DISCUSSION PAPER NO. 250

THE THEORY OF DOMESTIC INFLATION

by

Robert J. Gordon

October 1976

American Economic Review Proceedings
February 1977

American Economic Review Proceedings

February, 1977

THE THEORY OF DOMESTIC INFLATION

Robert J. Gordon*

Authors and readers of the thousands of articles and books published on inflation during the past decade may regard as audacious any attempt to survey the theory of domestic inflation in 3000 words. But far from requiring an apology, this format forces concentration on central issues and justifies skipping second-order questions. More leisurely expositions and extensive bibliographies are provided in recent surveys by David Laidler and Michael Parkin and by Robert J. Gordon (1976). The ground rules for this paper are a limitation to theory rather than empirical tests, to closed rather than open economies, and to causes of inflation rather than costs, consequences, or cures.

I. Inflation and Money in the Long Run

A simple set of definitions helps to separate noncontroversial from controversial issues. We begin with a national income identity, expressed in growth-rate form:

$$y \equiv p + q, \tag{1}$$

where lower-case letters represent rates of growth, and y, p, and q stand for, respectively, the rates of growth of nominal income, the aggregate price deflator, and real output. Subtracting the long-term trend growth rate of capacity (q^*) from both sides of (1), we obtain:

$$y - q^* \equiv p + q - q^*, \tag{2}$$

or

$$\hat{y} \equiv p + \hat{q},$$

where $\hat{y} = y - q^*$, and $\hat{q} = q - q^*$. Arthur Okun (1962) was the first to

establish the statistical relation now widely known as "Okun's Law" between the current unemployment rate (U), last period's unemployment rate (U_{-1}), and the output growth deviation ($\hat{q}$):

$$U = U_{-1} - \hat{q}/a. \tag{3}$$

When (3) is solved for $\hat{q}$, the result is substituted into (2), and then (2) is solved for the rate of inflation (p), we have:

$$p = \hat{y} + a(U - U_{-1}). \tag{4}$$

The sources of change in y can be decomposed if we once again invoke an identity:

$$\hat{y} \equiv \hat{m} + v, \tag{5}$$

where $\hat{m}$ is the growth rate of money adjusted for capacity growth ($\hat{m} = m - q^*$), and v is the growth rate of velocity. Combining (4) and (5), we obtain:

$$p = \hat{m} + v + a(U - U_{-1}). \tag{6}$$

Once the economy has settled down at any given unemployment rate ($U = U_{-1}$), the rate of inflation depends only on the adjusted growth rate of money ($\hat{m}$) and the growth rate of velocity (v). Shifts in fiscal policy can cause one-time-only changes in velocity, as even Milton Friedman (1966b) recognized long ago, but cannot cause permanent changes in the growth rate of velocity. Innovations in transactions technology, as well as an income elasticity of the demand for money differing from unity, could make v positive or negative, but these factors appear to exhibit only modest changes insufficient to account for marked accelerations or decelerations

in inflation.

Changes in the adjusted growth rate of money are thus isolated as a necessary concomitant of long-run changes in the inflation rate. It is in this carefully qualified sense that Friedman (1966a, p. 18) correctly labelled inflation as "always and everywhere a monetary phenomenon." But despite the attempts of some less subtle monetarists to treat this quotation as settling all questions, in fact it represents only a starting point. Accelerations in monetary growth are not usually autonomous whims of central bankers. In most classic wartime or postwar money-fueled inflations and hyperinflations, the role of the monetary authority has been passively to finance deficits resulting from the unwillingness or inability of politicians to finance expenditures through conventional taxation. In the same way, a "cost push" by unions or firms must be ratified continuously by the monetary authority if inflation is to continue.

A more general view, explicitly set out in Melvin Reder's classic analysis, attributes inflation to the passivity of the monetary authority in the face of a "tripartite" set of pressures emanating from all groups in society--labor, management, and government. Gordon (1975c) extends this theme by distinguishing the "demand for inflation," i.e., monetary accomodation, caused by government's refusal to tax and by pressure groups which attempt to increase their income share, from the "supply of inflation," the degree of response to these pressures, a result of the political balancing of the votes likely to be lost from higher inflation, as against the vote cost of the higher unemployment consequent upon a policy of non-accomodation.

II. The "Missing Equation"

For anything other than long-run analysis, equation (6) is incomplete. Even if $\hat{m}$ and v are known, there are two remaining unknowns (p and U) but only one equation. A decade ago it was usual to close the model by adding a Phillips curve:

$$(7) \qquad p = bp^e + f(U), \qquad 0<b<1, \quad f'<0.$$

Together (6) and (7) determine a menu of p,U combinations for different $\hat{m}$. It was common in the U. S. for economic advisers to Democratic Presidents to recommend a combination with higher p and lower U than the target of Republican advisers.

Friedman (1966a, p. 60) was the first explicitly to reject (7) and to state that "there is no long-run, stable trade-off between inflation and unemployment." On the grounds that workers supply labor by evaluating the expected real value of a wage offer, and that the expected and actual price levels and inflation rates cannot diverge in equilibrium, Friedman (1968), and Edmund Phelps (1967) argued that in equilibrium with $p = p^e$ only a single "natural rate of unemployment" (U^N) is possible:

$$(8) \qquad p = p^e + g(U - U^N), \qquad g'<0, \; g(0)=0.$$

The "natural rate hypothesis" (NRH) as embodied in (8) completely changed the framework of stabilization policy. No longer could an Administration choose its own favorite point on the p,U tradeoff curve. A rate of unemployment below U^N could not be achieved by aggregate demand policy through manipulations of $\hat{m}$, because inflation would continuously accelerate as long as p^e responds adaptively to past changes in p:

$$(9) \qquad p^e = h(p_{-1}, p_{-2}, \ldots)$$

A permanent reduction in actual unemployment could be achieved without

accelerating inflation only by operating directly on U^N, through manpower programs and other subsidies to reduce worker-job mismatch, and through reductions in the minimum wage and in other barriers to the flexibility of relative wages. It was not widely understood that the NRH did not establish a link between inflation and money where none existed before. Instead, p and $\hat{m}$ are linked together in (6), whether or not the "missing equation" is provided by the old-fashioned trade-off curve (7) or the NRH (8).

III. Short-run Price Inflexibility and the Role of Contracts

An important criticism of the NRH has been its apparent lack of validation in recession and depression episodes. Combining (8) and (9), a deceleration of inflation requires that actual U exceed U^N, since p^e cannot fall until p itself first experiences a decline. A period during which U remains above U^N for a substantial period should be characterized by an accelerating decline in p. But during the Great Depression the unemployment rate remained above 8.5 percent for twelve straight years in the U. S. without the slightest sign of such an acceleration. If the function g( ) in (8) were completely flat for high values of U, the NRH would remain valid only as long as U were kept below the flat range. Even if g( ) retains its negative slope in the range of U relevant for current policy, a relatively gentle slope nevertheless would make extremely costly any attempt to "beat the inflation out of the system" by the deliberate creation of a recession.

Until recently the apparent downward inflexibility of prices during periods of high unemployment constituted an empirical phenomenon in search of a theory. Okun (1975) distinguishes between "auction" markets (wheat,

peso futures) with instantaneous market clearing and "customer" markets in which economic incentives induce long-term contractual arrangements, infrequent price changes, and quantity rationing. Costly search makes customers willing to pay a premium to do business with customary suppliers. Firms, in turn, have an incentive to maintain stable prices to encourage customers to return, using yesterday's experience as a guide. "A kind of intertemporal comparison chopping" discourages firms from raising price in response to short-run increases in demand or decreases in productivity in order to avoid giving customers an incentive to begin exploring. Prices are not competely sticky, however. Widespread knowledge shared by customers and firms that costs have increased permanently allows price increases without providing an incentive for search, as was evident in the rapid response of final goods prices to the energy cost explosion of 1974.

While Gordon's (1975b) results support at least some role for changes in demand, nevertheless Okun's basic message is validated by the overwhelming share of the total variance of aggregate price inflation which is explained by changes in "standard" unit labor cost (defined for trend rather than actual productivity). Thus the search for an adequate theory of the downward inflexibility or inertia of inflation in the face of deep recessions and depressions turns to the labor market. Substantial attention has been attracted by the theory of implicit labor contracts independently developed by Costas Azariadis, Martin Baily, and Donald Gordon. Firms and workers engage in long-term contractual arrangements, which may be implicit and unwritten, and which specify wage rates in advance. Entrepreneurs are self-selected individuals who are relatively indifferent towards risk and are willing to provide insurance services for their risk-averse employees in the form of a fixed wage rate.

At present the wage contract models are incomplete and subject to criticism. Gordon (1976, p. 209) pointed out that the Azariadis-Baily-D. Gordon theory could not explain fixed-wage contracts without relying on government transfer payments paid to workers during unemployment, thus providing them with a higher total income over the cycle than they would receive if the wage varied to clear the labor market continuously. But government transfers would induce firms to respond to a recession in demand by laying off workers rather than cutting their wages even without any contractual arrangements, making the contract idea itself irrelevant. Robert Barro (1975) makes the important point that the adoption of fixed-wage contracts imposes dead-weight losses on participants by creating a divergence between the marginal product of labor and the marginal value of time. It is to the advantage of both firms and workers to maintain employment at its market-clearing level to maximize the total available product pie.

Ongoing theoretical work attempts to "rescue" the fixed-wage contract from these and other criticisms. Herschel Grossman has analyzed the attempt by firms to minimize the "default risk" of workers jumping from the fixed-wage labor contract into the auction part of the labor market when demand is high. Fruitful ideas introduced by various authors include the preference by firms for the relative certainty of the cost reduction achieved by layoffs compared to the uncertainty of the worker's response to a wage cut, and, perhaps most important, the role of employer profits made on the specific human capital of experienced employees, leading firms to maintain the wage rate of experienced employees, while achieving lower costs in a recession by laying off the least profitable inexperienced employees. The consensus appears to be shifting toward worker heterogeneity in the form of

differential risk of default, and differential endowments of specific human capital, as the most important elements motivating sticky wages, layoffs, and implicit contracts, and away from the completely homogeneous risk-averse workers featured in the earlier Azariadis-Baily-D. Gordon approach.

Whatever the precise details of the theory which explains wage and price inflexibility, the implications of such stickiness have been worked out in great detail by Barro and Grossman. Starting from an initial level of output (Q_0) and prices (P_0), let a decline in aggregate demand cut the "market clearing" price level (P*) at which Q_0 would be purchased. If the price level remains at P_0, firms want to produce as much as before but face a constraint on the amount which can be sold. Even if P drops below P_0, there will still be a sales constraint as long as P remains above P*. In the labor market the sales constraint forces firms to hire fewer workers than they would prefer at today's too-high sticky wages and prices. The requirement for the sales constraint to be lifted, and for firms to resume operating on their voluntary output supply and labor demand schedules, are (a) an increase in aggregate demand which raises P* back up to P, or (b) the passage of enough time to allow P to sink down to equal P*.

IV. Rational Expectations and Short-run Price Adjustment

The Application of Rational Expectations to Economic Policy (AREEP) constitutes a radical contribution to the theory of the short-run determinants of unemployment and inflation. The AREEP model begins with (6) above, often assuming $\underline{v} = 0$ to simplify the exposition, and thus has no bearing on our previous analysis of the long-run connection between p and $\hat{m}$. Equation (6)

is combined with the "Lucas supply function" (see Robert Lucas), which limits the source of output and unemployment changes to purely voluntary responses of firms and workers to deviations between actual and expected inflation:[1]

$$U = U^N + g^{-1}(p - p^e). \tag{10}$$

The supply function (10) is simply an inverted version of (8), describes the same long-run equilibrium conditions, and is implicit in expositions of the NRH by Friedman (1968) and others. While the idea of rational expectations has been fruitfully applied to the behavior of financial, primary commodity, and other "auction" markets, we argue here that AREEP goes badly astray by using (10) as a description of the conditions necessary for short-run output changes in the portion of the economy dominated by "customer" or "contract" markets and sluggish price adjustment.

Expectations are rational when the expectational error $(p - p^e)$ is unrelated to all information (I_{-1}) available when expectations were formed, including the autoregressive structure of all variables. The information set I_{-1} includes (6), which (when $v = 0$ and U is constant) implies:

$$p = \hat{m}, \text{ and } \qquad p^e = \hat{m}^e. \tag{11}$$

Substituting (11) into (10), we have:

$$U = U^N + g^{-1}(m - m^e). \tag{12}$$

Thus the monetary authority cannot influence unemployment, even in the short run, unless it acts in an unpredictable way. If it simply responds to an event by a formula known to the public in the previous period as part of the information set I_{-1}, the public will shift its expectation m^e by the exact amount of the change in m, the difference $(m - m^e)$ will be zero, and unemployment will not change.

The preceding argument, recently formalized by Thomas Sargent and Neil Wallace, requires for its validity that the price level (P) respond instantaneously to any change in the market-clearing price (P*). When P is sticky and fails to drop instantly to P*, the firm faces a sales constraint and cannot operate along its voluntary Lucas supply curve (10). The U. S. evidence in favor of sluggish price adjustment is strong. Gordon's (1975b) reduced-form regression between $\underline{p}$ and past values of $\underline{m}$ in the postwar U. S. has a mean lag of four years.[2] And Robert Hall has shown that only two percent of the quarterly variation in U. S. unemployment during 1954-74 remains unexplained in a simple two-quarter autoregression, in contrast to (10) above, in which U can differ from U^N only by the serially uncorrelated random error $(p - p^e)$.

Bennett McCallum has tried to argue that "recognition of price level stickiness does not, in and of itself, negate the Lucas-Sargent Proposition." His argument and its defects are most transparent for the extreme case of completely rigid prices in which $p = 0$ and a rational expectation $p^e = 0$ as well. The expectation error $(p - p^e)$ in (10) is zero, and thus unemployment is unaffected by any aggregate demand policy. But consider a policy which cuts nominal expenditure by half from E_0 to $.5E_0$. According to the McCallum argument, if prices are rigid the price level (P), unemployment, and output (Q) remain at their original level. If originally $E_0 = P_0Q_0$, now $E_1 = .5P_0Q_0$. Production is double the level of sales, and so an involuntary accumulation of inventories occurs and continues as long as E remains low and P remains rigid. Retention of the Lucas supply function in the face of price rigidity thus leads to the counterfactual conclusion that businessmen never cut production in response to involuntary inventory accumulation!

There is nothing wrong with the assumption of rational expectations itself, nor with its fruitful application to auction markets. But in light of widespread evidence that, except in a few scattered auction markets, prices adjust sluggishly to the market-clearing level in response to demand and supply shocks, it is hard to avoid the conclusion that for short-run analysis the Lucas supply function and with it AREEP should be relegated to the same scrap heap of discarded ideas where lie the earlier classical models of perfect market clearing laid to rest by Keynes forty years ago.

V. Cost Push, Controls, and Supply Shocks

Much attention in the popular press has been devoted to the positive correlation of inflation and unemployment during some years of the 1970's, and the alleged failure of economists to explain it. The straw man being attacked has only one arm, equation (8) of our two-equation inflation model, and lacks its other arm, equation (4). Further, inflation is necessarily negatively correlated with unemployment in (8) only when p^e is fixed. Inflation can increase while unemployment is rising, as in 1970 and early 1971, if expectations are formed adaptively and p^e is still rising in response to past realizations of p.

In contrast to equation (8), the dynamic supply schedule which plots a negative relation between p and U for given p^e and U^N, equation (4) is a dynamic demand schedule which plots a positive relation between p and U for given $\hat{y}$ and U_{-1}. Any event which shifts the supply curve up a fixed demand curve raises p and U simultaneously. We introduce the shift factor (Z) explicitly into (8):

$$p = p^e + g(U - U^N) + Z. \tag{13}$$

Z might be a cost-push pressure by unions, oil sheiks, or bauxite barons. As long as the authorities hold $\hat{y}$ constant, inflation and unemployment will increase simultaneously. The imposition of price controls may introduce a negative value of Z, which with $\hat{y}$ constant will cause inflation and unemployment to decrease simultaneously, as in the pre-election boom of 1971-72. The termination of controls raised inflation and unemployment simultaneously in 1974. Gordon (1975a) has shown in this context that crop failures or other supply shocks in general have multiplier effects which spread the loss of output into the nonfarm sector.

VI. Inertia and Policy Options

The same downward inertia of price adjustment which vitiates the conclusions of AREEP poses obstacles for policymakers. An economy inheriting a substantial fully anticipated inflation and operating at the natural unemployment rate has two problems--how to achieve price stability and how to reduce U^N to allow the creation of jobs for disadvantaged groups suffering from high unemployment rates. The direct remedy for inflation is the creation of a recession, which reduces p below p^e and allows the adaptive expectation of p^e to drift downwards. The permanent benefits of lower inflation must be weighed not only against the transitory output costs of a recession which might last for years, but against the permanent wealth loss caused by the recession-induced drop in saving.

Another remedy is the direct control of wages and prices. Price controls by themselves misallocate resources without permanently reducing inflation, because prices tend to be tied so closely to wage costs. Wage controls by themselves have proven to be politically infeasible; the present British experiment is possible only because it is structured to achieve a massive redistribution of income away from the rich. Recent proposals to "sell" wage controls include clever

tax schemes designed to offset the inevitable short-term losses of real income of workers who agree to allow their wages to be controlled.

Finally, the ongoing inflation can be accepted rather than resisted by allowing for the full indexing of financial assets, labor and product contracts, and all nominal dollar amounts (tax brackets, maxima, minima) written into private and government regulations. Preliminary research by Joanna Gray and others indicates that full indexing increases macroeconomic stability if the economy only suffers from demand shocks, but in the presence of supply shocks aggravates both inflation and recession. Thus from a social standpoint full indexing is not optimal, but as yet economists have failed to explain why private institutions have provided such an incomplete menu of indexed assets, liabilities, and contracts.

REFERENCES

C. Azariadis, "Implicit Contracts and Underemployment Equilibria," J. Polit. Econ., Dec. 1975, 83, 1183-1202.

M. N. Baily, "Wages and Employment Under Uncertain Demand," Rev. Econ. Stud., Jan. 1974, 41, 37-50.

R. J. Barro, "On Long-Term Contracting and the Phillips Curve," University of Rochester, unpublished, Dec. 1975.

_____ and H. Grossman, Money, Employment, and Inflation, Cambridge, 1976.

M. Friedman, "What Price Guideposts?" in G. P. Shultz and R. Z. Aliber, eds., Guidelines: Informal Controls and the Market Place (Chicago, 1966), 17-39 and 55-61. (a)

_____, "Interest Rates and the Demand for Money," J. Law Econ., Oct. 1966, 9. (b)

_____, "The Role of Monetary Policy," Amer. Econ. Rev., Mar. 1968, 58, 1-17.

D. F. Gordon, "A Neo-classical Theory of Keynesian Unemployment," Econ. Inquiry, Dec. 1974, 12, 431-59.

R. J. Gordon, "Alternative Responses of Policy to External Supply Shocks," Brookings Papers, 1975, 6, 183-206. (a)

_____, "The Effect of Aggregate Demand on Prices," Brookings Papers, 1975, 6, 613-62. (b)

_____, "The Demand for and Supply of Inflation," J. Law Econ., Dec. 1975, 18, 807-36. (c)

_____, "Recent Developments in the Theory of Inflation and Unemployment," J. Mon. Econ., Apr. 1976, 2, 185-219.

H. Grossman, "Risk Shifting and Reliability in Labor Markets," Scand. J. Econ, forthcoming, 1977.

J. Gray, "Wage Indexation: A Macroeconomic Approach," J. Mon. Econ., Apr. 1976, 2, 221-36.

R. E. Hall, "The Rigidity of Wages and the Persistence of Unemployment," Brookings Papers, 1975, 6, 301-35.

D. Laidler and J. M. Parkin, "Inflation: A Survey," Econ. J., Dec. 1975, 85, 741-809.

R. E. Lucas, Jr., "Expectations and the Neutrality of Money," J. Econ. Thy., Apr. 1972, 4, 103-24.

B. McCallum, " Price Level Stickiness and the Feasibility of Monetary Stabilization Policy with Potential Expectations," J. Polit. Econ., forthcoming.

A. Okun, "Potential GNP: Its Measurement and Significance," Proc. Amer. Statist. Ass., 1962, 98-116.

_____, "Inflation: Its Mechanics and Welfare Costs," Brookings Papers, 1975, 6, 351-90.

E. S. Phelps, "Phillips Curves, Expectations of Inflation, and Optimal Unemployment Over Time," Economica, August 1967, 34, 254-81.

M. W. Reder, "The Theoretical Problems of a National Wage-Price Policy," Can. J. Econ., Feb. 1948, 46-61.

T. J. Sargent and N. Wallace, "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," J. Polit. Econ., Apr. 1975, 83, 241-57.

FOOTNOTES

*Professor of Economics, Northwestern University

1. It is customary to include stochastic error terms in the structural equations (6) and (10), but no essential conclusions are changed by omitting these terms in this exposition.

2. Some AREEP theorists have pointed out another interpretation of my equation, that it represents a relation between p and m^e, with the lag distribution on m representing the adaptive formation of the expectation m^e. It is true that the long lag might represent expectation formation, not sluggish price adjustment. But then why should expectations on money take many years longer to form than expectations on inflation itself, which in interest rate regressions appears to be described by a mean lag of one year or less?