\chi$ can be explained successfully only as set of sustainable coordinated aggregate states ($\chi^*$), such that the given values of the exogenous variables plus only entail the GE solution quantities $\chi^* = \chi$.

The central issue in any methodological appraisal of GE explanation then concerns what classes of exogenous variables permitably might be used to explain $\chi$ (as $\chi^*$). With reference to the past tradition of GE debate, define the following general sets of variables as potential candidates for exogeneity (see Buchanan, 1962; and McKee and West, 1981 for a similar taxonomy):

- $N$: the set of “ultimate” limiting properties of physical nature (as distributed over all individuals);
- $S$: the set of limiting properties of social/institutional entities (as distributed over all individuals);
- $I$: the set of properties of all individuals.

The set, $N$, typically would be identified with a set of technology/endowment constraints which are deemed to be “ultimate” constraints on behavior in the sense that they cannot be further transformed by the productive actions of any or all individuals. The set, $I$, normally would be identified with the set of (all) individual “tastes”, these implicitly incorporating each agent’s “rationality” (e.g., the aim of maximizing utility or profits). Finally, in economic models, the set, $S$, traditionally would be identified with a set of “relative prices” of activities. It is also evident that $S$ must be seen to
include the wide variety of legal, normative, and regulatory structures observed in any society, which may operate in conjunction with, or as substitutes for, a price mechanism (see Arrow, 1969). Specifically, S must incorporate the mechanisms of government policy, the contractual arrangements underlying firm, factor, and market exchange, as well as other institutional mechanisms for defining property rights, incentives, and the conditions of information transfer (see Newman, 1976; Shetter, 1981). It is apparent that many different types of GE structure are potentially available to the theorist, depending on which elements of the above sets he chooses as his exogenous "fundamentals". The crucial question to be considered at this point is: What choice of exogenous elements satisfies individualistic postulates of explanatory completeness?

1.2. The Individualistic Postulates. — To construct an answer to the above question, it is useful to begin from three propositions: (i) in any individualistic explanation of \( \chi \), elements of the set, I, must be assigned exogenous status; (ii) in any non-trivial GE setting (i.e., one which involves scarcity and positive prices), the set, N, cannot be empty; and (iii) the most general GE structure would appear to be one which draws explanatory components from all three of N, S, and I.

There can be little debate about (i) and (ii). An individualistic explanatory structure logically can accept elements of I as exogenous, elements of N as exogenous (so long as they do not override completely the explanatory role for I), and all behavioral constraints on individuals which follow from N alone and N and I together. The important issues therefore concern (iii), since it is this purportedly general specification — that assigns explanatory power not only to tastes and technology, but also to institutional constraints — which is typically criticized as being behaviorally arbitrary.

The anomaly implied here is not to be taken as problematic; it serves only to highlight the traditional "individualistic" interpretation of the nature and status of social institutions (see Popper, 1945; Hayek, 1948; Agassi, 1960, 1973; also Arrow, 1959; and Buchanan and Tullock, 1952). This

view postulates that, since all social institutions are man-made and man-changeable, they can never be regarded as "ultimate" constraints on behavior (such as those defined by N). Accordingly, all institutional manifestations — whether they be relative prices, government policies, regulatory restrictions or informal behavioral norms — must be explained as a rational consequence of individual attempts to achieve aggregate coordination, and can never be separated from the process by which such coordination is achieved. Robert Lucas (1981) puts this idea succinctly when he states:

... conventions and institutions do not simply come out of the blue, arbitrarily imposing themselves on individual agents. On the contrary, institutions and customs are designed precisely in order to aid in matching preferences and opportunities satisfactorily (p. 4).

The important methodological implication of this argument therefore may be summarized as follows:

**Proposition A:**

In any logically-complete individualistic GE explanatory structure, all elements of the set, S, must be endogenously-codetermined with the set, \( \chi^* \), as explained by "more fundamental" factors.

It is noted that the phrases "successfully closed", "behaviorally nonarbitrary", and "complete with respect to rationality" all may be taken as synonymous with the term "logically complete". Moreover, as will be shown below, if the "more fundamental" factors specified to explain \( \{S^*, \chi^*\} \) are just the set \( \{N, I\} \) — as in the Lucas quote — then the explanation is Walrasian in form. The use of any different set of explanatory components for this purpose must imply some form of Non-Walrasian theory.

Proposition A therefore removes any mystery as to why a purportedly more general assignment of explanatory power that involves exogenous institutional constraints can be seen to offer only a behaviorally-arbitrary explanation of \( \chi^* \). By individualistic standards, such an explanation cannot be more general than its \( \{N, I\} \) counterpart, since the S-constraints employed are only artificial; they are not truly exogenous and have no independent explanatory power. The explanation is behaviorally arbitrary simply because it fails to explain why S is the way it is; i.e., why agents would rationally choose to endorse the posted institutional constraints over alternative configurations.

1.3. The Definition and Explanation of Institutional Distortions. — In
any individualistic framework, the concept of “institutional distortion” has a precise definition: any element of \( S \) which is not compatible with (fails to adjust to consistency with) given \( N \) and \( I \). Denote any such element by \( \tilde{s} \); the collection of all such elements being denoted by \( \tilde{S} \). Any exogenous element of \( S \) must belong to \( S \), except by accident.

The reason why individualists invariably identify institutional distortions with the existence of aggregate welfare losses and the denial of Pareto-efficiency is basic. It follows from the recognition that a necessary GE condition for the achievement of both a (relative) welfare maximum and Pareto-efficiency is that the number of (severity of) independent constraints on individual interaction be at a (relative) minimum \(^4\). A Walrasian structure which, given \( I \), posits only \( N \)-constraints as independently binding, defines this benchmark minimum, since it explains all other constraints (i.e., such as those in \( S \)) as derivable from \( N \) (and \( I \)). (Note that this framework implies absolute constraint minimization only in the special and trivial case where the set, \( N \), is empty). It follows, therefore, that in any case in which an \( I \)-constraint is also regarded as binding, the set of independent constraints must be larger than the benchmark minimum set defined by \( N \) alone, further restricting the opportunities for gainful individual interaction. Accordingly, aggregate welfare must be lower than its benchmark maximum and the necessary condition for Pareto-efficiency is sacrificed. The “efficiency loss” sustained emerges as the failure of (rational) individuals to adjust a potentially-changeable constraint to consistency with \( N \) and \( I \).

The above case must be clearly distinguished from one in which a more severe \( N \)-constraint is injected into the setting, supplanting pre-existing, less severe \( N \) constraints, and making the benchmark minimum set more severe. Here aggregate welfare is again lower but Pareto-efficiency still holds because individuals continue to maximize only with respect to constraints they cannot change. (Relative) welfare maximization is thus maintained \(^5\).

The above analysis is important not only in clarifying a fundamental distinction between \( N \)-constraints and \( I \)-constraints, but also because many economists of an individualist persuasion have objected to an institutional distortions perspective precisely on the basis of the welfare implications presented here: that such explanations do not square well with the classical postulates of efficiency, maximization, and rationality on the part of agents. Unfortunately, such a view, taken literally, leads to a tautological formulation of the individualist position. The only substantive reason why an individualist objects to these models is that they constitute logically-incomplete explanations: they fail to explain the determination of \( \tilde{s} \). If one could produce a theory of “rational” distortions (i.e., determine \( \tilde{s} \) elements of \( S \) as \( \tilde{s}^*, \) at \( \chi^* \)), there would be no reason in principle for the individualist to shy away from these models. To clarify the difficulty of such a venture and to provide an initial orientation to the logical structure of this problem, it is now useful to state and prove a further general proposition:

**Proposition B:**

If the set \( \{N, S, I\} \) describes an “exhaustive” set of exogenous variables by which to explain \( \chi^* (\chi) \), and the conditions of Proposition A hold, then there exists no logical possibility of a successful individualistic explanation of the existence of (rational) institutional distortions.

**Proof:**

Presume the (conjectured) existence of one \( \tilde{s} \), say \( \tilde{s}_0 \), \( \{N, I\} \) can offer an explanation of this institution in accord with Proposition A but only if it is construed as non-distortionary; \( \{N\} \) can offer only a non-individualistic explanation (since \( I \) is not exogenous) and \( \{I\} \) is just an economically-trivial version of \( \{N, I\} \). Thus, more explanatory power is needed. By assumption, this can only come from the set, \( S \). To explain \( \tilde{s}_0 \) (as \( \tilde{s}^* \)), it is thus required that at least one other element of \( S \) be assigned exogenous status. But this contradicts the conditions of Proposition A, since it implies that \( S \) can be only a subset of \( \tilde{S} \), which completes the proof.

The implication of Proposition B is therefore that a theory of rational institutional distortions can be successfully constructed if and only if a class of exogenous variables other than, and independent of \( \{N, S, I\} \) can be specified in the explanation of \( \chi^* (\chi) \). In principle, such a set must serve to augment the explanatory power of \( \{N, I\} \) while satisfying the condition that \( S = \tilde{S}^* \) at \( \chi^* \). To identify what this new explanatory component might be, and how it can be fitted into existing analysis, is the ultimate objective of this paper.

1.4. Institutionalist Explanations and Problems of Explanatory Equivocation. — Before proceeding to discuss the long-standing debates in theory and policy which have emerged from the above problem setting, it is useful to emphasize just how much of the above analysis rests on the enforcement of
individualistic postulates. Instead, suppose that one were to adopt the traditional “institutionalist” (or “holistic”) viewpoint which states that social institutions ultimately possess a “life of their own”, independent of individual will, and are not potentially open to gainful adjustment by any or all individuals. Presuming that this viewpoint might be one implicit rationale underlying a defense of descriptive realism (at least in “short-run” models), how would the analysis be affected?

While such a viewpoint does preclude the possibility that $\dot{S}$ may be determined by even more aggregative process, as far as the explanation of individual behavior and its coordination is concerned, $\dot{S}$ now becomes a fundamental constraint on individual action, on the same footing as $N$. Accordingly, the methodological distinction between $N$ and $\dot{S}$ is rendered insubstantial. Moreover, since $\dot{S}$ now constitutes an “acceptable” set of exogenous variables by which to explain $X'$, it can form part of an explanation which is logically complete, relative to institutionalist standards. On the other hand, the institutionalist conception of institutional distortions does not prove to be very meaningful here — since individuals have no potential to change $\dot{S}$. In turn, the pursuit of a theory of rational institutional distortions becomes irrelevant; this problem is only for institutionalists.

Now, if economists currently were to divide along the lines of the historical institutionalist/institutionalist schism, then it is apparent that the former probably would be satisfied with (complete) explanations defined only in terms of $(N, I)$, while the latter could be equally content with (complete) explanations using only $(N, \dot{S})$. What makes the need for the intermedation proposed in this paper so pressing is the equivocation between these two sets of explanatory standards which is necessarily present in all non-market clearing, policy and applied microeconomic formulations that treat elements of both $I$ and $\dot{S}$ as exogenous variables in a GE context. Here the institutionalist can only question why $\dot{S}$ is given exogenous treatment, while the institutionalist can only wonder why $I$ is assigned autonomous status.

To illustrate how such explanatory equivocation makes it difficult to come to grips with distortionary phenomena, or put forth any coherent policy viewpoint, one need look no further than “obvious” phenomena: for example, a monopolist protected by a barrier to entry. Traditional textbook expositions of this behavior typically start right from the presumption that the barrier which allows the monopolist to reap excess profits is an $\dot{S}$. It therefore follows axiomatically that a monopolist’s price and output decisions are distortionary and that welfare can be increased by removing the monopoly. Unfortunately, by institutionalist standards, this model is incomplete. In lieu of any theory concerning why the $I$ could be rationally sustained — as distortionary — it is difficult not to speculate that the welfare losses from monopolies may be more a consequence of the incompleteness of the model than the world. Completing this model therefore only can mean showing that the barrier actually embodies an $N$, not an $\dot{S}$-constraint (natural monopoly), or showing that in fact it is an optimal social contrivance, compatible with $N$ and $I$. In either case, the completed version of the model (as expected) reinstates the potential for Pareto-efficiency, the implication being that we need not gain from removing monopolies after all. Suppose now that institutionalist standards were placed into this discussion so that $\dot{S}$ was regarded as a “non-changeable” constraint (like $N$). Then, we would reach yet another set of conclusions; namely, that both the initial $\dot{S}$-case of contrived monopoly and the case of natural monopoly must be regarded as observationally-equivalent and distortion-free (since individuals can change neither $N$ nor $\dot{S}$), and that both emerge from a framework which is logically complete by institutionalist explanatory standards. Evidently, there are too many conclusions here.

1.5. The Tradition of Walrasian Theory and Policy. — It is now appropriate to define formally a Walrasian GE explanatory structure as one which exhibits the following (generic) causal form:

$$[N, I] \rightarrow [\dot{S}^*, Z^*],$$

where “$\rightarrow$” is to read “is just causally sufficient to explain”, and where $Z^*$ specifies $\chi^*$ for the Walrasian case.

It is evident that this type of structure can admit a variety of alternative mathematical specifications of the form of $N$ and $I$, as well as many different “sizes” of activity/commodity space (as defined by $Z^*$). So long as a GE-$Z^*$ “exists”, however, all these “models” must share the same underlying structural property: that all elements of $\dot{S}$ are endogenous and non-distortionary, and codetermined with the Pareto-efficient behavior, $Z^*$. Accordingly, the logical completeness requirement of Proposition A is satisfied. The (assumed) causal sufficiency of $(N, I)$ is what gives this framework all its important implications. It means, for example, that any attempt to give a hitherto-neglected variable explanatory power must inevitably fail: the new variable simply will be forced into (endogenous) codetermination with $[\dot{S}^*, Z^*]$. Suppose that such a variable was “individual expectations of $\dot{S}^*$” —
where any arbitrary individual's expectation is denoted by \( E(S^*) \) and its aggregate counterpart by \( E(S^*) \). Then the Walrasian causal form could be expressed in terms of an "optimal" expectations requirement:

\[
(N, I) \rightarrow \{ Z^*, S^*, E^*(S^*) \}
\]

The condition is actually more stringent however. The behavioral structure of any individualistic model entails that \( S^* \) and \( E^*(S^*) \) conform identically to one another, since the realization of \( S^* \) hinges on the condition that all agents mutually anticipate the \( S^* \); thus, bringing it about self-fulfillingly by their actions based on these expectations. Hence, the familiar (deterministic) "rational expectations" condition emerges:

\[
E^*(S^*) = S^*, \text{ for all } j.
\]

While this analysis is consistent with the parable that all agents form their expectations on the basis of a "true" model of the economy (based on some given specification of \( \{N, I\} \)), the only fundamental feature involved is the methodological necessity of determining (rational) expectations and (rational) institutions by the same given.

What does this GE structure imply about government policy? To answer this question, it is useful to partition the set, \( S^* \), as follows: \( S^* = \{S_p, S_g\} \). \( S_p \) denotes the set of coordination/allocation "rules" under the jurisdiction of the private sector (e.g., market prices) while \( S_g \) denotes a corresponding set of government-determined policy/planning "rules". It then follows directly that, at \( Z^* \), the condition, \( S^* = \{S_p, S_g\} \), must obtain.

The above condition rules out the possibility of any active or beneficial policy to alter \( Z^* \). What is generated here is a theory of endogenous (passive) policy to alter \( Z^* \) circular. It is optimal and can vary if and only if \( N \) and/or \( I \) change. Furthermore, since it is implied that \( E^*(S^*_g) = S^*_g \) at \( Z^* \), all policy is "anticipated". These are the properties which underlie both classical "invariance" theorems regarding real variables and a host of policy ineffectiveness propositions (see Sargent and Wallace, 1976). Policy can never affect \( Z^* \) simply because \( S^*_g \) is codetermined with \( S^* \). Conversely, any model which claims to produce positive policy multipliers in a setting where \( N \) and \( I \) are invariant must be an incomplete and misspecified model; it must be buying the role for active policy only through the imposition of artificial \( S_p \) - constraints or non-rational expectations (cf. Lucas, 1976).

There is a desirably self-contained quality to the above arguments. It is therefore ironic that while this framework puts forth a complete explanation of \( Z^* \), other issues of "incompleteness" easily can arise with respect to the explicit determination of \( S^* \) (and did arise in traditional formulations of this position). A classic instance is the recognition that the achievement of \( Z^* \)-invariance also means that the structure "solves" in real terms. As a consequence, it cannot explain all nominal variables/absolute prices without further exogenous restrictions. Thus, the \( S_g^* \) determined in this framework indeed is consistent with an optimal quantity of real balances (in the tradition of Friedman 1969), but need not reveal anything whatsoever about the institution of nominal monetary policy or the existence of a unique/stable path of absolute prices which may ensue from it (see Black, 1974; Gray, 1984; McCallum, 1987). There are potentially ways around this underdetermination — by employing transactions/computational costs (see McCallum, 1986) — but the point remains: more explanatory power is typically required to determine \( S^* \) explicitly than to pin down \( Z^* \) alone.

Perhaps an even more basic issue of "completeness" in a policy context arises from the recognition that any claims about \( S^*_g \) already presume a successful Walrasian demonstration that: (a) \( S^*_g \) is non-empty; and (b) all elements of \( S^* \) are allocated optimally between \( S_p^* \) and \( S_g^* \). Traditional theorizing in this area typically ignored the "optimal mix" question and thus left (a) indeterminate. As Coase (1937) and later Williamson (1975) pointed out, the appropriate additional explanatory power is a transactions technology (defined in \( N \)) which can generate a transactions cost differential between \( S_p^* \) rules and \( S_g^* \) rules that varies with the level of rule performance (output) and which makes it inefficient to use other than both types of rules. This result is clearly crucial to any formulation of a Walrasian theory of policy since an explanatory failure at this level must undermine macroeconomic invariance results entirely.

---

1 This condition can be abbreviated as \( E(S^*) = S^* \) in an aggregate context. The stochastic condition is \( E'(S^*) = S^* \), where \( E' \) denotes "average" expectations and \( E(S^*) \neq S^* \) in general. Note that each individual's action is explained by \( \{N, I, S^* \} \rightarrow Z^* \) consistent with \( Z^* \).

2 The partition of \( Z \) into public and private production is kept implicit here. For an explicit model, see Newman and Sopher (1979).

3 These properties also underlie a variety of other Walrasian theorems; e.g., those of Coase and Ricardo equivalence (see Ryus, 1979).

4 A good summary of this viewpoint is again Ryus (1979). Note that the popular idea that the Walrasian position implies "no role for government" was soundly criticized by Lange and
The final, and possibly most fundamental, incompleteness which requires mention follows from the arguments of Arrow (1959): that in any economy where “competitiveness” holds (i.e., all agents are assumed to treat all elements of the parameter set), there can be by definition no complete individualistic explanation of the dynamic process by which \( S^* \) is set.

All of the above examples have a common characteristic; they require the addition of explanatory power to remove the posited “incompleteness”, which in turn requires the injection of dynamic structure and further behavioral constraints into the pre-existing setting. This invites a stronger concept of explanatory completeness than that specified in Proposition A:

**Proposition C:** (Supercompleteness)

In any supercomplete individualistic GE explanatory structure, the determination (and adjustment) of all elements of \( S \) must be codetermined with \( \chi^{*} \) and generated from an explicit dynamic structure which specifies additional constraints on (i) transacting; (ii) procuring information; and (iii) institutional change, and defines individual “attitudes” towards these constraints.

From a history of thought perspective, Proposition C is most important since virtually all traditional Walrasian formulations failed to satisfy it. The emphasis was on the construction of static models of “ideal” resource allocation, where all allocational frictions were absent and where institutional processes were largely trivialized. Denote this specification as the Arrow-Debreu (AD) specification \(^{11}\), where \( N \) and \( I \) are defined over a commodity space which involves only “ordinary” private goods, ignoring the ancillary activities (costs) of transacting, adjusting, and procuring information, i.e.,

\[
(N_{AD}, I_{AD}) \rightarrow (S_{AD}^{*}, Z_{AD}^{*})
\]

Structurally, the outstanding feature of AD was its “ideal” character; \( N_{AD} \) constituted the least constraining specification of \( N \)-constraints available in a Walrasian context, so a maximum maximizer for aggregate welfare could be defined (i.e., \( V_{AD}^{*} = V_{*}^{*} \)). Methodologically, the significance of AD lay in the fact that it typically was taken to be the only form of Walrasian theorizing; that the properties of AD were synonymous with the properties of Walrasian explanation as a whole. This was unfortunate: since it was largely agreed that the observed world was “non-ideal” (at least in the short-run), it was straightforward for critics to argue for the falsity of Walrasian predictions (\( x_{AD} \neq Z_{AD}^{*} \)). Furthermore, if “non-ideal” was equated with “distortionary”, a virtual invitation was handed to critics to seek out maligned social institutions and other constraints blocking full maximization, and to appeal them to AD, simply to achieve enough expository power to probe observed history. Accordingly, a trade-off between individualistic completeness and descriptive realism became inevitable. Friedman’s (1953) classic essay on methodology did little more than entrench this viewpoint since, rather than arguing for Proposition C explicitly (which would have preserved individualistic completeness in non-ideal/realistic settings), it resorted ad hoc to a defense of instrumentalism (see Boland, 1979) that rationalized how descriptively-unrealistic assumptions could still produce real world predictive success \(^{12}\).

The major contribution of recent Walrasian thinkers such as Lucas (1980, 1981) is to do what Friedman did not do: to aim exclusively for the fulfillment of Proposition C. This frees Walrasian explanation from AD confines and opens up the possibility that alternative Walrasian specifications can illuminate non-ideal behavior as well as their traditional competitors while still maintaining Proposition C (and A). The strategy is apparent: insofar as there are no restrictions on the possible expansions of the elements of the sets \( N, I, S, \) and \( \chi \), Walrasian causal form, \( (N, I) \rightarrow (S^{*}, Z^{*}) \), can hold for an indefinite number of larger and more explanatory GE problems than AD, so long as any expansion of \( S^{*}, Z^{*} \) is just met with the “appropriate” expansion of \( (N, I) \).

Denote any such specification as being of Extended Walrasian (EW) form:

\[
(N_{EW}, I_{EW}) \rightarrow (S_{EW}^{*}, Z_{EW}^{*})
\]

where EW-equilibria are defined in a higher commodity (activity) space than that of AD. Since it follows that \( Z_{EW}^{*} \neq Z_{AD}^{*} \), the former is compared only over AD-activities, i.e., \( Z_{EW/AD}^{*} \) and \( S_{EW/AD}^{*} \neq S_{AD}^{*} \), except accidentally, non-ideal behavior is illuminated: There are no institutional distortions, so Pareto-efficiency exists. But \( N_{EW} \) contains \( N_{AD} \); hence, aggregate welfare, \( V_{EW/AD}^{*} \) is lower than \( V_{AD}^{*} \).\(^ {13}\) The “rational expectations” condition is still maintained; hence: \( E^{*}(S_{EW}^{*}) = S_{EW}^{*} \).

---

\(^{11}\) Taylor’s (1938) demonstration that the same GE problem could be solved for \( \{S_{1}\} = \{2\} \). This raised the “indeterminacy” problem. A classic, but neglected, statement of the “efficiency” of government intervention is Samuelson (1958, p. 491).

\(^{12}\) The terminology follows Fischer (1977b). A better term might be the “Hicks-Allen-Samuelson” specification.

\(^{13}\) Since instrumentalism has no intrinsic relation to the methods of economic theory while logical completeness is central, it is ironic that economists have paid much more attention to the former.

\(^{14}\) Note that \( V_{EW/AD}^{*} = V_{EW}^{*} > V_{AD}^{*} \) if the “new” activities yield no utility.
1.6. The "Institutional Distortions" (ID) Structure. — It is taken as axiomatic that traditional theorists who defended ID models — whether they were of a Keynesian or policy persuasion, or interested in analyzing non-competitive market forms — did so for strong reasons. If they accepted the ideas that AD was the only Walrasian specification (and explained only "ideal" circumstances), that the AD commodity space specification was appropriate and that seeking exogeneity beyond \( \{ N, S, I \} \) was impossible, then the use of ID-models to explain non-ideal settings became an inevitability. The generic form of this type of explanation is:

\[
\{ \hat{S}, N, I \} \rightarrow \{ \hat{S}^e, \hat{Z}^e \},
\]

where \( \hat{S} \) denotes the set of (exogenous) institutional constraints, \( \hat{S}^e \) is the set of explained institutional distortions, and \( \hat{Z}^e \) is the non-ideal and distortionary activity level. It is a matter of definition that \( \hat{S} \) be non-empty while it is characteristic of these formulations to assume that \( \hat{S}^e \) is non-empty. Noting \( \tilde{S} = \{ \hat{S}_p, \hat{S}_g \} \), and keeping implicit its subscript below, the traditionally-popular ID specification then becomes:

\[
(\tilde{S}_p, N_{AD}^e, I_{AD}^e) \rightarrow (\tilde{S}_p^e, \hat{S}_g^e, \hat{Z}_{AD}^e).
\]

Here \( \tilde{S}_p \) may be defined in terms of a fixed price or wage contract/structure, a "market failure", a barrier to entry, or more general types of institutional failure which induce free riding, incentive incompatibility, or moral hazard.  

The overriding generic properties of this Non-Walrasian structure are:

(a) it determines a "second best" equilibrium in which Pareto efficiency is denied and aggregate welfare is not maximized, relative to \( N \) and \( I \) (i.e. \( V_{AD} < V_{AD}^{*} \)), and (b) it violates the completeness (rationality) conditions of Propositions A and C from the outset. The latter point makes it clear that more recent attempts to force rational expectations into this framework via the general conditions:

\[
E^*(\hat{S}^e) = \hat{S}^e \quad \text{and} \quad E^*(\hat{S}_g^e) = \hat{S}_g^e,
\]

constitute only spurious attempts at model closure (e.g., Fischer, 1977a; Taylor, 1980). If individualistic standards are maintained, and a theory of rational distortions is not specified, the framework cannot explain why agents would sustain \( \hat{S} \) in the first place, let alone why they would ever treat \( \hat{S} \) parametrically in forming their expectations. Thus, this structure is incomplete with respect to rationality even with (assumed) rational expectations.

Given the pervasiveness of "active" policy prescriptions in the Keynesian/Pigovian tradition, it is easy to ignore the inherent "passivist" imprecision of this framework. Thus, in the GE described above, only the endogenous policy, \( S_g^e \), is implied. This is codetermined with \( Z^e \) and is "optimal" in the sense that any other policy configuration, given \( S_g^e \), must lower aggregate welfare. The historical arguments for active welfare-improving policy therefore can rest only on the additional presumption that some elements of \( S_g \) are exogenous, and can influence elements of \( \{ S_p, N_{AD}^e, I_{AD}^e \} \). This in turn permits \( Z^e \), to change. No matter whether a full employment \( Z^e \) or some other objective is targeted, or whether \( S_g^e \) is supposed to achieve this: (i) by directly influencing (offsetting) \( S_p \), or (ii) through a tax/transfer/spending structure which changes \( N_{AD}^e \) by redistributing existing resources towards government and away from the private sector, the methodological picture is unaltered. The causal sufficiency of this explanation requires that the original explanatory incompleteness associated with \( S_g^e \) simply be traded for a like one with respect to \( S_p^e \); the idea of welfare-improving government policy is therefore just a methodological contrivance when \( S_g^e \) is left unexplained. 

The above verdict of course does not bear on the intrinsic possibility of beneficial government policies; it simply points up the traditional failure to explain their emergence from individualistic foundations. And even Walrasian theories of endogenous policy must have problems with these foundations in an AD world where transactions/adjustment/information costs are uniformly set equal to zero. The ID-version of this structure just makes these problems more glaring; it cannot even provide a rationale for why \( S_g^e \) offers more flexible (and faster) adjustment instruments than \( S_p^e \) (Even if an ad hoc relative costs rationale was specified, the situation would still be difficult, since if \( S_p^e \) was assigned a uniform comparative advantage over \( S_g^e \), it would just beg the question of why the sluggish institutions comprising \( S_p^e \) were ever chosen to serve in the institutional "mix" in the first place).  

\[ 13 \quad \text{This might be regarded as "active" policy in the insubstantial sense that, starting from a Walrasian } Z^e, S_g^e \text{ must be adjusted to } Z_g^e. \text{ However, such a view, and the view that aggregate welfare improves as this adjustment takes place (i.e., en route to } Z^e) \text{ are second-best GE propositions; they are "third-best" arguments.} \]

\[ 14 \quad \text{This is the same as the argument which proves Proposition B. This point is independent of the additional idea that governments need have no incentive to engage in action which maximizes aggregate welfare (Nieselen, 1971; McKenzie and Macaulay, 1980).} \]
However, it is the coexistence of both exogenous and endogenous elements of \( S_e \) and \( S_e \) in ID models which opens up the more basic question: namely: Are individualistic, or institutionalistic, postulates really being used in these policy problems (see Brunner and Meckling, 1977)? If the individualist reacts to the asymmetrical treatment of \( S \) by wondering why the \( S_e \) can adjust so well while the \( S_e \) are so sluggish, then the institutionalist also questions why only some elements of \( S \) are granted a "life of their own". Moreover, just as the individualist will wonder why, if social constraints are fundamentally "changeable" by individuals, the initial \( S_e \) could ever be rationally sustained and generate a welfare loss relative to the AD "ideal", so the institutionalist will wonder why this welfare ideal is even relevant if \( S_e \) is fundamentally "non-changeable". The ambivalence between these two explanatory standards must make for an unhappy compromise: ID structure can be regarded as "arbitrary" by both camps.

1.7. The "Observational Equivalence" of Extended Walrasian (EW) and Institutional Distortions (ID) Approaches. – The explicit objective of EW explanation is to provide a first-best equivalent to any ID account of observed departures from AD outcomes. This approach, therefore, aims at satisfying Proposition C (and A) and constitutes one possible attempt to provide a theory of "rational" distortions. Alleged departures from Pareto-efficiency, such as wage-price rigidity, market failure, and the like, are no longer explained as institution failures (as in the ID approach) but as rational, efficient GE consequences of an extended set of Walrasian constraints. This section synthesizes earlier arguments by examining these "equivalence" and its implications.

Denote any observed "non-ideal" state of the economy by \( \hat{x} \), where \( \hat{x} \neq Z_{AD} \) and where, by definition, \( V^x < V^*_{AD} = V^{max} \). Since \( \hat{x} \) cannot be explained by \( \{N_{AD}, I_{AD}\} \), more explanatory power is needed. The ID approach therefore appeals \( \hat{x} \) to these given while keeping the AD commodity space intact. Thus,

\[
\{\hat{x}, N_{AD}, I_{AD}\} \rightarrow \{\hat{x}, \bar{x}, Z_{AD}\}
\]

such that \( \hat{x}, \bar{x} \in Z_{AD} \), \( \bar{x} < V^*_{AD} \), and where the existence of institutional distortions are the exclusive reason for why \( Z_{AD} \) is not achieved.

The key insight of the EW approach is that the only reason why the ID approach generates second-best outcomes is because it adds an extra (and artificial) set of constraints without extending the commodity space. Thus, it succeeds in removing the initial situation of explanatory underdetermination only by replacing it with one of overdetermination. Accordingly the initial EW strategy is to expand the commodity space by adding a set of (costly) "ancillary" activities \( (Z_e) \) to the problem, typically those connected with (i) transacting; (ii) procuring information and (iii) institutional adjustment. This then allows \( \{N_{AD}, I_{AD}\} \) to be expanded (without overdetermination) by appropriate technology, \( N_e \) and preference, \( L_e \) elements, corresponding to the new activities. The Walrasian form implied is:

\[
\{N_{EW}, I_{EW}\} \rightarrow \{\hat{x}_{EW}, Z_{EW}\}
\]

where \( N_{EW} = \{N_{AD}, N_e\}, I_{EW} = \{I_{AD}, I_e\}, \hat{x}_{EW} = \{\hat{x}_{AD}, \bar{x}\} \) and \( Z_{EW} = \{Z_{AD}, Z_e\} \), and where \( Z_{EW/AD} = \hat{x} \).

This EW explanation accounts for the same \( \hat{x} \) as the ID approach but the GE is non-distortionary and satisfies individualistic postulates. \( S \) is rendered endogenous (as part of \( S_{EW} \)) because it is explained as a "rational" consequence of an augmented environment which includes \( N_e \) and \( I_e \); the \( \hat{x} \) are now transformed into "optimal" rigidities. The simplest case in which aggregate welfare equivalence also obtains is where the ancillary activities yield no utility. Then, \( \{I_{EW}\} = \{I_{AD}\} \) and it follows that \( V^x_{EW} = V^*_{AD} \). The explanation for this result is that, at \( Z_{EW} \), the new costly activities, \( Z_e \), pull output away from the pre-existing AD activities by exactly the amount that the \( S \)-constraints create an "output loss" in \( Z_{AD} \) through inefficiency.

This equivalence argument is especially important, first, because it suggests that, given no restrictions on the expansions of the elements of, respectively, \( N_e \) and \( I_e \), and \( S_e \), both approaches may be explanatory for (infinitely) many \( \hat{x} \) instances. Second, it suggests the general proposition that any second-best argument for active policy intervention lies a corresponding first-best argument with only passive policy implications; that there exists a first-best analogue model underlying any second-best policy formulation. The other side of the coin is that, if some specification of both...

---

17 A typical example of this asymmetry is an IS/LM structure which assumes fixed nominal wages and flexible output prices. The structure generated by assigning all elements of \( S \)'s life of their own is \( \{N, S\} \rightarrow \{\hat{x}, \bar{x}\} \), which denies individualistic postulates outright.

---

18 The following section will provide a more explicit representation of the informational constraints in this structure; the last section deals with adjustment costs. In the latter, a dynamic EW-path of institutional change is implied and the additional condition, \( E^x(ds/dt) = (ds/dt) \), is implied.

19 For the given AD commodity space, there are definite limits on the "size" of the set \( S_e \), but not on the "values" of its elements, or on the possibility of transferring elements of \( \hat{x} \) back to \( \hat{x} \).
ID and EW potentially can "fit the facts", then the relative merits of the two approaches cannot be decided by empirical evidence. The choice between the two explanatory forms is methodological: it rests only on the issue of whether or not one wants individually-complete explanations.

Evidently, if the theorist wishes to satisfy Proposition C (A), but also feels he should be able to illustrate the existence of "true" behavioral distortions, then the fundamental dilemma indicated by Proposition B comes into play. Since EW provides a theory of "rational", yet only "apparent" distortions, any insight into "true" distortions requires an explanatory components beyond \( \{ N_{EW}, I_{EW} \} \) be specified. As elements of \( S \) cannot acceptably be used, the assumption that \( \{ N, S, I \} \) is an exhaustive set of (potentially) exogenous variables must be abandoned; new fundamentals must be found. In the following section, the candidate of "limited information" is considered. It is shown that this exogenous condition is necessary, but not sufficient, for the task at hand. Such a conclusion sets the stage for an individualistic analysis of institutional reform processes in the final section.

II. Limited Information and Institutional Distortions

Since the theme of limited/imperfect information may be seen to underlie any "sophisticated" defense of ID models (e.g., Keynes, 1936, 1957), it is now useful to consider the possibility that this type of constraint is "fundamentally" exogenous, may underwrite the sustainability of institutional distortions, and, along with \( N \) and \( I \), may provide the appropriate basis for a theory of rational distortions. The analysis will begin from the EW treatment of the topic, since it provides the benchmark case in which the above possibilities are denied.

2.1. The EW-Approach to Imperfect and Costly Information. -- The fundamental principle underlying this approach is that informational constraints cannot constitute an independent class of exogenous variables. Uncertainty, randomness, and a technology governing information acquisition and cost all can be regarded as elements of \( N_{EW} \), with a possible (Bayesian) option of placing some of this explanatory power in \( I_{EW} \) (see Poirier, 1988). Standard

EW form is therefore maintained, where rationality is complete with respect to all activities, including those connected with information acquisition. Informational constraints as such cannot be distortoraty, the level of informational activity is endogenous, and, since agents always maximize gains from trade relative to \( \{ N_{EW}, I_{EW} \} \), and \( S_{EW} \) obtains, no institutional distortions are possible either. (Once again, note that \( V_{EW/AD} \) is less than the \( V_{AD} \) determined with no informational constraints binding at all).

The essential feature of this structure is that it allows successful maximization processes to co-exist with informational limitations. Accordingly, it must endorse the postulate that the posited informational impediments can never be sufficiently constraining to block the determination of \( Z_{EW} \) (or render it accidental). In a stochastic context, this means that all agents in some way must achieve that state of knowledge which rules out all systematic errors in the maximization problem defined by \( N_{EW}, I_{EW} \), even though one of the parameters of this problem is an informational constraint. Denote such a state of knowledge as Walrasian Sufficient Knowledge (WSK). WSK therefore requires some assumption to the effect that agents can know "true" probability distributions and that they can distinguish systematic from random error (without systematic error). WSK is also the condition under which the rational expectations property, \( E^*(S_{EW}) = S_{EW} \), can be guaranteed to hold, where \( E^*(S_{EW}) \) denotes an "average" expectation. The "perfect" knowledge (foresight) assumptions of AD structure clearly satisfy, and are stronger than, the WSK condition.

The methodological thrust of this argument is therefore as follows:

**Proposition D:**

Under WSK, informational constraints are Pareto-irrelevant and cannot be used to explain any \( \{ S^*, Z^* \} \) outcome.

2.2. Non-Walrasian Approaches to Informational Constraints. -- The overriding implication of Proposition D is that any theory of "true" distortive phenomena must be based on the denial of WSK. The denial of WSK implies the contrary (except by accident) both \( E^*(S^*_{EW}) \neq S^*_{EW} \) and \( E^*(S^*_{AD}) \neq S^*_{AD} \), so that only a distortionary GE potentially can be explained. Here agents fail to maximize gains from trade and systematic expectation errors can be sustained. Denoting any state of knowledge which is less "informative" than WSK as Walrasian Insufficient Knowledge (WIK), a "sophisticated" version of an AD model then emerges. It attempts to explain both \( S \) (\( Z_{AD} \)) and \( S_0 \) (\( Z_{AD} \)) by the conjunction of \( N_{AD} \) and \( I_{AD} \) with WIK, i.e.,

\[ \{ N_{AD}, I_{AD}, \text{WIK} \} \rightarrow \{ S_0, Z_{AD} \} \]
In this structure, all elements of $S$ are endogenous and distortionary, which satisfies individualistic postulates, while WIK is assigned the status of a fundamental exogenous constraint (i.e., is Pareto-relevant). The explanation for why agents might sustain $\tilde{S}$ ($\tilde{S}^*$) is simply that they do not know any better.

This type of Non-Walrasian structure has had a variety of expressions (see Keynes, 1937; Hahn, 1978; Drazen, 1980; Fisher, 1981). If its (AD-based) form above could be shown to be observationally-equivalent to EW explanation, then the methodological superiority of the latter could be directly challenged and the required theory of “rational” distortions could come forth. Moreover, a distortionary version of EW could also be offered, i.e.,

$$\{N_{EW}, I_{EW}, WIK\} \rightarrow \{\tilde{s}_{EW}, \tilde{z}_{EW}\}$$

Why this framework has never really succeeded in its aim, however, is apparent: it does not specify any substantive concept of Non-Walrasian rationality (under WIK) which would illuminate why any particular $\{\tilde{S}^*, \tilde{Z}^*\}$ combination might be picked out of the many (infinite) available. Such a determination cannot be completed by reference to a given state of WIK alone, since each such state can be consistent with a whole spectrum of expected and realized $\{\tilde{S}^*, \tilde{Z}^*\}$ combinations. There are no “rational” criteria specified here to allow agents to pick between alternative equilibria. Moreover, this problem will not go away even if one makes recourse to a stronger specification of WIK: that the existing state of knowledge is (Non-Walrasian) sufficient to ensure:

$$\tilde{E}_n(\tilde{S}^*) = \tilde{S}^*$$

i.e., on average, all agents realize their anticipated institutional distortions. Suppose that agents modeled their expectations in ID form. There is still no process specified which explains how agents ever rationally arrived at a common model of $\{\tilde{S}, N, I\}$ by which to predict any particular $\tilde{S}^*$. There is therefore no explanation for why the aggregate of individuals would produce one mean expectation rather than another (see Frydman and Phelps, 1983).

The upshot of this argument is that the arbitrariness of expectations in this framework logically can be removed only by regarding some elements of $S$ as realized exogenously (in which case individualistic postulates are denied) or by assuming the additional knowledge required to attain WSK (in which case agents would rationally choose only non-distortionary institutions) $^2$. Without these problematic assumptions, this Non-Walrasian structure cannot be successfully closed. It cannot produce a theory of “rational” distortions.

The above problems may appear to be grave, but at least one basic proposition is consolidated:

Proposition E:

In any successful theory of rational distortions, the condition of WIK must be assumed to hold.

This states that endogenous knowledge must be a “fundamental” determinant of $\{\tilde{S}^*, \tilde{Z}^*\}$, along with $N$ and $I$, even if this constraint by itself is not sufficient to complete the explanation.

As for the structural deficiencies of the models considered above, they simply signal an unproductive approach to the problem. What these Non-Walrasian formulations miss is a basic appreciation of the individualistic idea that $S$ is changeable. Thus, even if one could guarantee that individual expectations of distortionary institutions are self-fulfilling, the overriding characteristic of the setting examined is that the institutional environment is sub-optimal and open to gainful institutional reform. Accordingly, the above framework at its best could never guarantee the sustenance of any particular anticipated $\tilde{S}$ ($\tilde{S}^*$) unless it removed the potentiality of rational reform altogether by way of non-individualistic agents. As will be shown explicitly in the next section, it is the specification of agents' reactions to the "constraints" on potentially-gainful reform, given WIK, which is the key to both formulating a concept of Non-Walrasian rationality and illuminating why agents may sustain distortionary institutions in a coordinated setting, even though they have the option of changing them. Here the realization of expectations holds only secondary interest as a closure device.

III. "Rational" Distortions as a Consequence of "Rational" Institutional Reform Under Walrasian-Insufficient Knowledge

This section proposes a structure for explaining the sustainability of institutional distortions which satisfies both the conditions of Proposition C (Supercompleteness) and Proposition E (existence of WIK). As in the classic contributions of Samuelson (1947) and Arrow (1959), a principal con-

$^2$ To assure the exogeneity of (long-term) expectations was Keynes' (1936) alternative.
cern is the possibility of determining an explicit Non-Walrasian model of dynamic institutional adjustment (rigidity) which does not violate individualistic tenets. As expected, the methodological situation here is not promising: one typically must choose between (incomplete) models of distortional institutional adjustment that employ fictitious "outside institutions" and other ad hoc mechanisms, and (complete) EW equilibrium dynamic structures that forfeit distortional character (see Gordon and Hynes, 1970; Rothschild, 1973; Fisher, 1976). For the individualist, however, success in this area is mandatory: a methodological view that rests on a fundamental distinction between $N$ and $S$ must provide a complete explanation of the process by which $S$ can, and is, changed by individuals. If no such theory can come forth (i.e., the adjustment process is merely "presumed"), then the sets, $N$ and $S$, cannot be distinguished substantively. The individualist can have little response to the ID-theorist who claims that an $S$ really is an element of $N$ (or vice versa). If these arguments were permitted, many years of doctrinal debate would be trivialized.

This paper invests the required additional explanatory power in a "reform technology", artfully specified so as to capture both the potential adjustability and rigidity of distortional institutions within one rational choice framework. The basic strategy is to view the prevailing distortional institutions of the WIK environment as both the objects of reform and the means to such reform. Thus, the process of institutional change becomes, in a sense, self-constraining. Under the concept of "Non-Walrasian" rationality implied, it now becomes rational for agents to attempt to "stabilize" (impede the change in) some elements of $S$ in order to reap beneficial changes in other elements of $S$, and quite irrational for them to try to adjust all distortional institutions as fast as possible towards a conjectured $S^*$. The "failure" of some distortional institutions to adjust then becomes the key to explaining any successful institutional reforms by individuals. In what follows, we highlight the "possibilities" of a framework built on this concept of rationality.

3.1. A useful reference for the proposed analysis is an EW structure which posits an explicit institutional adjustment constraint. Here positive adjustment costs underwrite non-infinite institutional adjustment velocities which, under WSK, define an optimal institutional reform path where $E^*$ $(dS/dt)^* = (dS/dt)^*$ obtains. A severely-constraining adjustment technology easily explains "sluggish" institutional reform as a Walrasian-rational, non-distortionary, response to the high costs of adjustment.

The setting to be examined bears a generic resemblance to the above, except when WIK is posited. Thus, agents are still in the process of "learning" WSK and (except by accident) are not on the EW-optimal path. There are now two types of institutional reforms to consider: (i) potential movements to the EW-path; and (ii) equilibrium movements along this path. Now, even if adjustment costs were zero (which would remove the dynamics in (ii)), reforms of type (i) potentially would still transpire. The following analysis attempts to account for rational, yet sluggish, adjustment in such a context.

3.2. Consider an economy with a large number of agents operating under WIK at a defined point in time. Given the existence of relevant ancillary technologies (but excluding those of adjustment), and given the values for $N$ and $I$, a stationary optimal set of institutions, $S^*$, is defined.

Each agent holds a conjecture concerning $S^*$, $E(S^*)$, which, given WIK, is not realized at the given date. Rather, $S = \{S_0, S_1, \ldots\}$ prevails, which implies the possibility of gainful institutional reform for all. While some agents may believe that $S = S^*$, presumably most agents will deem at least some elements of $S$ to be sub-optimal and will attempt to employ some fraction of their own total resources ($N_j$) to changing these entities. The set of institutions which any individual targets for reform (which may consist of none, some, or all elements of $S$) is denoted by the set, $S^*_j$, the time-rates of change of its elements being denoted by $S^*_j$.

Immediately, an impediment to pushing the explanation further emerges (Arrow, 1959):

**Proposition F: (Competitiveness)**

In any "large" economy, $N_j$ does not constitute sufficient resources for any one individual to successfully change any element of $S$.

Traditional proposals to circumvent this problem (in the context of price adjustment) have not been very interesting (see Arrow, 1959; Gordon and Hynes, 1970; Fisher, 1976; Drazen, 1980). To assume "monopoly power" is a submission outright to the ID-tradition and compromises individualistic postulates. To presume the existence of group (or lobby) coordination does not sit well with the knowledge limitations posited. The strategy employed here is to maintain Proposition F but to add to it a basic theoretical idea found in the theory of regulation (Stigler, 1971): that the prevailing
institutions of this environment – even if distortory – provide additional resources to any individual for the purposes of reform. Thus:

**Proposition G: (Power to Reform)**

In any “large” economy at any point in time, an individual can change elements of $S$ if and only if, given $N_j$, he uses at least one $j$ as a “means” to “input” into such reform.

Let $S_j^f$ denote the implied set of “input” institutions, with associated time rates of change $S_j^f$.

Proposition $G$ establishes a basic constraint on any agent’s reform plan: his potential reform of some institutions hinges on his decision not to reform other institutions. In general, given differences in $E_j(S^*)$, different reform plans will be pursued by different individuals. For example, one individual may attempt to use the existing institutions of government and government policy to effect a change in relative prices, another may try to use the existing relative price structure to procure a tariff, while a third may use prevailing levels of tariffs and prices in order to change current government policy. In all cases, the individuals in question do not have the power by themselves to change $S_j^f$ of the above institutions at the same point in time.

A basic implication of this type of structure is the following:

**Possibility I:**

It is logically possible that an agent rationally can endorse (not change) some $j$ even though he has the power to change it, and conjectures that he could gain from the change.

**Explanation:** Consider an individual who holds that, given $E_j(S^*)$, all elements of $S$ are distortory. By Proposition F, if he tried to change all such elements, he would relinquish his power to effect any gainful changes at all. Accordingly, he rationally must refrain from changing at least one $j$ so as to achieve the power to effect reforms in all his other targets.

Possibility I provides an important insight into how individual endorsement of (distortory) fixed prices/wages, monopolies, and even government policies can be explained as completely rational if these institutions can be seen to serve as inputs into, rather than outputs of, a reform plan. The static “welfare loss” typically attributed to each of these items is spurious in this context; it ignores the potential welfare improvement in the (reformed) output institutions which is made possible by stabilizing the input in question.

3.3. For any individual, the general relationship between alternative combinations of distortory reform “inputs” – own resources plus $S_j^f$ (changing at rate $S_j^f$), and intended reform outputs $S_j^r$ (changing at rate $S_j^r$), defines his “reform technology”. Each $(S_j^r, S_j^f)$ in turn may be associated with a unique expected mean return, $R_j$, and risk (variance), $\delta_j$. (Note that the return to changing any (allegedly) optimal institution is zero). Furthermore, each individual is presumed to possess well-defined preferences over $R_j$ and $\delta_j$.

Define the “scope” of an individual’s reform plan as the proportion of (total) $S$ which is allocated to $S_j^f$. For each individual, it is assumed that, as “scope” rises, the marginal gains from expanded institutional reform opportunities (at lower speeds of adjustment) fall, and are eventually outweighed by the rising marginal costs of lost inputs (i.e., reduced power). However, there is another relationship to consider: that the lower (higher) are the rates of change defined by $S_j^f$, the higher (lower) is $R_j$, and the lower (higher) is $\delta_j$. The rationale is that increasingly “stable” institutional inputs are more effective reform resources. A maximizing individual therefore will pick only relatively stable institutions to serve as inputs in his plan; he would never consider inputs that changed as quickly as his targeted reform outputs.

It is compatible with WIK to assume that each individual knows with certainty his own reform technology, resources, and preferences but has severely-limited information both on the reform parameters facing others and on the $S_j^f$ that he actually faces. The last item is especially important: given $j$’s personal parameters plus complete knowledge of (control over) the rates of change of his institutional inputs, $j$ could choose an ex ante optimal plan. But, in lieu of such knowledge (or control), he has no option but to determine his input choices, and their corresponding speeds of adjustment, on the basis of the conjecture, $E_j(S^*)$. Only if these expectations are realized does the intended reform output transpire. The above analysis immediately generates the following possibility:

**Possibility II: (Unanimity)**

If all agents have identical personal parameters [and $E_j(S^*)$] and uniformly

---

81 This view is also central to the theories of “piecemeal social change” found in Popper (1945); see also Agassi (1960).

82 For an informational rationale for this argument – to the effect that volatile institutional rules reduce “public information” – see Newman (1976).
employ as inputs only those institutions which are thought to be optimal, then the reform plans of all agents can be realized even under WIK.

The result follows from the fact that all individuals use the same model of reform and implicitly agree not to change each other’s inputs; input expectations are realized self-fulfillingly at a zero rate of change.

Possibility II is important because it serves both as a benchmark for the existence of EW-style results in a WIK setting and as an illustration of what will generally not occur in this setting. Any heterogeneity in personal parameters (or even $E_i(S^*)$) renders this realized expectations possibility accidental, since individuals will choose different ex ante reform plans and mutually undermine each other’s use of institutional inputs. If agents are not myopic, then they will recognize this fact: each agent will make his allocation conditional on his expectation of the extent to which any chosen set of institutional inputs could be reform outputs in the plans of others. This is what $E_i(S^*)$ now reflects.

Especially in the case where all agents pick “high scale” reform plans, the game theoretic possibilities of such a setting are unlimited. On the other hand, when “low scale” plans are uniformly used, a different dimension of this framework emerges: that many individuals will endorse, and thus stabilize, the same distortionary inputs even though their reform objectives are quite different. The “sharing” of stabilized inputs not only improves the probability of success of low scale reforms but it also supports a concept of (partial) institutional coordination which does not rely either on the perceived optimality of institutions or on the existence of a reform consensus. This view also may provide basic insight into why specific (“special interest”) reform plans can be more successful than general reform programmes, and why individuals typically prefer to use the stabilized, yet distortionary, institutional resources of government to their own ends rather than putting these institutions directly on the reform agenda.

3.4. It is useful to highlight two more important “possibilities” within this framework:

Possibility III:

In any WIK reform setting in which strategic interdependence exists, it can be completely rational for an agent to endorse a change in an institution which he thinks is currently “optimal”.

Explanation: Let $\bar{S}$ be such a (conjectured) optimal institution, given $E_i(S^*)$. By definition, $j$ anticipates zero return from changing $\bar{S}_j$; he would like to

employ it only as a reform input. However, $j$ conjectures that most other individuals would like to change $\bar{S}_j$ and could gain from the change. Then, it would be completely rational for $j$ to accept the anticipated change in $\bar{S}_j$, redefine his reform plan in the light of these expectations, and exploit the stability of the institutions which others are using to change $\bar{S}_j$, rather than attempting to employ his own resources and other institutional inputs to impede the change in $\bar{S}_j$ directly.

Possibility IV: (Aggregate Consistency Condition)

In any interdependent WIK reform setting, the failure of some group of individuals to effect institutional reforms implies ex post that other groups of individuals must have been (largely) successful in realizing their reform plans.

Explanation: Consider the simplest case of two possible types of individuals, forming groups A and B respectively, and only two institutions, $\bar{S}_1$ and $\bar{S}_2$. A and B have identical resources and preferences, but differ on $E_i(S^*)$: Group A thinks $\bar{S}_2$ is currently optimal and would like to use it to change $\bar{S}_1$. Group B has exactly the reverse situation. Now, suppose Group A failed to change $\bar{S}_1$. Then, this event can be explained only by A’s failure to stabilize its input, $\bar{S}_2$. But this implies that Group B must have achieved its targeted reform in $\bar{S}_2$, given A’s failure to adjust $\bar{S}_1$.

The significance of this latter possibility is that, even under the condition of WIK, the existence of (some) successful reform cannot be ruled out. Moreover, since there is no reason to believe that any one group of individuals would achieve success all of the time, it is implied that, on average, all groups of individuals will achieve their desired institutional adjustments some of the time. The important point about this idea is that it emerges from a context where there is no guarantee of realized reform expectations; successful institutional changes therefore may be largely unintended. Furthermore, while any given realized institutional change may be conjectured to be a welfare improvement by its protagonists, under WIK there is no guarantee of this outcome either. Hence, a successful institutional reform in principle can be welfare-decreasing. There is nothing particularly startling about this outcome; in fact, it is what drives the dynamics of this structure. Successful institutional change in one period does not stop any agent from playing “the reform game” in the next period with a new set of realized distortions, and a new reform plan. It is only the ultimate achievement of WSK which can shut off this process.

3.5. Following in the tradition of Kydland and Prescott (1977) and
Barro and Gordon (1983), it is unproblematic to extend this “game” to include government policy makers, since policy changes are no more than a reform in (a subset of) $S_a$. Analytically, the reform plans of government policy makers (given their reform technology) can be placed in a common strategic context with those of private sector agents; policy effects emerge as an endogenous consequence of this interaction. This framework invites a basic distinction. A “passive” policy change is one which emerges from a private sector reform of policy-related elements of $S_a$ given appropriately-stabilized private and non-policy related governmental inputs. An “active” policy change is initiated directly by the government, using some stabilized private sector inputs. (The speed and size of the policy change is a function of the extent of these inputs and their degree of volatility).

A standard “game” setting here is one where, under WIK, policy makers pursue an active policy strategy (as above) while private sector agents pursue private sector reforms using stabilized governmental inputs. Noting Possibility IV, an “impotence of active policy” result can emerge in any context where there is (widespread) private sector reform, since this must destabilize the inputs needed for government to effect substantive policy adjustment. The other half of this argument is that the private sector reforms have transpired precisely because the corresponding governmental inputs were so stable. Even though the assumed setting is so clearly distor- tionary, a traditional Friedman-Simons argument concerning the benefits of fixed or slow changing policy rules surfaces here, and goes hand-in-hand with claims about the impotence of active policy. The (Keynesian) corollary is of course that successful discretionary policy presumes the ineffectiveness of private sector reform. This is exactly the condition which motivates ID-models.

3.6. There are a wide variety of other interesting issues which might be explored within the above structure. This paper has aimed simply to demonstrate the “possibility” of producing explanations which are complete in individualistic terms and yet which still can admit (rational) institutional distortions. In doing so, it has criticized both the Walrasian viewpoint (which forfeits distortionary phenomena altogether) and the traditional “institutional distortions” viewpoint (which admits distortionary phenomena only on a non-individualistic basis) and indicated a pressing need to find methodological options which go beyond these two entrenched positions.

REFERENCES

Agassi J., “Methodological Individualism”, British Journal of Sociology, September 1960, 11, 244-70.
—, “Institutional Individualism”, British Journal of Sociology, June 1975, 26, 144-55.


VERSOSA TEORIA DELLE DISTORSIONI "RAZIONALI"

Questo articolo esamina il fondamento "individualistico" della teoria dell'equilibrio generale. Il problema più difficile in questa tradizione è quello di ricongiungere l'esistenza di fenomeni distorsivi con l'esistenza della razionalità (completa). Si mostra che questa ricongiunzione è logicamente possibile e si rompe così la dicotomia fra i modelli Walrasiani (che scartano i fenomeni distorsivi completamente) e i modelli con "distorsioni istituzionali" (che compromettono il postulato individualistico). In questo modo viene chiara un'ampia varietà di strutture della teoria e politica tradizionali e valutata in termini del concetto relativamente semplice di "completeness del modello".

---


---


---


---


---


---


---


STUDI IN MEMORIA DI TULLIO BAGIOTTI
ESSAYS IN MEMORY OF TULLIO BAGIOTTI

Comitato Promotore
Sponsoring Committee
F. Caffè
G. Demonti
G. De Nard
G. Ferrari
P. Silvio Labin

F. van Hayek
G. Myrdal
F. Perroux
P.A. Samuelson
R.S. Vaneey

Curatori
Editors
A. Agnati
D. Cantarelli
A. Montesano


1988. In 8°; pp. LX-1366 L. 120.000

EDIZIONI CEDAM - PADOVA