PART ONE

INFLATION AND STRUCTURE

CHAPTER 2

MACROECONOMIC PERSPECTIVES ON INFLATION AND UNEMPLOYMENT

Modern macroeconomic theories for inflation and unemployment have evolved in a dual love-hate relationship with the Phillips Curve. The notion that there exists a stable inverse relationship between inflation and the rate of unemployment -- dubbed as the Phillips Curve after the original work by Phillips (1958) -- was assimilated into macroeconomic models during the 1960s. The theoretical relationship was supported by observations stretching over close to a century, yet, as soon as macroeconomists put their new discovery into use, the Phillips Curve seemed to break down! During the late 1950s, when Phillips published his original article on the British experience, many macroeconomists in the United States were perplexed by the persistence of inflation in the midst of recession. Later, during the 1960s, inflation in most advanced capitalist economies accelerated with no apparent decline in the rate of unemployment. Finally, since the mid 1970s, after a dramatic rise in both inflation and unemployment, the two variables began to moved together, in an open defiance of the Phillips Curve.

The gradual emergence of stagflation and the progressive breakdown of the Phillips-Curve relationship presented mainstream macroeconomics with the most serious challenge since the Second World War. Macroeconomists launched a series of bitter attacks on the Phillips Curve, yet their criticism sought to modify, not nullify. Behind the theoretical Phillips Curve lay strong neoclassical convictions regarding the working of supply and demand. Although macroeconomics abstracted from the structure of underlying markets, the negative association between inflation and unemployment seemed to indicate that perfect competition was a useful assumption in the study of broad aggregates. The basic relationship between inflation and unemployment was simply too significant to discard. As a result, most macroeconomic challenges to the Phillips Curve have been half-hearted: they 'augmented' the elementary relationship with auxiliary factors.

By the early 1990s, after three decades of theoretical challenges, macroeconomic theories for inflation and unemployment still dominate the collective consciousness of economists and policy makers alike. In this sense, the struggle to save the Phillips Curve has been successful. Yet the achievement came at considerable cost. Amendments to Phillips Curve were never quite sufficient and additional modifications were constantly called for in order to accommodate changing realities. This repeated 'augmentation' of the Phillips Curve injured the apparent integrity of macroeconomics. The most serious damage, however, was caused by the nature of modifications. In order to explain the breakdown of the Phillips Curve, macroeconomists resorted to adversities such as 'disequilibria,' structural and informational 'imperfections,' and external 'shocks' delivered from outside the macroeconomic system. In other words, they abandoned the cardinal belief in equilibrium and perfect competition which previously characterized the 'neoclassical synthesis.'

In this chapter we deal with some of the key contributions to the macroeconomic literature on inflation and unemployment. Our aim is not to provide a comprehensive or even a partial survey. Instead, we focus our attention on fundamental methodological issues which arise as macroeconomists leave the ideal neoclassical domain of perfect competition and equilibrium and venture into alternative terrain. The first and second sections deal with the original Phillips Curve and its theoretical foundations. In the third section, we move from the labour market into the macroeconomic arena. The fourth section deals with the notion of structural imperfections. The fifth and sixth sections examine the integration of expectations and the natural rate of unemployment into the Phillips-Curve framework. In the seventh section, we appraise the rational-expectations framework. The eight section evaluates the effect of institutional instability on stagflation and, in the ninth section, we explore the notion of supply shocks.

2.1 The Original Phillips Curve

In 1958, A.W. Phillips published a careful empirical study examining the relation between unemployment and wage inflation in the United Kingdom over a period extending from 1861 to 1957. First he fitted a nonlinear function, negatively relating wage inflation to the rate of unemployment between 1861 and 1913 and then he demonstrated how this function could explain the relationship for

the subsequent period between 1913 and 1957. The stylized, stable relationship suggested that a 5.5 percent for unemployment was associated with zero wage inflation. When unemployment was above this threshold, there was a modest decline in nominal wages. On the other hand, when unemployment was below 5.5 percent, the rate of wage inflation increased rapidly.[1] Phillips also identified counter-clockwise 'loops' of data observations around the stylized fitted function. These loops indicated that when the rate of unemployment was falling, wage inflation exceeded the value given by the function and when unemployment was growing, the rate of change of wages was lower than values predicted by the function.

Phillips' results were assimilated quickly, partly because they provided strong confirmation for the working of competitive market forces, particularly for the way prices adjusted to 'excess demand' or 'excess supply.' The tentative theoretical hypothesis for this adjustment process is stated explicitly in Phillips' opening passage (1958, p. 283):

> When the demand for a commodity or service is high relative to the supply of it we expect the price to rise, the rate of the rise being greater the greater the excess demand. Conversely when the demand is low relatively to the supply we expect the price to fall, the rate of the fall being greater the greater the deficiency of demand. It seems plausible that this principle should operate as one of the factors determining the rate of change of money wage rates, which are the price of labour services.

Hence, it follows that if the rate of unemployment and its first derivative are taken as two independent proxies for 'excess supply' in the labour market, both should be negatively related to the rate of change in money wages. The rate of unemployment could explain wage inflation along the negatively-sloped Phillips Curve and the rate of change in unemployment would account for the counter-clockwise loops around it.

Most of the early literature that followed Phillips' original study emphasized this stylized relationship between wage inflation and unemployment but the Phillips Curve was significant also for what it failed to explain. In fact, Phillips took great pain to explain every *deviation* from the stylized loop. His explanations are interesting because they point to structural elements that are inconsistent with

---

[1] Phillips (1958, p. 290) fitted the following function to his data:

$\log (w + a) = \log b + c \log U$,

where w denoted the rate of change of wage rates and U measured the percentage unemployment. The estimated values for the parameters were 0.9 for a, 9.638 for b and 1.394 for c.

the assumption of perfect competition in labour and commodity markets. Several examples could be cited to illustrate this point. In the upswing between 1893 and 1896, for instance, wage rates rose more slowly than usual, a development that Phillips (p. 292) attributed to the rapid growth of employers' federations and the consequent rise in employers' resistance to trade-union demands. Similarly, the regular relationship was again disturbed in 1912, presumably by strike activity of union members in the coal-mining industry (*ibid.*). Another observation was the progressive narrowing of the cyclical loops between 1861 and 1909. Phillips (pp. 292-93) explained this in two ways; first, by the proliferation of wage-indexation and, second, by increasing time lags in the response of wage changes to changes in the level of unemployment. The significance of these lags, he argued, increased with the historical extension of collective bargaining and arbitration. Another illustration (pp. 293-94) points to the dramatic decline of wages in 1921 and 1922 (22.2 and 19.1 percent, respectively) which exceeded by far the moderate decreases suggested by the fitted curve. Phillips attributed much of these declines to automatic cost-of-living adjustments triggered by substantial decreases of import prices in those years. Finally, the observations for the 1948-1957 period appeared to generate a *reverse* loop, which Phillips (pp. 297-98) again explained by a lagged adjustment of wage rates to unemployment.

There is a common feature in these realistic supplementary explanations. Employers' federations, trade unions, collective bargaining, arbitrations, wage-indexation and lagged adjustments can be perceived as 'institutional rigidities' that *distort* the functioning of a *laissez faire* market system. In this sense, by recognizing such institutional realities, Phillips anticipated the subsequent dilemma that later macroeconomists often faced when they tried to relate the Phillips Curve to a changing world. The cost of being able to explain rising inflation (and, subsequently, stagflation) involved sacrificing the theoretical 'ideal' of perfect competition as its pristine simplicity was increasingly tainted by various realistic social and institutional 'distortions.'

2.2 From Disequilibrium to Equilibrium

The theoretical underpinning for Phillips' empirical findings was developed by Lipsey (1960). 'The usual argument,' writes Lipsey (p. 13), 'merely states that when there is excess demand . . . wage

rates will rise, while when there is excess supply . . . wages will fall. Nothing is said about the speed at which the adjustment takes place.' In other words, a theoretical framework where disequilibrium generates equilibrating forces is incomplete unless we specify a dynamic 'adjustment mechanism' to explain the *speed* at which the system moves toward equilibrium. Phillips indeed suggested that wage inflation was positively correlated with the magnitude of excess demand but, according to Lipsey (p. 2), he had not provided a 'model of market behaviour' that explained this relationship. Hence, in order to eliminate the potential for serious misinterpretation, the model underlying the Phillips Curve must be 'fully specified' (Lipsey, p. 12). In light of his emphasis on rigorous specification it is interesting to note that Lipsey does not specify the underlying market structure for his own model. Instead, he writes:

> We shall consider this relationship, first, for a single market, and then for the whole economy . . . We might analyze the market for *any* commodity since the argument at this stage is quite *general*. Since, however, the subject of the present article is the labour market we shall use the terminology appropriate to that market. (pp. 12-3, emphases added)

The use of such ambiguous language is unhelpful for it is hard to imagine a 'general' model for price adjustment that can be applied to 'any' market structure. The emphasis Lipsey puts on the role of 'excess supply' and 'excess demand' suggests that his own model may be applicable to perfect competition but is probably inadequate for other structures.[2]

The model for the single market contains three basic relations. One is the 'adjustment mechanism' which specifies the rate of change of wages as a linear function of the relative excess demand for labour:

$$(1) \qquad w = \alpha \, [(d - s) \, / \, s] \, ,$$

where w denotes the rate of change of wages, α is a fixed coefficient, d is the demand for labour and s is the supply of labour. The second relation is a curvilinear, negative function linking the rate of unemployment with the relative excess demand:

---

[2] In monopoly and monopolistic competition, there is no unique supply curve (supply depends on demand conditions) and, in oligopoly, the meaning of both supply and demand curves is ambiguous. Under these conditions there is no clear definition for *excess* supply or demand.

(2) $U = f_1 [(d - s) / s]$,

Equation (2) merely describes the relationship between the rate of unemployment and relative excess demand and it has no causal implications. When the market is in equilibrium (no excess demand or supply), there is only 'frictional unemployment,' with number of vacancies being just equal to the number of unemployed workers. When excess supply develops, unemployment increases linearly, while an increase in excess demand is associated with a curvilinear fall in unemployment (as excess demand increases, the fall in unemployment becomes progressively smaller because unemployment cannot become negative). The third relationship is the 'adjustment function' which is derived by combining equations (1) and (2):

(3) $w = \alpha f_2 (U)$.

This last equation is, of course, the standard Phillips Curve. It is interesting to examine the methodology employed in developing this model because some of its features reappear in subsequent macroeconomic theories of inflation and stagflation. Two aspects are worth noting: the central role assigned to non-observable variables, and the view that markets continuously move toward equilibrium. We consider each of these elements in turn.

According to Lipsey, the first problem for analysis stems from the continuous shifts of demand and supply curves, movements which make difficult the identification of these individual curves. Fortunately, he argues, this is not an unsurmountable obstacle for, in order to obtain Equation (1), it is 'only necessary to *know* demand and supply' at the existing market price and other points on curves can be ignored (p. 13, emphasis added). Note that even with this qualification, one may still ask the practical question as to how we could discover these two magnitudes. The theoretical analysis is cast in terms of supply and demand; that is, in terms of *desires*, or plans to sell and buy labour services. These

are psychological tendencies, not observable market outcomes. In this light it is unclear how could we solve the problem by limiting ourselves to the existing market price.[3]

Reliance on non-observable magnitudes introduces a strong axiomatic element into the analysis. Lipsey (p. 13) asserts that in order to *observe* the linear relation illustrated in Equation (1), 'it is necessary only that there be an unchanging *adjustment mechanism* in the market.' Unfortunately, even within Lipsey's own framework, this is only a necessary and not a sufficient condition, for in order to observe this relation we must first be able to observe the 'excess demand' variable. This is not always possible, admits Lipsey, but for practical not conceptual reasons. In his opinion, the difference between the number of unfilled vacancies and the number of unemployed workers could provide a 'reasonable *direct measurement* of excess demand' but, unfortunately, vacancy data are seldom available and even when they are available these data might be seriously flawed. As a practical solution, Lipsey suggests we relate excess demand only to unemployment (rather than to the difference between vacancies and unemployment). The solution is not very helpful, however, because Lipsey ignores the conceptual challenge altogether. His 'empirical' definition for supply and demand in the labour market is specified in terms of actual market outcomes rather than in terms of hypothetical desires. The quantity supplied is assumed to be equal to the sum of recorded employment and unemployment, while the quantity demanded is assumed to be equal to the sum of actual employment and vacancies. Hence, the difference between observed unemployment and vacancies is equal to 'excess demand' *by definition*. In this light, reliance on unemployment figures for want of vacancies data does not solve anything for it merely inserts an axiomatic link -- a negative curvilinear function between observable unemployment and non-observable excess demand -- in the theoretical chain.

The existence of these axiomatic elements weakens the scientific status of Lipsey's theory. One reason for developing this theory in the first place is that

---

[3] One could argue that 'quantity supplied' and 'quantity demanded' are observable when the existing market price is an equilibrium one because, when we define equilibrium as a 'chosen position,' we assert that the actual outcome is identical to the desired one. (See Asimakopulos, 1978, p. 43.) This reasoning, however, is quite misleading. Equilibrium here is defined in reference to desires and not the other way around, and unless we could *first* observe those desires we could not know that the market is indeed in equilibrium!

> if the relation ceases to hold, or changes, and we have no model to explain it, we can only say 'the relation has ceased to hold' or 'the relation has changed' and we will have learned nothing more than this. If we have a model explaining the relationship, we will know the conditions under which the relation is expected to remain unchanged. Then, if a change occurs, the model will predict *why* this has happened and this prediction will give rise to further tests from which we can learn. (Lipsey, p. 12)

Yet, can we really expect this model to tell us why the Phillips Curve changes? Consider, for instance, Lipsey's discussion of the impact that unions may have on the Phillips Curve (p. 17). In his opinion, unions may change the adjustment mechanism specified in Equation (1); for example, by making wage increases more responsive to excess demand and less responsive to excess supply. If this happens, the Phillips Curve itself should change. However, when we observe such a change in the empirical Phillips Curve, how can we know it originated from the influence of *unions* on the adjustment mechanism? For that purpose, any number of factors may affect the adjustment mechanism but we have no way of observing these effects because the adjustment mechanism itself remains defined only in terms of non-observable elements. Note that Equation (2) is also non-observable due to the presence of the excess demand variable. This introduces the further complication of not being able to associate changes in the empirical Phillips Curve with changes to either Equation (1) or Equation (2).

The second central feature of Lipsey's model is the emphasis on equilibrium. The labour market is subject to 'external' forces which shift demand and supply functions and create disequilibria. Fortunately, disequilibrium positions are inherently transient because the 'internal' forces, namely the 'laws of supply and demand' and the 'adjustment mechanism,' drive the system toward equilibrium. Wage inflation is the process by which stability is restored and, hence, even when wage inflation persists over lengthy periods of time, ultimately it is a *temporary* phenomenon -- it will disappear once equilibrium has been re-established.

The Phillips-Curve framework was rapidly incorporated into the mainstream of macroeconomics but this assimilation occurred amid criticism and consequent amendments. Attacks on the early Phillips Curve proceeded along two lines, both related to market 'imperfections.' One group of macroeconomists emphasized the significance of institutional rigidities in economic structure, while another analyzed the impact of imperfect information. We deal with these aspects in the following five sections.

2.3 Perfect Competition?

While Phillips (1958) and Lipsey (1960) focused their attention on the labour market, Samuelson and Solow (1960) suggested a macroeconomic framework by modifying the earlier formulation of the Phillips Curve. Instead of relating unemployment and *wage* inflation, the curve now linked unemployment with the overall *price* inflation. This 'modified Phillips Curve,' roughly estimated on the basis of 25 years of American data, was suggested by Samuelson and Solow (p. 192) as a 'menu of choice between different degrees of unemployment and price stability.' The relationship was considered to be significant because it appeared to be stable. This 'tradeoff relationship' suggested that the consequences of unemployment in terms of inflation (and vice versa) were *predetermined* and the politician had only to chose the desired combination that minimized social hardship (or maximized political gains).

While subsequent analysis of the Phillips Curve was concerned chiefly with such policy implications, much less attention was initially paid to the shift from wage to price inflation. Samuelson and Solow did not explain this transition explicitly and its rationale was only implicit in their article. Succeeding interpretations (for instance, Klein, 1967) used the assumption of a constant markup to explain this switch from wage to price. According to this later view, firms set their unit price as a *constant* markup over unit wage cost, so price inflation was just equal to wage inflation minus the growth in workers' productivity. Because productivity growth was relatively stable, price inflation could be interpreted as a relatively stable, linear function of wage inflation. In other words, you could move form the original to the modified Phillips-Curve equation simply by replacing wage inflation by price inflation on the lefthand side and subtracting productivity growth from the righthand side.

Of course, markup pricing was inconsistent with a rigid competitive model where prices respond to excess demand and supply. Indeed Samuelson and Solow argued that we must distinguish between the mechanism of demand-pull inflation which operated through competitive forces and cost-push inflation associated with 'market imperfections.' In its essentials, wrote Samuelson and Solow (p. 178), the demand-pull theory for inflation was based on the *a priori* presumption that real variables (real

outputs, inputs and relative prices for goods and factors) were determined by a set of competitive equations which were 'independent of the absolute level of prices.' The latter is determined by the money supply or, more broadly, by the overall level of money expenditures. This rigid neoclassical 'dichotomy' between the processes which determined real as opposed to nominal variables,

> would require that wages fall whenever there is unemployment of labor and that prices fall whenever excess capacity exists in the sense that marginal cost of the output that firms sell is less than the prices they receive. (p. 180)

Adherents of this position, wrote Samuelson and Solow (p. 177), were puzzled by the inflationary experience occurring between 1955 and 1958 in the United States. During that period, prices increased despite a growing overcapacity, slack labour markets, slow real growth and no apparent great buoyancy in overall demand. This historical episode was inconsistent with the conclusions of a strict competitive model so institutional friction and rigidities of the cost-push perspective gained a greater recognition:

> Some holders of this view attribute the push to wage boosts engineered unilaterally by strong unions. But others give as much or more weight to the co-operative action of all sellers -- organized and unorganized labor, semimonopolistic managements, oligopolistic sellers in imperfect commodity markets -- who raise prices and costs in an attempt by each to maintain or raise his share of national income, and who among themselves, by trying to get more than 100 per cent of the available output, create 'seller's inflation.' (p. 181)

Samuelson and Solow accepted the significance of these features and noted that

> to explain possible cost-push inflation, it would seem more economical from the very beginning to recognize that imperfect competition is the *essence of the problem* and drop the perfect competition assumptions. (emphasis added)

The introduction of a more realistic world-view into the macroeconomic framework enables Samuelson and Solow to use markup pricing as an implicit assumption for their modified Phillips Curve. The problem is that their modification requires that firms not only follow markup pricing, but also that the markup be *stable*, for otherwise, the modified curve need not remain fixed. Such instability will obviously destroy the explanatory power of the modified Phillips Curve and nullify its policy implications. On the other hand, the assumption of a fixed markup implies that Samuelson and Solow can partially conciliate demand-pull and cost-push theories: even when prices are 'pushed' by economic sellers in an imperfectly competitive world, stability of the *realized* markup indicates that, eventually, only the absolute costs and prices have risen while their relative levels remained unchanged. In other words, sellers' inflation does not cause a redistribution of sellers' incomes. Surely, this does not mean that

cost-push inflation is unrelated to 'real' variables as demand-pull theorists may argue. On the contrary, even with a fixed realized markup, reducing inflation has considerable costs in terms of unemployment and unused capacity. On this, Samuelson and Solow (p. 191) wrote:

> [I]f a mild demand repression checked cost and prices not at all or only mildly, so that considerable unemployment would have to be engineered before the price level updrift could be prevented, the cost push hypothesis would have received its most important confirmation.

Hence, the implication of the modified Phillips Curve that price stability requires a 'high' rate of unemployment is partly a result of social struggle between sellers in an imperfect world. But the struggle culminates not in redistribution *between* the different sellers but rather in the emergence of a cruel tradeoff between rising prices or curtailed output for society *as a whole*.

2.4 An Aggregate View of Market 'Imperfections'

During the 1960s, several researchers sought to encompass structural 'imperfections' into their empirical macroeconomic framework of the Phillips Curve. An early contributor to this literature was Perry (1966). His approach deserves a close examination because it was later adopted and extended by other writers, particularly in the National Bureau of Economic Research. Perry argues that the simple Phillips-Curve model where wage- inflation is explained by the single variable of unemployment is too restrictive. In the context of perfect competition, unemployment is a sufficient explanatory variable because '[a]ll economic forces must act on either the demand for or supply of labor, and their effect is already measured by unemployment' (p. 22), but in modern economies that are far from the 'competitive ideal,' wage inflation is affected by additional factors that must be considered. Hence, in a more realistic framework, writes Perry (p. 23),

> [e]ither the theory of adjustment must be modified or the assumption of perfect competition dropped. In fact, both can be done comfortably in the problem at hand with some confidence that we will be moving toward a more accurate specification of wage-determining process. . . . A model that acknowledges these points should yield more useful results, although it will necessarily represent a somewhat looser theoretical abstraction than the competitive one.

The question, of course, is what institutional features should be included to improve the simple Phillips-Curve relationship and how should they be modeled? Perry's answer to this question is

ambiguous. Initially he asserts that

> [t]he most realistic picture of the wage-setting institutions in manufacturing as a whole would undoubtedly include the *whole spectrum of degrees of market power.* In a few cases, the purely competitive model . . . might apply. At the other extreme, some wage bargains would be made under conditions virtual bilateral monopoly. In between would be various combinations of strong and weak labor bargaining units facing employers with different degrees of monopoly power in their product markets and monopsony power as hirers of labor. (p. 23, emphasis added)

But then such structural aspects are too difficult to deal with and Perry recants, quietly returning to the convenient world of aggregates:

> A theory explaining the behaviour of aggregate wages could not hope to encompass specifically all the different microeconomic theories of wage behaviour associated with these cases. But it need not do so to be effective for the present purpose. The problem *may* be intrinsically a macroeconomic one in the sense that the appropriate variables to explain changes in the general wage level may be *aggregate* ones, with any hypotheses about behavioral underpinnings at a microeconomic level affording no additional information. (*ibid.*, emphases added)

In other words, the industrial system suffers from a great many 'imperfections' but this should not introduce great theoretical and empirical hurdles. We can always assume either that the complex dynamics of 'monopoly power' are largely irrelevant to our question, or that the pertinent aspects of these dynamics *may* be reduced to movements of several 'aggregate' variables. In other words, market 'imperfections' need not be analyzed when they can be ignored or aggregated.

Wages in manufacturing industries are commonly set within a system of collective bargaining and, according to Perry (p. 50), this process for wage-determination can be adequately analyzed with the following aggregate equation:

$$(1) \qquad w_t = \beta_0 + \beta_1 U_t^{-1} + \beta_2 p_{t-1} + \beta_3 R_{t-1} + \beta_4 \Delta R_t + e_t ,$$

where w is the rate of change in money wages, U is the rate of unemployment, p is the rate of change of the cost of living (the CPI), R is the rate of profit on equity, ΔR is the change in the rate of profit, e is an error term and $\{\beta_i\}$ are fixed coefficients that need to be estimated.

What is the rationale behind Equation (1)? Perry argues that the rate of unemployment should be included in every realistic model because even under collective bargaining, excess supply still has a

negative effect on wage inflation. The three other aggregate variable -- increases in the cost of living, the rate of profit and the change in the rate of profit -- capture institutional imperfections introduced by collective bargaining. Higher values for such variables tend to strengthen the bargaining position of employees and soften the objection of employers toward workers' demands and, hence, each of these variables is expected to be positively related to wage inflation. Perry estimates the parameters of Equation (1) separately for durable-goods and nondurable-goods industries, as well as for the manufacturing sector as a whole, and finds that indeed they all have the expected signs and are different from zero at conventional significance levels.

This model, Perry argues, differs from earlier works which explored the impact 'structural' variables had on wage inflation. Those studies were deficient because, unlike his own model, they examined the isolated influence of each factor instead of their simultaneous effect.[4] In the context of a *multi*-variable model like that of Perry, one cannot interpret the empirical Phillips Curve between wage inflation and the rate of unemployment as a fixed relationship. Instead, it should be viewed as a *mutatis mutandis* locus of points taken from a *family* of curves. The position of each individual Phillips Curve depends on the other factors at play, namely, on the magnitudes of the last three carriers in Equation (1) and the values of their associated parameters.

Perry uses his model in order to explore the different possible relationships between the rate of unemployment, wage inflation and price inflation. For that purpose he assumes that we live in a 'stationary state' where the rate of profit is fixed (namely, $R_t = R_{t-1}$ and hence ΔR is zero), the rate of price inflation is fixed (namely, $p_t = p_{t-1}$), and the rate of productivity increases (ρ) is fixed (namely, $\rho_t = \rho_{t-1}$), and he further assumes that the price level (the CPI) is determined as a fixed markup over direct cost. With these postulates he then shows that wage inflation (w) and price inflation (p) each depend on the rate of unemployment (U), the rate of increase in productivity (ρ) and the rate of profit (R) as specified by the following equations:

---

[4] The earlier studies cited by Perry include Dicks-Mireaux and Dow (1959), Klein and Ball (1959), Bowen (1960) and Bathia (1962).

(2) $$w_t = \alpha_0 + \alpha_1 U_t^{-1} - \alpha_2 \rho_t + \alpha_3 R_t ,$$

(3) $$p_t = \alpha_0 + \alpha_1 U_t^{-1} - (1 + \alpha_2) \rho_t + \alpha_3 R_t ,$$

where $\alpha_i = \beta_i / (1 - \beta_2)$.[5]

If we use coefficient estimates from Equation (1) to assign values to each α_i in equations (1) and (2), we have enough information to assess the empirical implication of the model. The relationships among the different variables are given by the partial derivatives of each equation and Perry concentrates his attention on the basic Phillips-Curve relation between inflation and unemployment. Since the rate of growth of productivity is assumed to be fixed, the position of the Phillips Curve depends on the rate of profit. Note that price inflation is equal (by definition) to the difference between wage inflation and productivity growth. This enables Perry to use the same Phillips Curve to relate unemployment to *both* wage inflation and price inflation, with the difference between them being the fixed rate of productivity growth. Perry illustrates his approach by showing how lower rates of profit improve the tradeoff (causing lower unemployment for each level of inflation), and how higher rates of growth of productivity lead to both an improved tradeoff and lower price inflation associated with any rate of wage inflation (pp. 62-3).

The analysis indicates that policy makers may have more flexibility than initially assumed by Samuelson and Solow (1960). They can be satisfied with an existing inflation-unemployment tradeoff but they can also attempt to improve it. According to Perry (ch. 5), this can be done by affecting the variables or coefficients in equations (2) and (3). For example, governments can reduce corporate tax-rates or accelerate depreciation schedules in order to maintain existing cash-flows with a lower pre-tax rate of profit, or they can try to encourage productivity growth. They can also change the

[5] Under the stationary-state assumption, lagged values for the carriers in Equation (1) could be replaced by current values and the ΔR variable could be dropped. The assumption that prices are set with a fixed markup formula indicates that we can obtain Equation (2) by first substituting $w_t - \rho_t$ for p_t in Equation (1) and then solving for w_t. Similarly, Equation (3) can be derived by first substituting $p_t + \rho_t$ for w_t in Equation (2) and then solving for p_t.

institutional structure of wage and price determination by reducing the monopoly power of unions and firms, or by trying to persuade the general public toward a greater restraint.

Perry's model suffers from several shortcomings which arise because he acknowledges the significance of economic structure but then fails to deal with it effectively. First, the wage equations does not seem to reflect market 'imperfections' in any clear way. As we argued earlier, the observed rate of unemployment is not necessarily equivalent to the non-observable values for excess demand and, hence, there is room for other variables in explaining wage inflation even under perfect competition.

Second, Perry's explanation for price inflation is not constructed as a testable hypothesis but is rather based on the simple assumption that the aggregate price level is determined as a *fixed* markup over cost. Unfortunately, this assumption seems unwarranted for both theoretical and empirical reasons. The standard theory of the firm usually emphasizes the *ultimate* goal of maximizing return on investment. In this context, the markup is either an insignificant corollary of profit maximization or a means toward this end, but there is no reason to assume it is constant.[6] It seems only plausible for changes in the rate of profit to affect the markup. Indeed, why should firms be willing to grant larger wage increases that lower their markup when their rate of profit increases, but not attempt to raise their markup after the rate of profit falls? The empirical data for most capitalist economies clearly indicate that markups of price over prime-cost fluctuate through time. Under these conditions, why would one still insist on a fixed-markup assumption? Perry provides no explicit answer to this question but notes that it is the 'neutral standard' (p. 64). In other words, by assuming a fixed markup we imply that inflation has no effect on income distribution (it is 'neutral' in this sense) and a serious complication is resolved before it even arises. Unfortunately, these methodological manoeuvres are quite costly because they invalidate most of Perry's conclusions about the Phillips Curve tradeoff. His model indicates that

[6] In the model for perfect competition, firms are price takers not price makers. When the market price changes they alter their output in order to equate the new price with their marginal cost, but this also causes the average markup to change. In the long run, perfectly competitive firms reallocate their capital and production to follow the highest rate of profit and this often implies changes in the average markup. The standard model for monopoly also suggests that the markup changes with demand conditions when the monopolist equates marginal revenue and cost. For oligopolies, the results are more ambiguous; when oligopolies compete, interdependency between them may lead to any one of an infinite number of possible markup levels, whereas when they cooperate, they may set and alter the markup according to some arbitrary 'target' rate of return.

the root of price inflation is in the wage determination process, but that may be true only if we accept his assumption for fixed markup pricing. Otherwise, in the absence of a testable hypothesis about the markup, the pricing equation is incomplete and, hence, the Phillips-Curve tradeoff between unemployment and price inflation is unstable.[7]

A third problem concerns Perry's assumption that 'aggregate relationships exist' (p. 57). He agrees that wage determination in different industries may rely on different factors linked by different functional relationships, but argues that they can be safely ignored from a macroeconomic perspective. This assumption is unwarranted and may lead to misleading empirical results. For example, Perry (pp. 30-1) stipulates that 1/4th of all wage contracts are negotiated in each quarter, so the annual arithmetic average of wage inflation is a function of annual arithmetic averages for the carriers in Equation (1). This assertion has no empirical basis and, as Rowley and Wilton (1974) demonstrate, the particular distribution of wage settlements through the year has a dramatic effect on the sign of estimated coefficients, their magnitude and their associated levels of significance. Accounting for other aspects of heterogeneity (such as types or industrial activity or corporate size) will only introduce further instability into Perry's model.

The fourth problem we deal with is the assumption that underlying relationships between the variables are stable. Perry begins his dissertation by disassociating himself from the stable model for perfect competition and ventures toward a greater recognition of structural 'imperfections.' He concludes his analysis by arguing that the government can try to affect the Phillips-Curve tradeoff by altering the underlying economic structure. However, if the government can affect institutional patterns of wages, prices and profits why should we assume that these patterns are stable to begin with? For example, to have a stable Phillips-Curve relationship we need to have a stable rate of profit and Perry's use of only four different rates (10.0, 10.8, 11.8 and 12.5 percent) may give the incorrect impression that this rate

[7] This potential instability is heightened when Perry (p. 64) agrees that '[a]ctual price behavior may not conform to this standard' and discusses the possible implications of deviations from a fixed markup. For instance, when half of all prices increase 'autonomously' by 2 percent (independently of changes in cost and productivity), the position and slope of the Phillips-Curve between unemployment and inflation are altered. The problem, as Perry (p. 68) admits, is that this result is only *hypothetical* and 'has no empirical foundation.'

is indeed stable. According to Figure 3.7 (p. 48) however, the rate of profit during the 1948-62 period fluctuated between 8 and 16 percent! Unfortunately, the rate of profit in Perry's model is 'exogenously given' and, hence, such temporal fluctuations make it hard to predict inflation and unemployment, or design policy to improve the tradeoff between them.[8] The source of instability is not limited to the rate of profit. The parameters in Equation (2) and (3) are also determined exogenously by the underlying institutional structure and Perry does not explain why they should remain stable over time. These comments indicate that in order to analyze the effects of institutional structures on aggregate unemployment and inflation, we must first carefully analyze these structures, something that Perry failed to do.

The study by Perry suggested that there was not one but many potential 'Phillips Curves,' each corresponding to a particular set of institutional parameters. These underlying parameters were presumed to be relatively stable and, unless the government affected their values, the tradeoff between inflation and unemployment could remain stable over a substantial period of time. This idea of stability was not unanimously accepted. Several mainstream macroeconomists argued that indeed there were many potential Phillips Curve, though the reason for this multiplicity was to be found in informational, not structural 'imperfections.' Furthermore, while there were many possible Phillips Curves, all of them were inherently unstable.

2.5 Expectations: Economic Agents Strike Back

From the early 1960s, many developed capitalist economies began to experience rising rates of inflation with little or no decline in the rate of unemployment. This was a significant development because it put into question the time-honoured link between scarcity and price movements. Was it possible for prices and wages to be independent of excess supply or demand? According to Friedman (1968) and Phelps (1968) the answer was negative but the reason was not 'structural imperfections.' Phelps (p. 678) argued that most existing explanations for wage movements (like that of Perry)

[8] In his discussion of the dynamic properties of his model, Perry specifies an equation for changes in the rate of profit but does not explain the rate of profit itself (pp. 90-2).

contained countless independent variables in numerous combinations and it was difficult to choose among the different models because they often lacked any clear rationale. Instead, he suggested we move toward a 'unified and empirically applicable theory of money-wage dynamics,' where individual markets were competitive but economic outcomes were still 'distorted' because the flow of *information* was imperfect.

According to Friedman and Phelps, the vertical Phillips Curve did not constitute an anomaly in economic theory simply because the very construction of this curve involved a basic confusion: unemployment depended on *real*, not *nominal* wages and prices. The nuisance for economic theory, wrote Friedman in his Noble Lecture (1977, p. 12), was that nominal and real values need not move together:

> Low unemployment would, indeed, mean pressure for a higher real wage -- but real wages could be higher even if nominal wages were lower, provided that prices were still lower. Similarly, high unemployment would, indeed, mean pressure for a lower real wage -- but real wages could be lower, even if nominal wages were higher, provided prices were still higher.

So why did earlier observations indicate that the Phillips Curve was negatively sloped? Friedman and Phelps answered this question by making the curve a special case within a broader theoretical framework. The argument of the two theorists was similar and we focus mainly on the work by Friedman (1968; 1977).

Because information regarding employment opportunities and the availability of workers is costly (Stigler, 1961; 1962b) and because workers possess specific human capital (Becker, 1964), employees and employers enter into explicit or implicit *long-term* contracts. Although both sides seek to denominate their agreement in real terms, most labour contracts are signed in nominal dollars.[9] Consequently, the real wage over the life of the contract depends on an *unknown* future price level. Under these conditions, the desired nominal wage rate is set equal to the product of the desired real wage and the *expected* price index. The hallmark of the new theory, then, is this emphasis on price expectations formed by economic agents. If agents are always successful in correctly anticipating future

[9] Some collective agreements incorporate a COLA clause but the relative significance of such contracts has often been limited. Contracts can also be 'reopened' in special circumstances.

prices, the realized real wage is always equal the desired one. Since in this case the real wage is independent of inflation, it follows that the rate of unemployment -- which responds only to the real wage -- is also independent of inflation. Inflation ceases to be neutral, however, when economic agents err in their predictions. When price changes are unanticipated, the realized real wage differs from the desired real wage that is embodied in labour contracts, and until these long-term contracts expire, employment and unemployment deviate from their equilibrium relationship with real wages.

Why do errors in expectations lead to a tradeoff between inflation and unemployment? According to Friedman (1977, p. 13), the answer could be found by examining how both workers and employers *misinterpret* the effect of an unanticipated change in market conditions. For example, when the growth rate of nominal aggregate demand increases unexpectedly, each producer feels this increase primarily through rising demand for his own commodity. Although there is an overall expansion, the single producer myopically misinterprets it as an improvement in his own relative position. He believes that his own prices will be rising faster than the overall price level and, hence, is willing to raise the wage rate to attract additional workers. Workers fall in a similar trap when they believe that their wages increase faster than prices in general. As a result,

> a rise in nominal wages may be perceived by workers as a rise in real wages and hence call forth an increased supply, at the same time that it is perceived by employers as a fall in real wages and hence calls forth an increased offer of jobs. (*ibid.*)

For the economy as a whole the net result is a new position with lower unemployment and higher wages and prices. In other words, the economy moves up and to the left on the graph for the Phillips Curve. But this new situation is inherently unstable because it is based on an open 'lie.' If nominal demand continues to grow at its new higher pace, producers and workers will eventually realize they have been fooled by the market. The price for their own commodity is indeed rising but so too are all other prices and, hence, the *real* price for their commodity may not change at all! With this new, correct information, unemployment becomes artificially low. As agents adjust their expectations and revise their contracts to reflect the new rate of inflation, the Phillips Curve itself moves upward. The curve will stabilize in its new higher position when all contracts embody the new rate of inflation. When this happens, the economy will return to its original, 'natural rate of unemployment':

> At any moment of time, there is some level of unemployment which has the property that it is consistent with equilibrium in the structure of *real* wage rates . . . The 'natural rate of unemployment,' in other words, is the level that would be ground out by the

> Walrasian system of general equilibrium equations, provided there is embedded in them the actual structural characteristics of the labor and commodity markets, including market imperfections, stochastic variability in demands and supplies, the cost of gathering information about job vacancies and labor availability, the costs of mobility, and so on. (Friedman, 1968, p. 8)

Hence, the Phillips-Curve tradeoff is only a temporary relation based on the element of surprise. The authorities can use this tradeoff to reduce unemployment below its 'natural rate' only because they can fool all the people some of the time. But such efforts are self-defeating because no one can fool all the people all of the time. Eventually, economic agents will strike back, forcing policy-makers to cope with the original level of unemployment coupled with a *higher* rate of inflation. In the long run, there is no tradeoff and the Phillips Curve is vertical. Unemployment can thus be kept below its natural rate only at the cost of *accelerating* inflation.

The roots of inflation, then, are not 'imperfections' in the market structure. Such imperfections, to the extent they exist, affect mainly the natural rate of unemployment and beyond this influence, the market operates largely as a perfectly competitive system. Variations in demand and supply for factors or products can change only relative prices, so the source of *overall* price increases must be exogenous increases in available means of payment. Inflation is *caused* by expansionary demand policies when governments try to keep unemployment at an artificially low level, but it is *perpetuated* through expectations. In other words, inflation persists because agents expect it to persist.

Several features in this expectation-adjusted Phillips Curve are worth noting and deserve close examination. First, Friedman (1977, p. 12) emphasizes that 'only surprises matter.' It is the surprise of *unanticipated* inflation which confuses economic agents and causes them to misinterpret their relative situation. Friedman (1977, p. 13) explains that both workers and their employers 'are likely to adjust more slowly their perception of prices in general -- because it is more costly to acquire information about that -- than their perception of the price of the particular good they produce.' But this appears contrary to common experience. In practice, overall price indexes are published monthly and announced in the printed and electronic media. The cost of finding out what is the overall rate of inflation is surely redundant, especially considering the crucial significance of this information for the formation of

long-term contracts. Furthermore, in a perfectly competitive market, sellers and buyers are assumed to be 'price takers,' so why should they believe that their own price rises faster than prices of other market participants? Clearly, such a collective error cannot stem from a perfectly competitive framework.

Second, it is not clear why a rise (decline) in the rate of expansion of nominal aggregate demand *must* lead to an increase (decrease) in the rate of inflation. Friedman (1977, p. 13) asserts that when aggregate demand increases, commodity prices rise (or are expected to rise) and producers raise their wage offers to workers. This may be a likely outcome if we assume that commodity and labour markets operate at 'full' capacity and employment to begin with, but there is very little reason to expect prices and wages to rise when capacity utilization is 'very' low and unemployment is 'excessively' high. Of course, if the increase in demand growth is sufficiently large, bottlenecks may eventually be reached and, as we approach the 'natural' rate of unemployment, prices and wages may start to rise. In this light, the expectation-adjusted Phillips Curve involves a circular argument: an increase in the growth of nominal aggregate demand cannot cause a permanent reduction in unemployment because unemployment is *already* at its permanent 'natural' rate! If the economy indeed operates as a perfectly competitive Walrasian system, then excess nominal demand could lead only to rising prices as the neoclassical dichotomy asserts. But under such assumptions, the expectation-augmented Phillips Curve cannot be used to *prove* that there is no long-run tradeoff since this was already *assumed*. Friedman's expectation theory then merely asserts how the 'real' economy supposedly shields itself from the influence of 'monetary' forces.

Third, the assertion by Friedman (1977, p. 12, emphases added) that if 'everyone *anticipated* that prices would rise at, say 20 per cent a year, then this anticipation would be *embodied* in future wage (and other) contracts,' is impossible to prove. Friedman argues that 'real wages would then behave precisely as they would if everyone anticipated no price rise, and there would be no reason for the 20 per cent rate of inflation to be associated with a different level of unemployment than a zero rate.' This could be a meaningful assertion for a hypothetical economy where the real wage is equal, by definition, to both the marginal product of labour and the marginal disutility of work. In such an economy the real wage is clearly independent of the overall rate of inflation, but reality is slightly more complicated than

this fictitious world. In practice, the marginal values for productivity and utility are not observable and there is a continuous dispute between employers and employees on the 'appropriate' level for real factor prices. The determinants of real wages are quite 'arbitrary' and may involve elements of 'power.' There is no basis for an *a priori* assumption that factors such as 'bargaining strength' are independent of inflation, even when this inflation is fully anticipated by all sides. Furthermore, even if we ignore these difficulties, the statement by Friedman is still irrefutable because in practice we cannot distinguish between anticipated and unanticipated inflation.

To illustrate these predicaments, consider the following hypothetical example. Suppose General Motors and the United Auto Workers' union agreed for a nominal wage increase of 25 percent over the term of the contract and, suppose further, that the actual rate of inflation over that period was 20 percent. Could we test the proposition that this rate of inflation was in fact 'fully embodied' in the contract? To do that we must know whether or not both sides had the *same* anticipation for inflation, whether or not they expected this rate to be 20 percent, and whether or not the negotiations proceeded in 'real terms,' independently of these expectations. Unless we have all of this information, the neutrality proposition cannot be proven.

Fourth, the introduction of additional non-observable variables further diminishes the scientific character of the Phillips Curve framework. Friedman (1968, p. 10) quite openly admits that we *cannot* know what the natural rate is. 'Unfortunately,' he writes 'we have as yet devised no method to estimate accurately and readily the natural rate of either interest or unemployment.' A further complication is introduced when Friedman asserts that the natural rate of unemployment is not fixed and 'will itself change from time to time.' Under these assumptions, where the 'natural rate' is an invisible moving target, the hypothesis of a vertical long run Phillips Curve cannot be refuted. For instance, suppose that the government increases the pace of growth of nominal aggregate demand and, some time later, unemployment declines and inflation rises. Proponents of the natural-rate hypothesis can argue that the fall in unemployment was in fact a reduction in the natural rate itself and, hence, government policy was merely inflationary, precisely as predicted by the theory. This reasoning raises one simple but disturbing question: what empirical observation will be *inconsistent* with the natural-rate theory? The argument is

'flawless' simply because it cannot be empirically refuted! So unless we can specify the conditions under which this hypothesis *fails*, the natural-rate framework must be viewed as a mere tautology. Some macroeconomists such as Gordon (1985) and Fortin (1989), for instance, have attempted to estimate the natural rate of unemployment from regression analyses based on the expectation-adjusted Phillips Curve. Such estimates cannot be used to test the natural-rate hypothesis for inflation because the latter was already assumed to be valid when the estimates were derived.

Expectations create another serious problem for measurement because, like demand and supply, they also cannot be observed directly. If expectations and, hence, changes in expectations cannot be observed, how could we test the hypothesis that such adjustments cause the Phillips Curve to shift? Many economists attempted to tackle the problem by simply substituting *specifications* for *observations* but, unfortunately, they only replaced one problem with another. For example, suppose we impose an adaptive expectation mechanism on market prices and discover it has a substantial explanatory power. Can we conclude on the basis of such evidence that prices are determined by adaptive expectations of *market participants*? The answer to this question is negative because the statistical framework contains observations on prices but not on expectations. In fact, we never demonstrated that economic agents form adaptive expectations (or any other expectations), or that they act on the basis of such expectations. For that matter, current prices are 'determined' by past prices and market participants play no explicit role in the model!

The fifth and final issue concerns the 'neutrality' proposition associated with the expectation-augmented Phillips Curve. The statement by Friedman that demand policy cannot have a permanent 'real' effect on the economy has been challenged by several macroeconomists,[10] but their criticism refers mainly to the final impact and ignores the initial nature of the policy itself. Consider what happens when the government increases its demand for goods and services by raising military spending, for instance. Most of the new orders will typically go to a group of 50 or 100 corporations which, in turn, will subcontract some of the work to a few hundred additional firms. The remaining companies in the economy will be excluded from this initial injection of spending. Or consider the direct

[10] See Buiter (1980, pp. 39-40) for a summary statement on such criticism.

effect of open market operations by the central bank. In the United States, government bonds are not evenly distributed between households in the economy but rather are concentrated mainly in the hands of large institutional investors. An attempt of the central bank to increase the money supply by buying bonds requires that the bank bid up their prices. So the immediate beneficiaries of this monetary expansion are the large institutional investors while other economic agents remain unaffected. Clearly, the direct effect of such macroeconomic policies is to alter the existing distribution of income, assets, production and relative prices between market participants. In fact, it is hard to think of a single macroeconomic policy which does *not* have such initial 'real' effects on the economy.

To summarize, the expectation-adjusted Phillips Curve and the related natural-rate hypothesis are based on some rigid explicit or implicit assumptions regarding economic structure and scientific methodology. The economy is assumed to operate 'as if' it was a Walrasian competitive system where agents responds to 'real' stimulus and are impartial to 'nominal' ones. When macroeconomic demand policies are executed, their initial effect is assumed to be evenly distributed among all economic agents, so as not to upset the original 'real' structure of the market. An increase in the pace of aggregate demand growth causes inflation to accelerate because markets already operate at full capacity and employment. Information about aggregate price and inflation indices is available at no cost, but price-taking sellers and buyers are nevertheless confused by this initial turn of events and fail to realize that prices around them rise as fast as the price of their own commodity. As a result, they increase their supply and demand for products and factors and cause the overall level of unemployment to fall below its natural rate. Ultimately, agents discover their collective error and seek to reduce demand and supply as soon as their long term-contracts expire. This causes a gradual upward shift in the Phillips Curve reflecting the adjustment of expectations and contracts to the new level of inflation. When the adjustment is complete, the economy returns to its original, 'real' Walrasian equilibrium but with a higher rate of inflation. Unfortunately, this process of adjustment cannot be tracked down because both expectations and the natural rate of unemployment are not observable.

2.6 In Quest for Information: The Unemployed as an Investor

Although the 'natural' rate of unemployment could not be observed empirically, many economists still felt it was a crucial concept which deserved rigorous theoretical elaboration. The first systematic discussion on the topic appeared in an important collection of articles edited by Phelps in 1970 and titled *Microeconomic Foundations of Employment and Inflation Theory*. In the introduction, Phelps (1970, pp. 4-5) talked about a major theoretical breakthrough in the making:

> The theoretical departure that is common to these otherwise neoclassical papers is their removal of the Walrasian postulate of complete information, . . . [and] . . . With the postulate of perfect information removed, the way is at last open to formal study of general disequilibrium.

With this minor 'informational' amendment, the persistence of unemployment was no longer to be perceived as a condemnation of capitalism and an embarrassment to neoclassical theory. Instead, unemployment became a *desirable* aspect of economic activity and an integral part of conventional theory.

To set a framework for the new microfoundations, Phelps (1970, p. 6) describes our economy as a collection of islands. Competition on each individual island is impeccable:

> [L]abor is technically homogeneous in production functions and indifferent among the many heterogeneous jobs of producing a variety of products. Producers on each island are in pure competition in the labor and product markets. Each morning, on each island, workers 'shape up' for an auction that determines the market-clearing money wage and employment level.

Unfortunately, the virtues of such system are distorted because the flow of information between different islands in the archipelago is not free. According to Alchian (1970, p. 29), information is a commodity like any other and, as such, it is subject to standard economic laws of production and cost:

> Dissemination and acquisition (i.e., the production) of information conforms to the ordinary laws of costs of production: faster dissemination, or acquisition costs more . . . [and] . . . Like any other production activity, specialization in information is efficient. Gathering and dissemination information about goods or about oneself is in some circumstances more efficiently done while the good or person is not employed, and thus able to specialize (i.e., while specializing) in the production of information.

Phelps' archipelago economy presents no exception to these postulated rules. Presumably there are no modern means of communication (such as telephone, newspapers or telex) between the islands and, hence, workers who want to know more about job offers must 'specialize' in gathering this information

by rowing from island to island:

> To learn the wage paid on an adjacent island, the worker must spend the day travelling to that island to sample its wage instead of spending the day at work. (Phelps, 1970, p. 6)

In this context, unemployed workers rowing between the islands are not seeking 'jobs' but 'job information.' According to Alchian (1970, p. 30):

> Jobs are always easily available. Timely information about the pay, working conditions, and life expectancy of all available jobs is not cheap. In a sense, *this* kind of unemployment is self-employment in information collection.

Since jobs are always available, worker are under no pressure to accept any particular offer. Instead, the choice of employment is based on a careful optimization strategy. Like any other investor, a typical worker in the archipelago tries to maximize the present value of his investment, namely, of his labour power. Under certain circumstances, this worker may find it highly advantageous to withdraw the services of his commodity. Such unemployment then constitutes a form of investment activity. In his taxonomy for different types of unemployment, Phelps (1972, p. 3) candidly suggests to categorize the motives for unemployment 'much as economists are used to classify people's motive for holding money.' He argues that when workers avoid the workplace they are involve in one or more of the following forms of unemployment: 'search unemployment,' 'precautionary unemployment' (or 'wait unemployment'), 'speculative unemployment,' or 'queue unemployment.' Let us briefly examine each of these concepts.

At any point in time, a labourer has a certain perception about the distribution of wage rates in the archipelago. On the basis of this perception, he formulates what Holt (1970a, p. 96) designates as the 'wage aspiration level.' A worker with strong entrepreneurial drives who finds current wage offers to be below his own aspiration may chose not to work and instead row between the islands and sample different job offers. By 'searching,' the unemployed worker produces information necessary to update his perception of the wage distribution and his associated wage-aspiration level. Of course, the search is costly, mainly because the worker does not earn money while searching. When a current wage offer exceeds the difference between the wage aspiration level and the cost of continued 'searching,' the search is called off and the worker accepts the job offer.

A worker who wants to follow a more 'precautionary' investment strategy has a second alternative: he can specialize in 'waiting.' Gordon and Hynes (1970) for example, approach this strategy as an inventory management problem which they first apply to landlords and then, with an equal vigour, to workers. The owner of an apartment building commonly leases his apartments for a fixed period of a year. This landlord can always lease his vacant apartments at a low rent, but then he runs the risk of not being able to rent them at a higher price later, if demand picks up. The worker faces much the same problem because Gordon and Hynes make the (unrealistic) assumption that labour contracts are also binding for a fixed period of time. Under such conditions, a worker who accepts a job at less than his wage-aspiration level could find himself locked in a disadvantageous position if demand for labour revives. Hence, although such workers can easily find work at a substandard wage, they might decide to 'accept leisure' and enter the passive state of 'wait unemployment' until they receive a 'proper' wage offer.

In between the aggressive 'search unemployment' and the precautionary 'wait unemployment' there is a third form of investment: 'speculative unemployment.' According to Phelps (1972, p. 3), when a worker is engaged in precautionary unemployment, he or she chose to wait for the *unpredictable* arrival of a more lucrative job offer. A worker may also withhold his labour services for speculative reasons when he or she *predicts* that future offers will indeed be higher than current ones.

Finally, the unemployed may perceive himself as standing in a queue with other unemployed workers waiting to be hired. Workers in the queue are ordered by employers according to their perceived skills and, naturally, it is the low-skilled workers who are likely to suffer the longest spells of unemployment. This type 'queue unemployment,' first discussed by Thurow (1969), is different from the previous ones because the unemployed worker does not think he can obtain a job by reducing his wage rate. Yet according to Phelps (1972, p. 29), even 'queue unemployment' stems, at least in part, from the curse of imperfect information:

> There is some question of whether queue-unemployment can stand for long as a distinct type of unemployment. It tends to blend into the other types if we acknowledge that most workers of however little skill could, perhaps only after lengthy and arduous search, reasonably expect to find employment somewhere, in some kind of paying job,

at some wage not beneath consideration.

Hence, the problem is that workers at the bottom of the queue become (irrationally) desperate and then, 'to our distress we find that labor markets are less imperfect than we thought' (*ibid.*).

When information is more costly to obtain when employed, individual workers may voluntary chose to invest in unemployment in order to search or wait for such information. Eventually, however, when sufficient data are obtained, these individuals will accept job offers and there should be no unemployment. In other words, the 'natural' rate of unemployment should converge to zero. At first, this explanation appears to be inconsistent with the observation that actual rates of unemployment are always positive, but this appearance is deceptive because it refers to a static world. In a dynamic economy, argue adherents of the new microfoundations, the numerous supply and demand curves for individual commodities are never stable. The continuous stochastic shifting of such curves means that new information about job opportunities is constantly being generated. Workers are aware of this viability and, naturally, devote some of their time toward productive unemployment in quest for new information about fresh opportunities. In summarizing views on the subject, Phelps (1970, p. 17) argues that a certain rate of unemployment is not only 'natural' but also *desirable* for our economy:

> It would be as senselessly puritanical to wipe out unemployment as it would be to raise taxes in a deep depression. Today's unemployment is an investment in a better allocation of any given quantity of employed persons tomorrow; its opportunity cost, like that of any other investment, is present consumption.

Hence, instead of a direct attack on unemployment through aggregate demand policies, Holt (1970b) recommends to decrease market 'friction.' This could be done by policies that improve economic stability, increase search efficiency and introduce computer-aided counselling and placement, for example.

The notion of unemployment as investment in information is not limited, however, to the 'natural' rate of unemployment. According to proponents of the new microfoundations, unemployment may deviate from its natural rate, but this difference also stems from a rational choice by workers to seek further information. An explanation for this phenomenon, consistent with the expectation-augmented Phillips Curve, is outlined by Phelps (1970, pp. 6-7). When aggregate demand in the archipelago economy falls, workers are mislead to believe that this decline is at least partially

specific to their own island (recall the lack of perfect, costless information). Consequently they intensify their quest for information by increasing their search and wait activity. Unemployment rises above its natural rate until workers finally realize (as the new information is collated and analyzed) that their investment was futile and then go back to work, this time with a lower nominal wage rate. Similarly, an increase in aggregate demand will cause workers to reduce their search or wait activity and will generate a temporary fall in unemployment below its natural rate. Note that because they are unable to distinguish relative from aggregate changes, the underlying strategy of workers is perfectly rational, despite its subsequent failure.

So far, the discussion emphasized voluntary aspects of unemployment. Considerable unemployment is generated, however, when firms lay off workers who presumably would like to retain their current jobs. Such unemployment would appear as 'involuntary' yet, according to Alchian (1970, p. 39), this may be a misleading interpretation. In his opinion, even layoffs can be attributed to latent unemployment aspirations of employees! To illustrate his argument, Alchian considers the hypothetical case where, after demand for cars dropped, General Motors lays off 20,000 workers without even negotiating with them the possibility of a temporary wage cut. One may blame labour unions or assert that these workers could not be employed profitably at any wage rate, writes Alchian, but layoffs are a 'sensible' policy quite independent of such qualifications. In his opinion, General Motors lays off workers because it knows workers will simply leave if their wages are cut:

> Employers learn that wage cuts sufficient to justify profitable maintenance of the prior rate of output and employment would be too deep to keep employee beliefs about alternatives. And so layoffs are announced without fruitless wage renegotiations.

The views about voluntary unemployment examined in this section were criticized almost as soon as they emerged in the early 1970s.[11] Here we wish only to stress the inadequacy of these new microeconomic foundations for empirical research. The problem arises because unemployment is explained in reference to human 'motives' but these are unknown. The argument that workers 'voluntarily' chose to become unemployed can be accepted or rejected as an article of faith, but it cannot be proven or refuted because the psychological drives of workers are not observable. Note that even

[11] See for instance the accounts by Gordon (1976), Hall (1980) and Tobin (1972).

layoffs cannot be considered as leading to 'involuntary' unemployment because such layoffs are considered to be quits in disguise. Indeed, Lucas (1978, p. 355) goes even further to asserts that

> it does not appear possible *even in principle*, to classify individual unemployed people as either voluntary or involuntary unemployed depending on the characteristics of the decision problems they face. One cannot, even conceptually, arrive at a useable definition of full employment as a state in which no involuntary unemployment exists. (emphasis added)

Yet in his view, having no tools to distinguish between 'voluntary' and 'involuntary' unemployment is a methodological bliss, not a curse. In fact, there are considerable benefits to be gained once we accept that *all unemployment is voluntary* and discard the concept of full employment:

> First, one dispenses with that entire meaningless vocabulary associated with full employment, phrases like potential output, full capacity, slack and so on, which suggested that there was some *technical* reason why we couldn't all return to the 1890 workweek and produce half again the GNP we now produce. Second, one finds to ones relief that treating unemployment as a voluntary response to an unwelcome situation does not commit oneself to normative nonsense like blaming depressions on lazy workers. (p. 356)

The greatest benefit, however, is that policy-makers no longer have to be concerned with the average rate of unemployment because, by definition, this is also the 'natural' rate of unemployment. The focus thus shifts to preventing distortions that cause the actual rate to fluctuate around its natural level:

> On this view, the average (or natural, or equilibrium) rate of unemployment is viewed as raising policy issues only insofar as *it can be shown* to be 'distorted' in an undesirable way by taxes, external effects, and so on. Nine percent unemployment is then viewed as too high in the same sense that 2 percent is viewed as 'too low': both are symptoms of costly and preventable instability in general economic activity. (p. 353, emphasis added)

Unfortunately, Lucas replaces Keynes' vocabulary with barren unscientific jargon. If, as Lucas (p. 355) argues so forcefully, 'the "thing" to be measured [the natural rate] does not exist,' how could we discover its determinants? How could we distinguish changes in the natural rate itself from fluctuations around it? In this context, how could an unknown natural rate 'be shown' to be 'distorted' by policy? How could we establish whether 'distortions' to the natural rate are desirable or not? What may constitute desirable as opposed to undesirable distortions here? Lucas and other founders of the new microfoundations do not provide answers to these methodological questions.

2.7 A Rational Expectations 'Revolution'?

The apparent failure of stabilization policies during the 1970s influenced a growing number of macroeconomists to accept the neoclassical dichotomy between a stable domain of 'real' activity and an erratic environment of 'nominal' variables. Many began to argue that interventionist demand-policies were ineffective *even* in the short run and their sole effect was additional price instability. Governments were increasingly called to take their hands off the real economy and limit their activity to a stable expansion of monetary aggregates.

The gradual return to rigid pre-Keynesian convictions (now labelled as 'new classical' macroeconomics), was partly affected by developments in the theory of expectations. Although fierce opponents of stabilization policy accepted the 'natural-rate hypothesis,' they could not use the expectation-adjusted Phillips Curve as developed by Friedman and Phelps to fully support their point. That framework still allowed governments to affect the real economy in the short run and, unfortunately, that short run was much too long. According to Friedman (1968, p. 11), the 'temporary' effect of government policy could last anywhere between two to twenty years, so the case for stabilization policy could not be totally dismissed. To overcome this obstacle, adherents of the new classical approach needed to modify the expectation-adjusted Phillips Curve even further. In particular, they focused on the factors determining the speed at which the Phillips Curve shifted from one long-run position to the next. For Friedman, the reallocation of the Phillips Curve was not instantaneous because institutional arrangements (such as long-term contracts) created friction and because price expectations were slow to adapt to evolving reality. Hence, the nullifying of these two obstacles became an essential step toward accepting the conclusions of new classical macroeconomics.

The first of these impediments was removed by eliminating all institutional distortions and installing a new form of friction-free Walrasian system as the normal state of the economy. Lucas (1972) begins his seminal contribution to the new classical literature with the following paragraph:

> This paper provides a simple example of an economy in which equilibrium prices and quantities exhibit what may be the central feature of the modern business cycle: a systematic relation between the rate of change in nominal prices and the level of real output. The relationship, essentially a variant of the well-known Phillips curve, is derived within a framework from which all forms of 'money illusion' are rigorously

> excluded: all prices are market clearing, all agents behave optimally in light of their objectives and expectations, and expectations are formed optimally. (p. 103)

To derive his results Lucas (pp. 104-6) defines the 'structure of the economy' in highly abstract terms. In his economy there are N identical individuals, each of whom lives for two periods; each person has n units of labour and can produce n units of output; the output cannot be stored but can be freely disposed of; there exists a government with only one function, namely the issuance of fiat money; this money is transferred from the government to individuals in the beginning of the period and from individuals back to the government in the end of the period; there is no inheritance; finally, trade is carried out with an auctioneer at a single market-clearing price. This framework may be fascinating for intellectual reasons but its usefulness toward understanding the 'modern business cycle' is unclear. What can *pragmatic* macroeconomists and policy-makers learn from such a hypothetical economy that definitely never existed and will never exist? Lucas fails to deal with this question but from what he explains in a footnote (p. 105, emphasis added), it seems that these and other simplifications are necessary 'to keep the laws governing the *transition* of the economy from state to state as simple as possible.' In other words, the assumption that the economy is always in a 'state' of market-clearing equilibrium is admittedly artificial, but this moderate sacrifice of realism is fully justified because such 'abstraction' clears the way for the more important task of describing movements from one equilibrium to the next!

In the absence of any institutional rigidities, this movement from one state of equilibrium toward the next is governed solely by the way individual agents form their future expectations. Friedman (1977, p. 24) argued that because of prolonged pre-war price stability, individuals in the United States and the United Kingdom expected the 'normal' price level to persist. This element of inertia remained strong even when inflation began to increase and, consequently, individuals were systematically disappointed when their price expectations underestimated the actual changes. Since price expectations adapted only gradually, the drift of the Phillips Curve between successive long-run equilibrium positions was painfully slow.[12]

---

[12] A simple 'adaptive-expectation' mechanism can be described by the following equation:

$$e_t = e_{t-1} + \alpha(p_{t-1} - e_{t-1}), \qquad 0 < \alpha < 1$$

New classical macroeconomists criticized the validity of adaptive expectations because they implied that economic agents were hopelessly 'irrational': they continued to use a model which was likely to generate systematic prediction errors and they 'wasted' non-price information that could have been used to improve their price forecast. The critiques pointed out that if individuals were indeed rational decision makers, they should also formulate 'rational expectations.' The rational-expectation hypothesis was first suggested by Muth (1961, p. 316) who argued that 'expectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory.' In his opinion, this meant that

> expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the 'objective' probability distributions of outcomes).

Muth's emphasis on rational expectations was ignored by macroeconomists for over a decade until it was picked up by Lucas, Sargent and others in the early 1970s. Sargent (1973, p. 431) for instance, asserted that

> expectations of inflation are assumed to be endogenous to the system in a very particular way: they are assumed to be 'rational' in Muth's sense -- which is to say that the public's expectations are not systematically worse that the predictions of economic models. This amounts to supposing that the public expectations depend, in the proper way, on the things that economic theory says they ought to.

Since these early formulations, the idea of rational expectation depended critically on two key concepts: the 'objective distribution of outcomes' and the 'relevant economic theory' associated with it. Despite close to two decades of theorizing, these concepts remained surprisingly enigmatic. The language used in the rational-expectation literature is often cryptic and the emphasis on mathematical symbolism helps to further cloak substantive issues. In our examination we present some of the basic claims advanced in the rational-expectations literature and assess their merits.

---

where the actual rate of inflation is denoted by p, the expected rate by e and α is a fixed 'disappointment coefficient.' This means that individuals set their current expectation to last period expectations plus an allowance proportionate to last period's disillusionment. The speed at which agents 'learn' from their errors depends on the magnitude of the 'disappointment coefficient' α: a low value means a strong inertial bias and a high value indicates a short memory and quicker adjustment. Clearly, when the rate of inflation is rising (falling), adaptive expectations will underestimate (overestimate) inflation.

A simple description of the rational-expectation framework could run as follows: The economy is a closed system with its own 'laws of motion.' These laws of motion determine how the endogenous variables of the system interact with the exogenous and predetermined ones (in other words, these laws determine the reduced form for the simultaneous equation system).[13] The economic system interacts with other systems like 'nature' and 'politics.' These systems determine values for the exogenous variables. Some of these exogenous variables follow systematic patterns while others are random variables with given distributions. If history could have been 're-run' with given laws of motion, given values for the systematic exogenous variables and given values for the predetermined variables, it would have generated an 'objective distribution of outcomes.' The mean of this distribution would reflect the impact of predetermined and systematic exogenous variables, and the dispersion would be affected by the distribution of exogenous disturbances.[14] In accepting such a framework, the rational-expectations theorists merely follow the standard approach toward macroeconomic modelling. The difference between standard macroeconomic models and ones based on rational expectations stems from assumptions regarding what people know about the economic system.

In a world of rational expectations, people possess considerable knowledge about the system. They understand the system's laws of motion (in other words, they know the 'relevant theory' and the values of its parameters). They also know all about the past history of the system (they know the values for the predetermined variables). They further know the values for those exogenous variables which follow a systematic pattern. They do not know the values for the random exogenous variables but they know the distribution from which these variables are drawn. Under these conditions, a simple rational expectations hypothesis for inflation can be summarized by the following equations:

$$(1) \qquad e_t = E(p_t \mid I_{t-1}) ,$$

---

[13] The term 'laws of motion' (as used by Sargent, 1986, p. 3, for instance) refers to a description of a stationary process and has nothing to do with Marx's original reference to principles governing the dynamic transformation of society.

[14] The characteristics of this distribution are based on the conventional assumption that the mean impact of random shocks on endogenous variables is zero.

(2) $p_t = E(p_t \mid I_{t-1}) + u_t$.

In these equations the expected rate of inflation is denoted by e and the actual rate by p; E is the conditional expectation operator and I is the 'set of relevant information available' (an all-inclusive term for the 'relevant theory' regarding the 'laws of motion,' the parameters of that theory and the values for predetermined and systematic exogenous variables); finally, u is the effect on inflation of random exogenous shocks. Under these conditions, expectational errors stem only from these unpredictable shocks and have no systematic component.

What happens in a Walrasian, frictionless community of rational economic agents when, starting from equilibrium, the government attempts to increase the pace of growth of aggregate demand? If these intentions become known before they are executed (for example when the government follows a 'policy rule'), the effects of the policy are *immediately* neutralized by the counteractions of private economic agents. The reason for this 'policy-ineffectiveness' is straightforward. According to the neoclassical dichotomy between the 'real' and 'nominal' domains, the ultimate effect of demand policy is on the price level. Since this demand policy is part of I when expectations are formed, the impact of such policy on next period's prices can be accurately predicted by Equation (1). When agents adjust their 'real' supply and demand schedules in anticipation of the new policy measures, they at the same time make these measures ineffective (in other words, by altering their 'decision rules' they also alter the system's 'laws of motion'). This instantaneous adjustment means that policy does not inflict even a short-run disequilibrium and the economy shifts smoothly from one long-term equilibrium into the next. Note that 'surprise' policy can affect the real economy. In the absence of a 'policy rule' for instance, policy changes constitute a random shock to the system and affect prices through u. Since rational expectations do not account for such unpredictable jolts, the real economy is distorted by the nominal impetus. Fortunately, this effect is very short-lived because the execution of the policy makes it part of the system's laws of motion and, hence, an ingredient of the 'relevant theory.' The conclusion of this new-classical scheme resembles the famous *Catch-22*: in order to stabilize the economy, policy must be related to events in some systematic way. But a systematic policy is predictable and a predictable policy is neutral. To put it somewhat differently, in order to *stabilize* the economy the

government must be able to affect it, but this calls for an erratic, unpredictable policy which can only *destabilize* the economy! The circle is closed and the case against demand management is complete.

The rational expectation hypothesis has been often hailed as a 'revolution' in macroeconomic thinking. Many of its leading lights downplay their contribution, however, stating it is merely a natural evolution toward a greater consistency of macroeconomic models with basic microeconomic tenets. Taylor (1985, p. 393) asserts that macro-models with rational expectations are now the 'rule rather than the exception,' yet several key features suggest that embracing the new classical framework may in fact hinder rather than enhance our understanding of how a modern economy works. These aspects deserve some closer examination and we consider them now.

The first question concerns the 'relevant theory.' In a Lucas-type abstract economy, the problem does not even arise simply because the economy is *defined* by the theory, but in a complex, modern economy like that of the United States, the question can no longer be ignored. Reality has no enclosed set of blue prints and, indeed, economists rarely agree about it. There exists a rich menu of different theories and it is not clear which theory (if any) provides an accurate description of the economy's alleged 'laws of motion.'

A second question regards the assimilation of a 'relevant theory.' Even if a 'correct' theory does exist, why should it become common knowledge? Again, in a Lucas-type economy, agents are simply *assumed* to possess all the necessary information about the economy's blue prints and its historical evolution, but what occurs in a real economy? Muth (1961, p. 330) stated that expectations must be at least 'moderately rational' for otherwise 'there would be opportunities for economists to make profits in commodity speculation, running a firm, or selling the information to present owners.' In other words, by taking advantage of their superior understanding, economists turn their private correct theory into

common knowledge and the 'relevant' theory is assimilated.[15] There are two difficulties with this logic. One, when the economy is changing, the relevant theory of today need not be the relevant one for tomorrow and, hence, this process may mislead economic agents to adopt outdated views.[16] Two, the assimilation of theories has a very 'real' effect on the economy because it presumably *redistributes* income (particularly profit) from those who cannot read the market to those who can. Muth (p. 316, emphasis added) argues that a 'public prediction' has no substantial effect on the operation of the economic system '*unless* it is based on inside information' but his own view on the assimilation of market knowledge suggests that every relevant theory grows from 'inside information.' Hence, whether assimilated or not, 'relevant' theories must have a substantial impact on the economy.[17]

Third, the rational-expectations hypothesis asserts that people's expectations constitute part of the system's laws of motion. This implies that interdependency between the 'objective distribution of outcomes' and the 'relevant theory' is potentially destabilizing. Frydman and Phelps (1983) argued that the 'average opinion' of economic agents is one of the exogenous variable in the economic system, so when agents attempt to determine this 'average opinion,' they get entangled in an infinite-regress problem and may drive the system toward a permanent state of disequilibrium. Cagan (1983, p. 45), commenting on the same point, wrote that

> Maximizing behavior requires that economic agents can in fact find the maximum position on their own. If that position is affected by the expectations of others, I do not see that maximizing behavior under such circumstances, even with Bayesian learning, is any longer well defined.

---

[15] In arguing that irrational expectations are necessarily short-lived, Maddock and Carter (1982, p. 45) invoke the authority of Keynes (1930, p. 160) who wrote that 'actions based on inaccurate anticipations will not long survive experiences of a contrary character, so that fact will soon override anticipation except when they agree.' However, this merely suggests that people may realize they were wrong, not that they will necessarily learn from their mistakes. As argued bellow, the convergence of expectations toward rational expectation is not inevitable.

[16] Economists have been continuously altering their models yet their predictions published in the popular and scientific media do not seem to converge toward any single, 'correct' vector. For instance, a recent survey of 'What Economists are Predicting for 1990' published in *Business Week* for December, 25, 1989, reports 25 predictions for real-GNP growth ranging between a high of 5.1 percent and a low of -3.2 percent. Predictions for inflation range between 2.5 to 6.3 percent, predictions for the interest rate vary between 6.5 and 12.5 percent and prediction for the rate of unemployment run between 4.4 and 8.8 percent. Note that these predictions were not made by 'ivory tower' economists but by business economists working for large companies who stood to lose from erroneous forecasts.

[17] 'Inside information' on the stock market generated and continues to generate substantial profits but after such information is used, it becomes *useless* rather than *relevant* public knowledge.

Even Taylor, an orthodox adherent of rational expectations, admitted that '[b]ecause of the self-fulfilling feature of rational expectations, there is generally a continuum of solutions to rational expectations models' (1985, p. 419).

A fourth problem arises when we examine how the private sector responds to public-sector initiatives in a 'game theoretic' structure. For instance, if the government can revoke its policy commitments (when it follows an *unconstrained* rather than constrained 'policy rule'), the neutrality proposition fails. Kydland and Prescott (1977) argue that in a dynamic game between two agents (the private sector against the government rather that against 'nature'), rational expectations may lead to 'inconsistency of optimal plans.' Buiter (1980, p. 36) concludes that traditional optimal control techniques 'fail to take account of the impact of future policy measures on current events through the changes in current behaviour induced by anticipation of these future policy measures.' This cultivated language conveys a simple message: when human beings are allowed discretion and there is some interdependency between their economic decisions, there may be no 'objective distribution of economic outcomes.'

Fifth, the rational-expectations framework focuses on how private-sector agents *respond* to public-sector initiatives, while little or no attention is paid to dynamic *initiatives* in the private sector itself. This choice of emphasis is common in much of the macroeconomic literature on expectations but it is striking in the new classical writings. In its crude formulation, the rational-expectations hypothesis examines only one type of initiative: government attempts to change aggregate demand. Every other economic action is 'automatic.' Private agents with a fixed set of preferences are locked in their uncompromising drive to maximize utility. To achieve this goal under prefect competition they must follow one pre-determined course of optimal action. Theirs is a 'game of man against nature' where nature changes 'technology' and man responds following fixed, known rules of conduct. If we discard this perverted animism and recognize that initiative, discretion and interdependency exist in the relation between agents such as firms, consumers, workers and investors, we open a pandora's box of disturbing questions. For example, what rational expectations can agents formulate on a world dominated by oligopolies with complex business ties? What prices should we expect to see when managers tell us they follow a rule-of-thumb in setting profit markups? What are the expected 'objective outcomes' from

attempts by private agents to form coalitions or to influence the government toward a redistribution of income?

To our knowledge, there is no definition for the 'objective distribution of outcomes' in the rational-expectations literature. The idea seems to imply that the experience of our economy in any 'sample period' is generated by some specified 'laws of motion,' and that this actual 'history' is merely one observation drawn from a infinite sample of potential outcomes, with a stable mean and a given dispersion.[18] This framework becomes meaningless when we view the economic process as a qualitative *transformation* or evolution rather than a 'draw' from a *stationary* process. When there is human initiative, historical change has few if any 'deterministic' components and even rational agents cannot 'jump over Rhodes' to discover the future. 'About these matters,' argued Keynes (1937, p. 185) 'there is no scientific basis of which to form any capable probability whatsoever. We simply do not know.'[19] These criticisms should not be interpreted as suggestions towards improvements of the rational-expectations framework. We believe that new classical economics is barren and misleading, and that theoretical 'improvements' to this approach are simply further steps in the wrong direction.

The danger of accepting the legitimacy of such 'improvements' is illustrated by recent attempts to incorporate seemingly 'realistic' features into a rational-expectations theory. The prominence of new classical ideas also brought them under the magnifying glass of macroeconomists. Scholars like Tobin (1980), Buiter (1980), Frydman (1981) and Gordon (1981) argued that the policy-ineffectiveness conclusion depended not only on the assumption of rational expectations but also on the existence of a Walrasian, market-clearing system of prices. When a system with sluggish wage or price adjustment was substituted for the Walrasian construct, the short-run Phillips Curve reappeared even under rational

---

[18] Note that these presumptions underlie the notion of 'functional relationships' in the social sciences and are common in conventional econometric approaches to estimation, testing and predictions of macroeconomic models. The significance of the rational-expectations framework is in making these presumptions explicit.

[19] According to Georgescu-Roegen (1979, p. 322), the most notable feature of the economic process is the continuous emergence of *novelty*, or qualitative change. Unfortunately, he argues, 'no analytical model can deal with the emergence of novelty, for everything that can be derived from such a model can only concern quantitative variations . . . noting can be derived from an analytical model that is not logically contained in its axiomatic basis.' Contrary to the new-classical euphoria, Georgescu-Roegen concludes that 'we cannot possibly have a bird's eye view of the future evolution of mankind' (p. 325).

expectations (recall that institutional rigidities were one of the elements in Friedman's model). Fischer (1977), for instance, introduced multiperiod contracts in the labour market and concluded that the authorities could affect real variable provided the policy duration was shorter that the length of contracts. Phelps and Taylor (1977) reached a similar conclusion when they examined the consequences of prices and wages being set one period in advance. In these models *future* prices are set to clear the market on the basis of *current* information but when new information about policy arrives, prices are too 'sticky' to adjust immediately and the policy becomes effective. Talyor (1979) introduced overlapping, or staggered wage contracts into the rational-expectations framework and concluded that policy can be effective even if its announced lead-time is longer than the duration of the longest contract (Taylor, 1985, p. 414).

According to Taylor (1985, p. 411), the algebra of these models retains the long-run neutrality of policy but allows the same policy to be effective in the short run. Hence, such models can be viewed as attempts to resolve what Gordon (1981, p. 509) labelled the 'persistence dilemma' of the rational-expectations hypothesis. The acknowledgment of contracts and price stickiness may appear to reconcile the rational-expections hypothesis with persistent deviations of actual unemployment from its trend. Unfortunately, this aura of realism is a rather deceptive decoration for a barren axiomatic model that has very little to do with dynamics of complex market structures. Taylor must be aware that real-life contracts have numerous institutional and dynamic aspects which cannot allow stable ARMA representations. Yet, finding such time-invariant representations are crucial for his model so real contracts must give way to axiomatic ones, where all dangerous actuality has been conveniently removed. The model apparently dresses in 'realism' while, in fact, it is shallow.

2.8 'Institutional Instability' and Stagflation

The history of the Phillips Curve could be described as an ongoing duel between reality and theory, in which the cunning of history has proven to be no match to the ingenuity of macroeconomists. When, during the late 1960s and early 1970s, inflation accelerated with no apparent decline in unemployment, macroeconomists responded by modifying the downward-sloping Phillips Curve into a

vertical one. To do so, they introduced expectations and the 'natural rate' axiom into the framework. Subsequently, when history staged a combination of rising inflation and rising unemployment, macroeconomists responded by trying to bend the Phillips Curve into an upward-sloping position using concepts such as 'institutional instability' and 'exogenous shocks.' We consider these latter modifications in this and the following section.

In his 1977 Nobel lecture, Friedman asserted that the vertical curve could survive the new reality of stagflation with only a 'modest elaboration of the natural-rate hypothesis.' The element missing from his own original formulation was the requirement for the rate of inflation itself be *stable*. When the same rate of inflation prevails for 'many decades,' wrote Friedman (p. 24), we could expect that prices be fully anticipated and fully adjusted. These conditions for a vertical Phillips Curve are likely to be met in what Friedman calls the 'long-long run' but the interim phase of transition toward inflation stability may involve some unpleasant complications. The increase in the rate of inflation during the post-war period in Europe and the United States also brought with it increased *fluctuations* in that rate. Friedman speculates that this increase in inflation instability led to rising *institutional* instability, whereby the optimum length of unindexed commitments was shortened, the efficiency of the price system in coordinating economic activity was reduced, public policies became increasingly confused, and the extent of government intervention in free markets was greatly increased. Friedman argues that such developments had adverse consequences for economic efficiency, but he admits that they do not really explain the apparent drift of unemployment.[20] In other words, accepting the proposition that the Phillips Curve is vertical in the *long-long run* does not help us resolve the puzzle of *contemporary* stagflation.

In our opinion, the weakness of Friedman's analysis stems not from his failure to further amend the Phillips Curve framework but rather from his very attempt to do so. Friedman (1977, pp. 7-8) implies that his theoretical manoeuvres are constructive steps in scientific progress but it seems that, instead of directing us towards better understanding of stagflation, his 'patching-up' leads us into a

[20] Some authors (like Fischer, 1981) tested and rejected the presumed link between inflation, inflation instability and unemployment.

theoretical vacuum. Each successive interpretation of the Phillips Curve turns the existing construct into a 'special case' of a 'more general' framework. The 'short run' that extended from the late 19th century and until the middle of this century became a special case of a 'long run' that embraced us between the late 1960s and early 1970s, but even this 'long run' was merely a subset of a 'long-long run' phase which we entered in the mid-1970s. The first shift was created when economists discovered that information was 'imperfect.' The second transformation was instituted when economists realized that institutions were slow to 'adjust.' This leads us to pose one simple question: if economic life amounts to a continuous and progressive 'departure' from some enigmatic equilibrium relationships and if these relationships will be valid only in some *imaginary* future when stationarity replaces history, why should such equilibrium relationships be useful in explaining *real* phenomena? The predicament is well illustrated in Friedman's own writings. On the one hand, he painfully acknowledges the 'real' consequences of a high, variable inflation:

> . . . some groups gain . . . other lose. . . . The society is polarized; one group is set against another. Political unrest increases. The capacity of any government to govern is reduced at the same time that the pressure for strong action grows. (p. 26)

On the other hand, he has very little to say on these issues, since conflict of interests and continuous redistribution cannot be integrated into a framework which patiently looks forward toward some long-long run state of bliss, when full 'adjustments' and restored social harmony reinstate the neoclassical dichotomy between inflation and unemployment.

2.9 The Stagflationary Menace of 'Exogenous Forces'

Although macroeconomics was criticized during the 1970s for its failure to effectively deal with stagflation, macroeconomists were not ready to take the blame. Blinder (1979, pp. 3 and 5-6) for example, insists that there is nothing wrong with macroeconomics for, by using the very rudimentary aggregate demand and supply curves, one can provide a 'fairly simple and general theory of stagflation' that 'can indeed explain what has happened.' In his opinion, critiques have often erred by confusing the problem with its solution: stagflation could be easily explained but it could not be easily cured. Moreover, politicians generally failed to understand this and made a difficult situation even worse.

Within what has now become the standard macroeconomic model, stagflation arises either as an adjustment process following an earlier expansion of aggregate demand, or as a result of adverse contraction in aggregate supply. Both cases begin and end in a long-run macroeconomic equilibrium, but they differ in the source of disruption and in the way the economy responds to it. Consider the first case, where the initial equilibrium is upset by an 'autonomous' expansion of aggregate demand. In the short-run, the expansionary demand 'shock' causes output to rise beyond its 'natural' rate with no parallel increase in prices. This is a false tranquillity, however. As time passes, the economy moves into an 'intermediate run' and input prices begin to rise, pulling output prices with them. Unfortunately, this is not the end of the story. Eventually, output starts to *fall* because, by definition, the economy must converge to its 'potential,' or 'natural' rate of output. Hence, we move through three phases in the following order: growth without inflation, growth with inflation and, finally, stagnation with inflation, or stagflation. 'Because wages and prices move sluggishly,' asserts Blinder (1979, p. 14, emphasis added), 'real output *must* overshoot its *eventual* position' and stagflation is merely the inevitable process by which this 'eventual' position is reached. The explanation also suggests that stagflation is in fact implicit in the augmented Phillips-Curve framework: an expansionary demand-policy causes the economy to climb up the short-run Phillips Curve, but the subsequent shift of the curve causes unemployment to *increase* back to its 'natural' rate in the midst of *rising* inflation.

While most macroeconomists accepted the theoretical validity of demand-induced stagflation, it was the 'supply-shock' rationale that captured their imagination. Why complicate the analysis, many asked, when the 'laws of supply and demand' offered the most simple solution to the stagflation riddle? If prices and output move in opposite directions, it was only natural to associate this outcome with changes in supply, not in demand. A 'supply shock' which shifted the intermediate and long-run aggregate supply curves to the left, would cause stagflation with rising prices and falling output. Moreover, since we assume that the natural rate of output itself is reduced, the situation is often irreversible and the adverse effects of the original shock may be with us 'for ever' (Blinder, 1979, p. 16). Finally, a supply shock may create a lengthy wage-price spiral that will further aggravate the initial effects of that shock. All of this means that when supply-shocks hit the economy, politicians are cornered into a policy nightmare:

> The limited capability of policy to influence supply poses a particulary vexing problem in a stagflationary world since any stabilization policy adopted in response to stagflation is bound to aggravate one of the problems [inflation or unemployment] even as it helps cure the other. Such is the policy dilemma of stagflation. (Blinder, 1979, pp. 20-1)

Many macroeconomists were excited by this alternative theoretical avenue though only few were fully aware of its wider methodological implications.

Both aggregate demand and aggregate supply are 'convenient' tools for analyzing the neoclassical synthesis. In this framework, we can always argue that aggregate demand increased or that aggregate supply decreased and it is practically impossible to refute such assertions since 'desired' magnitudes for spending or production are not observable. Yet beyond this convenience, there lies a disturbing asymmetry between the two concepts of supply and demand. It seems that aggregate demand can shift for a host of 'subjective' reasons; for instance when consumers change their 'preferences' or 'propensities,' when investors experience a burst of 'animal spirits,' or when politicians make an 'autonomous' policy move. Thus, since the world of demand is supposedly at the mercy of human impulse, it can be easily blamed for much of our instability. A similar hypothesis for supply is not very convincing, however. The aggregate supply curve often emerges from a rational, efficient sphere of activity with no room for destabilizing elements of human fancy. Shifts in the curve occur for 'objective' reasons, such as changes in the production function or the availability of factors of production. This asymmetry poses an obstacle for a supply-based theory of stagflation, for how could the turmoil of stagflation originate from this stable domain of activity?

Disturbing as this question might have been, few macroeconomists were discouraged by its implications. For decades, macroeconomics made an efficient use of assorted 'imperfections' to patch up the theory of aggregate demand, and there was very little reason not to use this very approach in making necessary adjustments to the theory of aggregate supply. Supply prices depend on factor costs. In an ideal neoclassical world, such input costs are 'endogenous' to the system for they emerge as simple derivatives from the production function: the wage rate is equal to the marginal product of labour, the rate of profit is equal to the marginal product of capital, and so on. Unfortunately, noted many macroeconomists, our own market system was far from this ideal because some factors had the power to set their prices higher than their corresponding marginal products. In principle, such imperfections

could distort the pricing of every factor yet practical macroeconomists prefer to emphasize the pivotal role of raw materials and labour. Bruno and Sachs (1985, p. 7) are typical when they point their first 'blaming finger' at the weather and the oil sheiks:

> A clear and central villain of the piece is the historically unprecedented rise in commodity prices (mainly food and oil) in 1973-74 and again in 1979-80 that not coincidentally accompanied the two great burst of stagflation.

When the raw-material price shock hits the system, it causes the aggregate supply curve to shift to the left, raising prices and lowering output. The turbulence could have been lessened somewhat, argue Bruno and Sachs (ch. 1), if other factor prices were fully flexible. With such flexibility, an increase in the price of raw materials would have led to a reduction in the use of those inputs, to a consequent decline in the marginal product of accompanying factors and, hence, to a subsequent fall in the prices of these latter factors. Such cost reductions would have created a compensatory rightward shift in the supply curve and could reduce the severity of stagflation. Unfortunately, the price of the most important factor -- labour -- is far from being flexible, at least in the downward direction. When a raw-material price shock creates a burst of stagflation, workers not only refuse a necessary reduction in their real wages but they also demand and obtain real-wage increases! This causes the supply curve to shift even further to the right.

For many macroeconomists, labour is responsible for more than just aggravating an ongoing stagflation. According to Blinder (1979, p. 14), for instance, workers can generate their own supply shock when they 'suddenly become more aggressive and demand higher wages.' Bruno and Sachs (1985, p. 7) associate this undesirable power with evolving 'institutional rigidities' in the labour market:

> . . . one of the variables that set the stage for the 1970s stagflation was the rise in union power and militancy at the end of the 1960s . . . . A real wage boom resulted, which started a squeeze on profits even before 1973 . . . . It strikes us as misguided to consider the labor market as a perfectly competitive bourse when in almost every OECD economy much of the labor force is unionized and governments play an enormous role in affecting labour compensation.

The supply-shock theory for stagflation raises many interesting questions and we consider some of them now. First, the argument that supply shocks are created by 'excessive' factor prices has no empirical meaning. For instance, Bruno and Sachs (p. 178) argue that 'an important supply factor has been the persistent excess of real wage levels above the marginal product of labor at full employment'

and then devote an entire chapter to estimate this 'wage gap.' Under the heroic assumptions of 'output-clearing markets and competitive firms' and together with knowledge of the production function, this would have been a mundane task. In reality, admit the authors, there are 'technical limitations' which make this a somewhat difficult exercise. For example, markets may not clear 'on a year-to-year basis' and there are 'data problems' regarding the capital stock. We may also add that production functions as frontiers have a vague empirical meaning, that marginal productivity is not observable, and that we rarely if ever reach full employment where the level of marginal productivity should be measured. Despite these unsurmountable obstacles, Bruno and Sachs remain undeterred and proceed with a simple 'practical' solution. They observe that both unemployment and real wages were higher during the 1970s and early 1980s than during the late 1960s. Next, they make the convenient but totally arbitrary assumption that, during the 1965-69 period, unemployment was at its 'full employment' level and hence that wages were at their 'right level' (i.e., there was no 'wage gap'). Finally, they use various hypothetical production functions and measures for productivity changes to estimate by how much the actual wage exceeded the 'full-employment wage' in subsequent years. Naturally, they find that the 'wage gap' in most OECD countries was positive after 1973, but then how could it *not* be positive under these definitions?

A second question concerns the source of different supply shocks. If the weather, Arab oil-sheiks and labour unions can engineer a supply shock, why should we not explore the possibility of a 'corporate profit-shock'? Bruno and Sachs (pp. 19-20) agree that, in principle, the roles of labour and capital are 'entirely symmetric' but suggest that this is not a matter for concern in practice. In their opinion, world supply and demand for saving determine the real rate of interest on world markets and 'competition among firms in the economy will ensure that the rate of profit will eventually equalize itself to this external rate of interest.' The picture emerging from this set of presumptions is perplexing. Most key industries in OECD countries are oligopolistic and large companies interact with each other in many different markets. Furthermore, market structures and the interrelations between large corporations experience continuous changes. Finally, governments are involved with these firms through procurement, subsidies, loans, taxation, the granting of certificates and so on. According to Bruno and Sachs, however, all these institutional features can be safely ignored. The combination of union power and government

involvement affects relative and aggregate wage levels in the labour market but, for some mysterious reason, market power and government activity in the product market have no similar consequences for the rate of profit. The return on capital simply cannot be contaminated by rigid institutions. Despite the heterogeneity of their experience, firms are somehow compelled to adjust their actions in order to ensure their own rate of profit converges toward the 'normal' world rate. Unfortunately, even this absurd assumption is not very helpful for empirical analysis. The 'normal rate of return' is supposedly determined by supply and demand for saving but these are neither observable nor stable. As a result, we can never verify that the 'normal' (average?) rate of profit is indeed an equilibrium value equal to the marginal productivity of capital. Since Bruno and Sachs admit that markets can occasionally be out of equilibrium, it is possible to have a profit shock even in this framework.

Third, the discussion suggests that supply shocks generate a redistribution of income from the 'shocked' to the 'shocker.' In analyzing the U.S. case, Blinder (1979, p. 17) draws attention to a 'massive redistribution of real income away from urban workers and toward farmers and oil producers.' Presumably, the latter have a lower propensity to spend than the former, so the effect of this redistribution is to reduce aggregate demand. Blinder (p. 18-9) is quick to point out that such demand-reducing effects of supply shocks are 'not permanent':

> The farmers who do the high saving are probably accumulating the means to finance subsequent investments in their farms, not to add to their estates. Oil companies will not sit on top of a pile of cash for long. They will either pay it out in dividends (to stockholders who will then spend it), spend it on additional investment goods, or use it to finance internally some investment projects that would otherwise have been financed externally. Like the oil companies, the OPEC nations too cannot be expected to allow the massive buildup of liquid assets to continue indefinitely. Gradually, these countries can be expected to find more and more ways to spend their oil earnings, thus returning demand in the form of exports to the countries that lost demand in the form of consumption . . . [hence] . . . For the long run, we have only the permanent shift of the supply curve to contend with.

The notion that income redistribution can have temporary but no lasting effect on the level of economic activity is inconsistent even with standard Keynesian views. The failure of oil companies (or other firms that increase prices faster than costs) to promptly re-invest their increased savings will lead to a fall in overall levels of activity, which may further reduce the incentive for future investment. A serious recession can easily eliminate previously-accumulated 'piles of cash' so their eventual long-term investment may never materialize. Similarly, there is no reason to assume that petrodollars accumulated

by OPEC and subsequently spent on western products had only a 'neutral' effect on economic activity. For example, we may discover that dollars earned from U.S. consumers may have been spent by OPEC countries on European-made products, or that price inflation for finished goods may have eroded the real purchasing power of perviously earned petrodollars. These considerations are of utmost importance but are generally neglected in the supply-shock literature.

A fourth and final issue concerns the notion of 'exogenous' shocks and its relation to equilibrium analysis. The spectra of increasing price instability, higher rates of unemployment and, finally, the puzzle of stagflation presented macroeconomists with a difficult dilemma. They could try to explain these as 'endogenous' phenomena but then this would amount to admitting the economic system was inherently unstable. Alternatively, they could maintain their stubborn emphasis on equilibrium and blame all the havoc on 'exogenous' forces that jolt the system. Most macroeconomists chose the second avenue but, by doing so, they have effectively admitted that explanation for important aggregate phenomena lied outside the realm of mainstream macroeconomic theory itself! Furthermore, an emphasis on equilibrium could appear meaningful when the economy is generally stable, with only occasional 'disruptions.' When there is a continuum of dynamic instability, however, attempts of macroeconomists to depict it as a rapid transition from one equilibrium to the next seem rather pathetic.

2.10 Concluding Comments

The progressive disintegration of the Phillips Curve helped unveil some pristine simplicities which characterized the 'neoclassical synthesis' of microeconomics and macroeconomics. The belief in the equilibrating force of perfect competition proved decreasingly useful in an era of stagflation. In their attempts to defend their paradigm, mainstream macroeconomists were forced to transcend previously sacred boundaries and acknowledge that underlying microeconomic structures and non-equilibrating changes were significant for macroeconomic analysis.

Alas, the departure of macroeconomists from equilibrium and perfect competition seems hopelessly circumscribed. For most macroeconomists, the deviation from ideal market conditions, even

when such a deviation persists for a long period of time, is an *exception*. For them stagflation is ultimately an alien phenomenon. Its roots lay not in the 'economic system' but rather in impediments imposed on that system. Given this assessment, it is not surprising that most attempts to examine the broader structural causes and implications of inflation and unemployment were perceived as challenges to mainstream macroeconomics. We examine some of this structural literature in the next two chapters.