Econometric Policy Evaluation: A Critique

1 Introduction

The fact that nominal prices and wages tend to rise more rapidly at the peak of the business cycle than they do in the trough has been well recognized from the time when the cycle was first perceived as a distinct phenomenon. The inference that permanent inflation will therefore induce a permanent economic high is no doubt equally ancient, yet it is only recently that this notion has undergone the mysterious transformation from obvious fallacy to cornerstone of the theory of economic policy.

This transformation did not arise from new developments in economic theory. On the contrary, as soon as Phelps and others made the first serious attempts to rationalize the apparent trade-off in modern theoretical terms, the zero-degree homogeneity of demand and supply functions was re-discovered in this new context (as Friedman predicted it would be) and re-named the "natural rate hypothesis".[1] It arose, instead, from the younger tradition of the econometric forecasting models, and from the commitment on the part of a large fraction of economists to the use of these models for quantitative policy evaluation. These models have implied the existence of long-run unemployment-inflation trade-offs ever since the "wage-price sectors" were first incorporated and they promise to do so in the future although the "terms" of the trade-off continue to shift.[2]

Reprinted from *The Phillips Curve and Labor Markets,* vol. 1 of Carnegie-Rochester Conference Series on Public Policy, eds. Karl Brunner and Allan H. Meltzer, Amsterdam: North-Holland Publishing Company, pp. 19–46, by permission.

This clear-cut conflict between two rightly respected traditions—theoretical and econometric—caught those of us who viewed the two as harmoniously complementary quite by surprise. At first, it seemed that the conflict might be resolved by somewhat fancier econometric footwork. On the theoretical level, one hears talk of a "disequilibrium dynamics" which will somehow make money illusion respectable while going beyond the sterility of $dp/dt = k(p - p^e)$. Without underestimating the ingenuity of either econometricians or theorists, it seems to me appropriate to entertain the possibility that reconciliation along both of these lines will fail, and that one of these traditions is fundamentally in error.

The thesis of this essay is that it is the econometric tradition, or more precisely, the "theory of economic policy" based on this tradition, which is in need of major revision. More particularly, I shall argue that the features which lead to success in short-term forecasting are unrelated to quantitative policy evaluation, that the major econometric models are (well) designed to perform the former task only, and that simulations using these models can, in principle, provide *no* useful information as to the actual consequences of alternative economic policies. These contentions will be based not on deviations between estimated and "true" structure prior to a policy change but on the deviations between the prior "true" structure and the "true" structure prevailing afterwards.

Before turning to details, I should like to advance two disclaimers. First, as is true with any technically difficult and novel area of science, econometric model building is subject to a great deal of ill-informed and casual criticism. Thus models are condemned as being "too big" (with equal insight, I suppose one could fault smaller models for being "too little"), too messy, too simplistic (that is, not messy enough), and, the ultimate blow, inferior to "naive" models. Surely the increasing sophistication of the "naive" alternatives to the major forecasting models is the highest of tributes to the remarkable success of the latter. I hope I can succeed in disassociating the criticism which follows from any denial of the very important advances in forecasting ability recorded by the econometric models, and of the promise they offer for advancement of comparable importance in the future.

One may well define a critique as a paper which does not fully engage the vanity of its author. In this spirit, let me offer a second disclaimer. There is little in this essay which is not implicit (and

perhaps to more discerning readers, explicit) in Friedman [11], Muth [29] and, still earlier, in Knight [21]. For that matter, the criticisms I shall raise against currently popular applications of econometric theory have, for the most part, been anticipated by the major original contributors to that theory.[3] Nevertheless, the case for sustained inflation, based entirely on econometric simulations, is attended now with a seriousness it has not commanded for many decades. It may, therefore, be worthwhile to attempt to trace this case back to its foundation, and then to examine again the scientific basis of this foundation itself.

2 The Theory of Economic Policy

Virtually all quantitative macro-economic policy discussions today are conducted within a theoretical framework which I shall call "the theory of economic policy" (following Tinbergen [35]). The essentials of this framework are so widely known and subscribed to that it may be superfluous to devote space to their review. On the other hand, since the main theme of this paper is the inadequacy of this framework, it is probably best to have an explicit version before us.

One describes the economy in a time period t by a vector y_t of state variables, a vector x_t of exogeneous forcing variables, and a vector ϵ_t of independent (through time), identically distributed random shocks. The motion of the economy is determined by a difference equation

$$y_{t+1} = f(y_t, x_t, \epsilon_t),$$

the distribution of ϵ_t, and a description of the temporal behavior of the forcing variables, x_t. The function f is taken to be fixed but not directly known; the task of empiricists is then to estimate f. For practical purposes, one usually thinks of estimating the values of a fixed parameter vector θ, with

$$f(y,x,\epsilon) \equiv F(y,x,\theta,\epsilon)$$

and F being specified in advance.

Mathematically, the sequence $\{x_t\}$ of forcing vectors is regarded as being "arbitrary" (that is, it is not characterized stochastically). Since the past x_t values are observed, this causes no difficulty in estimating θ, and in fact simplifies the theoretical estimation problem slightly. For forecasting, one is obliged to insert forecasted x_t values into F.

With knowledge of the function F and θ, policy evaluation is a straightforward matter. A *policy* is viewed as a specification of present and future values of some components of $\{x_t\}$. With the other components somehow specified, the stochastic behavior of $\{y_t, x_t, \epsilon_t\}$ from the present on is specified, and functionals defined on this sequence are well-defined random variables, whose moments may be calculated theoretically or obtained by numerical simulation. Sometimes, for example, one wishes to examine the mean value of a hypothetical "social objective function," such as

$$\sum_{t=0}^{\infty} \beta^t u(y_t, x_t, \epsilon_t)$$

under alternative policies. More usually, one is interested in the "operating characteristics" of the system under alternative policies. Thus, in this standard context, a "long-run Phillips curve" is simply a plot of average inflation–unemployment pairs under a range of hypothetical policies.[4]

Since one cannot treat θ as known in practice, the actual problem of policy evaluation is somewhat more complicated. The fact that θ is estimated from past sample values affects the above moment calculations for small samples; it also makes policies which promise to sharpen estimates of θ relatively more attractive. These considerations complicate without, I think, essentially altering the theory of economic policy as sketched above.

Two features of this theoretical framework deserve special comment. The first is the uneasy relationship between this theory of economic policy and traditional economic theory. The components of the vector-valued function F are behavioral relationships–demand functions; the role of theory may thus be viewed as suggesting forms for F, or in Samuelson's terms, distributing zeros throughout the Jacobian of F. This role for theory is decidedly secondary: microeconomics shows surprising power to rationalize individual econometric relationships in a variety of ways. More significantly, this micro-economic role for theory abdicates the task of describing the aggregate behavior of the system entirely to the econometrician. Theorists suggest forms for consumption, investment, price and wage setting functions separately; these suggestions, if useful, influence individual components of F. The aggregate behavior of the system then is what it is.[5] Surely this point of view (though I doubt if many would now endorse it in so bald a form) accounts for the demise of traditional "business cycle theory" and the widespread acceptance of a Phillips "trade-off" in

the absence of *any* aggregative theoretical model embodying such a relationship.

Secondly, one must emphasize the intimate link between short-term forecasting and long-term simulations within this standard framework. The variance of short-term forecasts tend to zero with the variance of ϵ_t; as the latter becomes small, so also does the variance of estimated behavior of $\{y_t\}$ conditional on hypothetical policies $\{x_t\}$. Thus forecasting accuracy in the short-run implies reliability of long-term policy evaluation.

3 Adaptive Forecasting

There are many signs that practicing econometricians pay little more than lip-service to the theory outlined in the preceding section. The most striking is the indifference of econometric forecasters to data series prior to 1947. Within the theory of economic policy, more observations *always* sharpen parameter estimates and forecasts, and observations on "extreme" x_t values particularly so; yet even the readily available annual series from 1929–1946 are rarely used as a check on the post-war fits.

A second sign is the frequent and frequently important refitting of econometric relationships. The revisions of the wage-price sector now in progress are a good example.[6] The continuously improving precision of the estimates of θ within the fixed structure F, predicted by the theory, does not seem to be occurring in practice.

Finally, and most suggestively, is the practice of using patterns in recent residuals to revise intercept estimates for forecasting purposes. For example, if a "run" of positive residuals (predicted less actual) arises in an equation in recent periods, one revises the estimated intercept downward by their average amount. This practice accounts, for example, for the superiority of the actual Wharton forecasts as compared to forecasts based on the published version of the model.[7]

It should be emphasized that recounting these discrepancies between theory and practice is not to be taken as criticism of econometric forecasters. Certainly if new observations are better accounted for by new or modified equations, it would be foolish to continue to forecast using the old relationships. The point is simply that, econometrics textbooks notwithstanding, current forecasting practice is *not* conducted within the framework of the theory of economic policy, and the unquestioned success of the forecasters should *not* be construed

as evidence for the soundness or reliability of the structure proposed in that theory.

An alternative structure to that underlying the theory of economic policy has recently been proposed (in [3] and [5]) by Cooley and Prescott. The structure is of interest in the present context, since optimal forecasting within it shares many features with current forecasting practice as just described. Instead of treating the parameter vector θ as fixed, Cooley and Prescott view it as a random variable following the random walk

$$\theta_{t+1} = \theta_t + \eta_{t+1},$$

where $\{\eta_t\}$ is a sequence of independent, identically distributed random variables.

Maximum likelihood forecasting under this alternative framework ("adaptive regression") resembles "exponential smoothing" on the observations, with observations in the distant past receiving a small "weight"—very much as in usual econometric practice; similarly, recent forecast errors are used to adjust the estimates. Using both artificial data and economic time series, Cooley and Prescott have shown (in [4]) that adaptive methods have good short-term forecasting properties when compared to even relatively sophisticated versions of the "fixed θ" regression model. As Klein and others have remarked, this advantage is shared by actual large-model forecasts (that is, model forecasts modified by the forecaster's judgment) over mechanical forecasts using the published versions of the model.[8]

Cooley and Prescott have proposed adaptive regression as a normative forecasting method. I am using it here in a positive sense: as an idealized "model" of the behavior of large-model forecasters. If the model is, as I believe, roughly accurate, it serves to reconcile the assertion that long-term policy evaluations based on econometric models are meaningless with the acknowledgment that the forecast accuracy of these models is good and likely to become even better. Under the adaptive structure, a small standard error of short-term forecasts is consistent with *infinite* variance of the long-term operating characteristics of the system.

4 Theoretical Considerations: General

To this point, I have argued simply that the standard, stable-parameter view of econometric theory and quantitative policy evaluation appears

not to match several important characteristics of econometric practice, while an alternative general structure, embodying stochastic parameter drift, matches these characteristics very closely. This argument is, if accepted, sufficient to establish that the "long-run" implications of current forecasting models are without content, and that the short-term forecasting ability of these models provides no evidence of the accuracy to be expected from simulations of hypothetical policy rules.

These points are, I think, important, but their implications for the future are unclear. After all, the major econometric models are still in their first, highly successful, decade. No one, surely, expected the initial parameterizations of these models to stand forever, even under the most optimistic view of the stability of the unknown, underlying structure. Perhaps the adaptive character of this early stage of macroeconomic forecasting is merely the initial groping for the true structure which, however ignored in statistical theory, all practitioners knew to be necessary. If so, the arguments of this paper are transitory debating points, obsolete soon after they are written down. Personally, I would not be sorry if this were the case, but I do not believe it is. I shall try to explain why, beginning with generalities, and then, in the following section, introducing examples.

In section 2, we discussed an economy characterized by

$$y_{t+1} = F(y_t, x_t, \theta, \epsilon_t).$$

The function F and parameter vector θ are derived from decision rules (demand and supply functions) of agents in the economy, and these decisions are, theoretically, optimal given the situation in which each agent is placed. There is, as remarked above, no presumption that (F,θ) will be easy to discover, but it *is* the central assumption of the theory of economic policy that once they *are* (approximately) known, they will remain stable under arbitrary changes in the behavior of the forcing sequence $\{x_t\}$.

For example, suppose a reliable model (F,θ) is in hand, and one wishes to use it to assess the consequences of alternative monetary and fiscal policy rules (choices of $x_0, x_1, x_2, \ldots$, where $t = 0$ is "now"). According to the theory of economic policy, one then simulates the system under alternative policies (theoretically or numerically) and compares outcomes by some criterion. For such comparisons to have any meaning, it is essential that the structure (F,θ) not vary systematically with the choice of $\{x_t\}$.

Everything we know about dynamic economic theory indicates that this presumption is unjustified. First, the individual decision problem: "find an optimal decision rule when certain parameters (future prices, say) follow 'arbitrary' paths" is simply not well formulated. Only trivial problems in which agents can safely ignore the future can be formulated under such a vague description of market constraints. Even to obtain the decision rules underlying (F,θ) then, we have to attribute to individuals some view of the behavior of the future values of variables of concern to them. This view, in conjunction with other factors, determines their optimum decision rules. To assume stability of (F,θ) under alternative policy rules is thus to assume that agents' views about the behavior of shocks to the system are invariant under changes in the true behavior of these shocks. Without this extreme assumption, the kinds of policy simulations called for by the theory of economic policy are meaningless.

It is likely that the "drift" in θ which the adaptive models describe stochastically reflects, in part, the adaptation of the decision rules of agents to the changing character of the series they are trying to forecast.[9] Since this adaptation will be in most (though not all) cases slow, one is not surprised that adaptive methods can improve the short-term forecasting abilities of the econometric models. For longer term forecasting and policy simulations, however, ignoring the systematic sources of drift will lead to large, unpredictable errors.

5 Theoretical Considerations: Examples

If these general theoretical observations on the likelihood of systematic "parametric drift" in the face of variations in the structure of shocks are correct, it should be possible to confirm them by examination of the specific decision problems underlying the major components of aggregative models. I shall discuss in turn consumption, investment, and the wage-price sector, or Phillips curve. In each case, the "right-hand variables" will, for simplicity, be taken as "exogenous" (as components of $\{x_t\}$). The thought-experiments matching this assumption, and the adaptations necessary for simultaneous equations, are too well known to require comment.

5.1 Consumption

The easiest example to discuss with confidence is the aggregate consumption function since, due to Friedman [11], Muth [28] and Modi-

gliani, Brumberg and Ando [2], [27], it has both a sound theoretical rationale and an unusually high degree of empirical success. Adopting Friedman's formulation, permanent consumption is proportional to permanent income (an estimate of a discounted future income stream),

$$c_{pt} = ky_{pt}; \tag{1}$$

actual consumption is

$$c_t = c_{pt} + u_t; \tag{2}$$

and actual, current income is

$$y_t = y_{pt} + v_t. \tag{3}$$

The variables u_t,v_t are independent temporally and of each other and of y_{pt}.

An empirical "short-run" marginal propensity to consume is the sample moment corresponding to $\text{Cov}(c_t,y_t)/\text{Var}(y_t)$, or

$$k\,\frac{\text{Var}(y_{pt})}{\text{Var}(y_{pt}) + \text{Var}(v_t)}.$$

Now as long as these moments are viewed as subjective parameters in the heads of consumers, this model lacks content. Friedman, however, viewed them as true moments, known to consumers, the logical step which led to the cross-sectional tests which provided the most striking confirmation of his permanent income hypothesis.[10]

This central equating of a true probability distribution and the subjective distribution on which decisions are based was termed rational expectations by Muth, who developed its implications more generally (in [29]). In particular, in [28], Muth found the stochastic behavior of income over time under which Friedman's identification of permanent income as an exponentially weighted sum of current and lagged observations on actual income was consistent with optimal forecasting on the part of agents.[11]

To review Muth's results, we begin by recalling that permanent income is that constant flow y_{pt} which has the same value, with the subjective discount factor β, as the forecasted actual income stream:

$$y_{pt} = (1 - \beta)\sum_{i=0}^{\infty} \beta^i E(y_{t+i}\,|\,I_t) \tag{4}$$

where each expectation is conditioned on information I_t available at t.

Now let actual income y_t be a sum of three terms

$$y_t = a + w_t + v_t, \tag{5}$$

where v_t is transitory income, a is a constant, and w_t is a sum of independent *increments*, each with zero mean and constant variance. Muth showed that the minimum variance estimator of y_{t+i} for all $i = 1,2, \ldots$ is $(1 - \lambda)\,\Sigma_{j=0}^{\infty}\lambda^j y_{t-j}$ where λ depends in a known way on the relative variances of w_t and v_t.[12] Inserting this estimator into (4) and summing the series gives the empirical consumption function

$$c_t = k(1 - \beta)y_t + k\beta(1 - \lambda)\sum_{j=0}^{\infty}\lambda^j y_{t-j} + u_t. \tag{6}$$

(This formula differs slightly from Muth's because Muth implicitly assumed that c_t was determined prior to realizing y_t. The difference is not important in the sequel.)

Now let us imagine a consumer of this type, with a current income generated by an "experimenter" according to the pattern described by Muth (so that the premises of the theory of economic policy are correct for a single equation consumption function). An econometrician observing this consumer over many periods will have good success describing him by (6) whether he arrives at this equation by the Friedman-Muth reasoning, or simply hits on it by trial-and-error. Next consider *policies* taking the form of a sequence of supplements $\{x_t\}$ to this consumer's income from time T on. Whether $\{x_t\}$ is specified deterministically or by some stochastic law, whether it is announced in advance to the consumer or not, the theory of economic policy prescribes the *same* method for evaluating its consequences: add x_t to the forecasts of y_t for each $t > T$, insert into (6), and obtain the new forecasts of c_t.

If the consumer knows of the policy change in advance, it is clear that this standard method gives incorrect forecasts. For example, suppose the policy consists of a constant increase, $x_t = \bar{x}$, in income over the entire future. From (4), this leads to an increase in consumption of $k\bar{x}$. The forecast based on (6), however, is of an effect in period t of

$$(\Delta c)_t = k\bar{x}\{(1 - \beta) + \beta(1 - \lambda)\sum_{i=0}^{t-T}\lambda^i\}.$$

Since this effect tends to the correct forecast, $k\bar{x}$, as t tends to infinity, one might conjecture that the difficulty vanishes in the "long run." To see that this conjecture is false, consider an exponentially growing supplement $x_t = \bar{x}a^t$, $1 < a < 1/\beta$. The true effect in $t - T$ is, from (1) and (4),

$$(\Delta c)_t = k\bar{x}\frac{(1-\beta)a^t}{1-a\beta}$$

The effect as forecast by (6) is

$$(\Delta c)_t = k\bar{x}\left\{(1-\beta) + \beta(1-\lambda)\sum_{j=0}^{t-T}\left(\frac{\lambda}{a}\right)^j\right\}a^t.$$

Neither effect tends to zero, as t tends to infinity; the ratio (forecast over actual) tends to

$$(1-a\beta)\left\{1 + \frac{a\beta(1-\lambda)}{(1-\beta)(a-\lambda)}\right\},$$

which may lie on either side of unity.

More interesting divergences between forecasts and reality emerge when the policy is stochastic, but with characteristics known in advance. For example, let $\{x_t\}$ be a sequence of independent random variables, with zero mean and constant variance, distributed independently of u_t, v_t and w_t. This policy amounts to an increase in the variance of transitory income, lowering the weight λ in a manner given by the Muth formula. Average consumption, in fact and as forecast by (6), is not affected, but the variance of consumption is. The correct estimate of this variance effect requires revision of the weight λ; evidently the standard, fixed-parameter prediction based on (6) will again yield the wrong answer, and the error will not tend to vanish for large t.

The list of deterministic and stochastic policy changes, and their combination is inexhaustible but one need not proceed further to establish the point: for *any* policy change which is understood in advance, extrapolation or simulation based on (6) yields an incorrect forecast, and what is more, a *correctibly* incorrect forecast. What of changes in policy which are *not* understood in advance? As Fisher observes, "the notion that one cannot fool all of the people all of the time [need not] imply that one cannot fool all the people even some of the time."[13]

The observation is, if obvious, true enough; but it provides no support whatever for the standard forecasting method of extrapolating on the basis of (6). Our knowledge of consumption behavior is summarized in (1)–(4). For certain policy changes we can, with some confidence, guess at the permanent income recalculations consumers will go through and hope to predict their consumption responses with some accuracy. For other types of policies, particularly those involving deliberate "fooling" of consumers, it will not be at all clear how to apply (1)–(4), and hence impossible to forecast. Obviously, in such cases, there is no reason to imagine that forecasting with (6) will be accurate either.

5.2 Taxation and Investment Demand

In [15], Hall and Jorgenson provided quantitative estimates of the consequences, current and lagged, of various tax policies on the demand for producers' durable equipment. Their work is an example of the current state of the art of conditional forecasting at its best. The general method is to use econometric estimates of a Jorgensonian investment function, which captures all of the relevant tax structure in a single implicit rental price variable, to simulate the effects of alternative tax policies.

An implicit assumption in this work is that any tax change is regarded as a permanent, once-and-for-all change. Insofar as this assumption is false over the sample period, the econometric estimates are subject to bias.[14] More important for this discussion, the conditional forecasts will be valid *only* for tax changes *believed* to be permanent by taxpaying corporations.

For many issues in public finance, this obvious qualification would properly be regarded as a mere technicality. For Keynesian countercyclical policy, however, it is the very heart of the issue. The whole point, after all, of the investment tax credit is that it be viewed as temporary, so that it can serve as an inducement to firms to reschedule their investment projects. It should be clear that the forecasting methods used by Hall and Jorgenson (and, of course, by other econometricians) cannot be expected to yield even order-of-magnitude estimates of the effects of explicitly temporary tax adjustments.

To pursue this issue further, it will be useful to begin with an explicit version of the standard accelerator model of investment behavior. We imagine a constant returns industry in which each firm has a constant

output-capital ratio λ. Using a common notation for variables at both the firm and industry level, let k_t denote capital at the beginning of year t. Output during t is λk_t. Investment during the year, i_t, affects next period's capital according to

$$k_{t+1} = i_t + (1 - \delta)k_t,$$

where δ is a constant physical rate of depreciation. Output is sold on a perfect market at a price p_t; investment goods are purchased at a constant price of unity. Profits (sales less depreciation) are taxed at the rate θ_t; there is an investment tax credit at the rate Ψ_t.

The firm is interested in maximizing the expected present value of receipts net of taxes, discounted at the constant cost of capital r. In the absence (assumed here) of adjustment costs, this involves equating the current cost of an additional unit of investment to the expected discounted net return. Assuming that the current tax bill is always large enough to cover the credit, the current *cost* of acquiring an additional unit of capital is $(1 - \Psi_t)$, independent of the volume of investment goods purchased. Each unit of investment yields λ units of output, to be sold next period at the (unknown) price p_{t+1}. Offsetting this profit is a tax bill of $\theta_{t+1}[\lambda p_{t+1} - \delta]$. In addition, $(1 - \delta)$ units of the investment good remain for use after period $t + 1$; with perfect capital goods markets, these units are valued at $(1 - \Psi_{t+1})$. Thus letting $E_t(\cdot)$ denote an expectation conditional on information up to period t, the expected discounted return per unit of investment in t is

$$\frac{1}{1 + r}E_t[\lambda p_{t+1}(1 - \theta_{t+1}) + \delta\theta_{t+1} + (1 - \delta)(1 - \Psi_{t+1})].$$

Since a change in next period's tax rate θ_{t+1} which is not anticipated in t is a "pure profit tax," θ_{t+1} and p_{t+1} will be uncorrelated. Hence, equating costs and returns, one equilibrium condition for the industry is

$$1 - \Psi_t = \frac{1}{1 + r}\{\lambda E_t(p_{t+1})[1 - E_t(\theta_{t+1})] + \delta E_t(\theta_{t+1}) + (1 - \delta)[1 - E_t(\Psi_{t+1})]\}. \tag{7}$$

A second equilibrium condition is obtained from the assumption that the product market is cleared each period. Let industry demand be given by a linear function, with a stochastically shifting intercept a_t and a constant slope b, so that quantity demanded next period will be

$a_{t+1} - bp_{t+1}$. Quantity supplied will be λ times next period's capital. Then a second equilibrium condition is

$$\lambda[i_t + (1 - \delta)k_t] = a_{t+1} - bp_{t+1}.$$

Taking mean values of both sides,

$$\lambda[i_t + (1 - \delta)k_t] = E_t(a_{t+1}) - bE_t(p_{t+1}). \tag{8}$$

Since our interest is in the industry investment function, we eliminate $E_t(p_{t+1})$ between (7) and (8) to obtain:

$$i_t + (1 - \delta)k_t = \frac{1}{\lambda}E_t(a_{t+1}) - \frac{b}{\lambda^2}\left[\frac{r}{1 - E_t(\theta_{t+1})} + \delta\right] + \frac{b}{\lambda^2}\left[\frac{(1 + r)\Psi_t - (1 - \delta)E_t(\Psi_{t+1})}{1 - E_t(\theta_{t+1})}\right]. \tag{9}$$

Equation (9) gives the industry's "desired" stock of capital, $i_t + (1 - \delta)k_t$, as a function of the expected future state of demand and the current and expected future tax structure, as well as of the cost of capital r, taken in this illustration to be constant. The second and third terms on the right are the product of the slope of the demand curve for capital, $-b\lambda^{-2}$, and the familiar Jorgensonian implicit rental price; the second term includes "interest" and depreciation costs, net of taxes; the third includes the expected capital gain (or loss) due to changes in the investment tax credit rate.

In most empirical investment studies, firms are assumed to move gradually from k_t to the desired stock given by (9), due to costs of adjustment, delivery lags, and the like. We assume here, purely for convenience, that the full adjustment occurs in a single period.

Equation (9) is operationally at the same level as equations (1) and (4) of the preceding section: it relates current behavior to unobserved expectations of future variables. To move to a testable hypothesis, one must specify the time series behavior of a_t, θ_t and Ψ_t (as was done for income in consumption theory), obtain the optimal forecasting rule, and obtain the analogue to the consumption function (6). Let us imagine that this has been accomplished, and estimates of the parameters λ and b have been obtained. How would one use these estimates to evaluate the consequences of a particular investment tax credit policy?

The method used by Hall and Jorgenson is to treat the credit as a permanent or once-and-for-all change, or implicitly to set $E_t(\Psi_{t+1})$ equal to Ψ_t. Holding θ_t constant at θ, the effect of a change in the

credit from 0 to Ψ (say) would be the same as a permanent lowering of the price of investment goods to $1 - \Psi$ or, from (9), an increase in the desired capital stock of $(b/\lambda^2)(r + \delta)/(1 - \theta)$. If the credit is in fact believed by corporations to be permanent, this forecast will be correct; otherwise it will not be.

To consider alternatives, imagine a stochastic tax credit policy which switches from 0 to a fixed number Ψ in a Markovian fashion, with transitions given by $\Pr\{\Psi_{t+1} = \Psi \mid \Psi_t = 0\} = q$ and $\Pr\{\Psi_{t+1} = \Psi \mid \Psi_t = \Psi\} = p$.[15] Then if expectations on next period's tax credit are formed rationally, conditional on the presence or absence of the credit in the current period, we have

$$E_t(\Psi_{t+1}) = \begin{cases} q\Psi & \text{if } \Psi_t = 0 \\ p\Psi & \text{if } \Psi_t = \Psi. \end{cases}$$

The third term on the right of (9) is then

$$\frac{b\Psi}{\lambda^2(1-\theta)}[-q(1-\delta)] \qquad \text{if } \Psi_t = 0,$$

$$\frac{b\Psi}{\lambda^2(1-\theta)}[1 + r - p(1-\delta)] \qquad \text{if } \Psi_t = \Psi.$$

The difference between these terms is given by the expression

$$\frac{b\Psi}{\lambda^2(1-\theta)}[1 + r + (q - p)(1 - \delta)]. \tag{10}$$

The expression (10) gives the increment to desired capital stock (and, with immediate adjustment, to current investment) when the tax credit is switched from zero to Ψ *in an economy where the credit operates, and is known to operate, in the stochastic fashion described above*. It does *not* measure the effect of a switch in policy from a no-credit regime to the stochastic regime used here. (The difference arises because even when the credit is set at zero in the stochastic regime, the possibility of capital loss, due to the introduction of the credit in the future, increases the implicit rental on capital, relative to the situation in which the credit is expected to remain at zero forever.)

By examining extreme values of p and q one can get a good idea of the quantitative importance of expectations in measuring the effect of the credit. At one extreme, consider the case where the credit is expected almost *never* to be offered (q near 0), but once offered, it is

permanent (p near 1). The effect of a switch from 0 to Ψ is, in this case, approximately

$$\frac{b\Psi}{\lambda^2(1-\theta)}[r + \delta],$$

using (10). This is the situation assumed, implicitly, by Hall and Jorgenson. At the other extreme, consider the case of a frequently imposed but always transitory credit (q near 1, p near 0). Applying (10), the effect of a switch in this case is approximately

$$\frac{b\Psi}{\lambda^2(1-\theta)}[2 + r - \delta].$$

The ratio of effects is then $(2 + r - \delta)/(r + \delta)$. With $r = .14$ and $\delta = .15$, this ratio is about 7.[16] We are not, then, discussing a quantitatively minor issue.

For a more realistic estimate, consider a credit which remains "off" for an average period of 5 years, and when "switched on" remains for an average of one year. These assumptions correspond to setting $p \approx 0$ and $q = \frac{1}{5}$. The ratio of the effect (from (10)), under these assumptions versus those used by Hall and Jorgenson is now $[1 + r + \frac{1}{5}(1 - \delta)]/(r + \delta)$. With $r = .14$ and $\delta = .15$, this ratio is approximately 4.5. This ratio would probably be somewhat smaller under a more satisfactory lag structure[17], but even taking this into account, it appears likely that the potential stimulus of the investment tax credit may well be several times greater than the Hall-Jorgenson estimates would indicate.[18]

As was the case in the discussion of consumption behavior, estimation of a policy effect along the above lines presupposes a policy generated by a fixed, relatively simple *rule*, known by forecasters (ourselves) and by the agents subject to the policy (an assumption which is not only convenient analytically but consistent with Article 1, Section 7 of the U.S. Constitution). To go beyond the kind of order-of-magnitude calculations used here to an accurate assessment of the effects of the 1962 credit studied by Hall and Jorgenson, one would have to infer the implicit rule which generated (or was thought by corporations to generate) that policy, a task made difficult, or perhaps impossible, by the novelty of the policy at the time it was introduced. Similarly, there is no reason to hope that we can accurately forecast the effects of future *ad hoc* tax policies on investment behavior. On

the other hand, there is every reason to believe that good quantitative assessments of counter-cyclical fiscal rules, which are built into the tax structure in a stable and well-understood way, can be obtained.

5.3 Phillips Curves

A third example is suggested by the recent controversy over the Phelps-Friedman hypothesis that permanent changes in the inflation rate will not alter the average rate of unemployment. Most of the major econometric models have been used in simulation experiments to test this proposition; the results are uniformly negative. Since expectations are involved in an essential way in labor and product market supply behavior, one would presume, on the basis of the considerations raised in section 4, that these tests are beside the point.[19] This presumption is correct, as the following example illustrates.

It will be helpful to utilize a simple, parametric model which captures the main features of the expectational view of aggregate supply—rational agents, cleared markets, incomplete information.[20] We imagine suppliers of goods to be distributed over N distinct markets i, $i = 1, \ldots, N$. To avoid index number problems, suppose that the same (except for location) good is traded in each market, and let y_{it} be the log of quantity supplied in market i in period t. Assume, further, that the supply y_{it} is composed of two factors

$$y_{it} = y_{it}^{p} + y_{it}^{c}.$$

where y_{it}^{p} denotes normal or permanent supply, and y_{it}^{c} cyclical or transitory supply (both, again, in logs). We take y_{it}^{p} to be unresponsive to all but permanent relative price changes or, since the latter have been defined away by assuming a single good, simply unresponsive to price changes. Transitory supply y_{it}^{c} varies with perceived changes in the *relative* price of goods in i:

$$y_{it}^{c} = \beta(p_{it} - p_{it}^{e}),$$

where p_{it} is the log of the actual price in i at t, and p_{it}^{e} is the log of the general (geometric average) price level in the economy as a whole, *as perceived in market i.*[21]

Prices will vary from market to market for each t, due to the usual sources of fluctuation in relative demands. They will also fluctuate over time, due to movements in aggregate demand. We shall not explore the sources of these price movements (although this is easy

enough to do) but simply postulate that the actual price in i at t consists of two components:

$$p_{it} = p_t + z_{it}.$$

Sellers observe the actual price p_{it}; the two components cannot be separately observed. The component p_t varies with time, but is common to all markets. Based on information obtained prior to t (call it I_{t-1}) traders in all markets take p_t to be a normally distributed random variable, with mean $\bar{p}_t$ (reflecting this past information) and variance σ^2. The component z_{it} reflects relative price variation across markets and time: z_{it} is normally distributed, independent of p_t and z_{js} (unless $i = j$, $s = t$), with mean 0 and variance τ^2.

The *actual* general price level at t is the average over markets of individual prices,

$$\frac{1}{N}\sum_{i=1}^{N} p_{it} = p_t + \frac{1}{N}\sum_{i=1}^{N} z_{it}.$$

We take the number of markets N to be large, so that the second term can be neglected, and p_t is the general price level. To form the supply decision, suppliers estimate p_t; assume that this estimate p_{it}^{e} is the mean of the true conditional distribution of p_t. The latter is calculated using the observation that p_{it} is the sum of two independent normal variates, one with mean 0 and variance τ^2; one with mean $\bar{p}_t$ and variance σ^2. It follows that

$$p_{it}^{e} = E\{p_t | p_{it}, I_{t-1}\} = (1 - \theta)p_{it} + \theta\bar{p}_t,$$

where $\theta = \tau^2/(\sigma^2 + \tau^2)$.

Based on this unbiased but generally inaccurate estimate of the current general level of prices, suppliers in i follow

$$y_{it}^{c} = \beta[p_{it} - ((1 - \theta)p_{it} + \theta\bar{p}_t)] = \theta\beta[p_{it} - \bar{p}_t].$$

Now averaging over markets, and invoking the law of large numbers again, we have the cyclical component of *aggregate* supply:

$$y_t^{c} = \theta\beta(p_t - \bar{p}_t).$$

Re-introducing the permanent components,

$$y_t = \theta\beta(p_t - \bar{p}_t) + y_{pt}. \tag{11}$$

Though simple, (11) captures the main features of the expectational or "natural rate" view of aggregate supply. The supply of goods is viewed as following a trend path y_{pt} which is not dependent on nominal price movements. Deviations from this path are induced whenever the nominal price deviates from the level which was expected to prevail on the basis of past information. These deviations occur because agents are obliged to infer current general price movements on the basis of incomplete information.

It is worth speculating as to the sort of empirical performance one would *expect* from (11). In doing so, we ignore the trend component y_{pt}, concentrating on the determinants of p_t, β and θ. The parameter β reflects intertemporal substitution possibilities in supply: technological factors such as storability of production, and tastes for substituting labor supplied today for supply tomorrow. One would expect β to be reasonably stable over time and across economies at a similar level of development. The parameter θ is the ratio $\tau^2/(\sigma^2 + \tau^2)$. τ^2 reflects the variability of relative prices within the economy; there is no reason to expect it to vary systematically with demand policy. σ^2 is the variance of the general price level about its expected level; it will obviously increase with increases in the volatility of demand.[22] Similarly, $\bar{p}_t$, the expected price level conditional on past information, will vary with actual, average inflation rates.

Turning to a specific example, suppose that actual prices follow the random walk

$$p_t = p_{t-1} + \epsilon_t \tag{12}$$

where ϵ_t is normal with mean π and variance σ^2. Then $p_t = p_{t-1} + \pi$ and (11) becomes

$$y_t = \theta\beta(p_t - p_{t-1}) - \theta\beta\pi + y_{pt}. \tag{13}$$

Over a sample period during which π and σ^2 remain roughly constant, and if y_{pt} can be effectively controlled for, (13) will appear to the econometrician to describe a stable trade-off between inflation and real output. The addition of lagged inflation rates will not improve the fit, or alter this conclusion in any way. Yet is is evident from (13) that a sustained increase in the inflation rate (an increase in π) will not affect real output.

This is not to say that a distributed lag version of (11) might not perform better empirically. Thus let the actual rate of inflation follow

a first-order autoregressive scheme

$$\Delta p_t = \rho\,\Delta p_{t-1} + \epsilon_t$$

or

$$p_t = (1 + \rho)p_{t-1} - \rho p_{t-2} + \epsilon_t \tag{14}$$

where $0 < \rho < 1$ and ϵ_t is distributed as before.

Then combining (11) and (14):

$$y_t = \theta\beta\,\Delta p_t - \theta\beta\rho\,\Delta p_{t-1} - \theta\beta\pi + y_{pt}. \tag{15}$$

In econometric terms, the "long-run" slope, or trade-off, would be the *sum* of the inflation coefficients, or $\theta\beta(1 - \rho)$, which will not, if (14) is stable, be zero.

In short, one can imagine situations in which empirical Phillips curves exhibit long lags and situations in which there are no lagged effects. In either case, the "long-run" output-inflation relationship as calculated or simulated in the conventional way has *no* bearing on the actual consequences of pursuing a policy of inflation.

As in the consumption and investment examples, the ability to use (13) or (15) to forecast the consequences of a change in policy rests crucially on the assumption that the parameters describing the new policy (in this case π, σ^2 and ρ) are known by agents. Over periods for which this assumption is not approximately valid (obviously there have been, and will continue to be, many such periods) empirical Phillips curves will appear subject to "parameter drift," describable over the sample period, but unpredictable for all but the very near future.

6 Policy Considerations

In preceding sections, I have argued in general and by example that there are compelling empirical and theoretical reasons for believing that a structure of the form

$$y_{t+1} = F(y_t, x_t, \theta, \epsilon_t)$$

(F known, θ fixed, x_t "arbitrary") will not be of use for forecasting and policy evaluation in actual economies. For short-term forecasting, these arguments have long been anticipated in practice, and models with good (and improvable) tracking properties have been obtained by permitting and measuring "drift" in the parameter vector θ. Under

adaptive models which rationalize these tracking procedures, however, long-run policy simulations are acknowledged to have infinite variance, which leaves open the question of quantitative policy evaluation.

One response to this situation, seldom defended explicitly today though in implicit form probably dominant at the most "practical" level of economic advice-giving, is simply to dismiss questions of the long-term behavior of the economy under alternative policies and focus instead on obtaining what is viewed as desirable behavior in the next few quarters. The hope is that the changes in θ induced by policy changes will occur slowly, and that conditional forecasting based on tracking models will therefore be roughly accurate for a few periods. This hope is both false and misleading. First, some policy changes induce immediate jumps in θ: for example, an explicitly temporary personal income tax surcharge will (cf. section 5.1) induce an *immediate* rise in propensity to consume out of disposable income and consequent errors in short-term conditional forecasts.[23] Second, even if the induced changes in θ are slow to occur, they should be counted in the short-term "objective function", yet rarely are. Thus econometric Phillips curves roughly forecast the initial phase of the current inflation, but not the "adverse" shift in the curve to which that inflation led.

What kind of structure might be at once consistent with the theoretical considerations raised in section 4 and with operational, accurate policy evaluation? One hesitates to indulge the common illusion that "general" structures are more useful than specific, empirically verified ones; nevertheless, a provisional structure, cautiously used, will facilitate the remainder of the discussion.

As observed in section 4, one cannot meaningfully discuss optimal decisions of agents under arbitrary sequences $\{x_t\}$ of future shocks. As an alternative characterization, then, let policies and other disturbances be viewed as stochastically disturbed functions of the state of the system, or (parametrically)

$$x_t = G(y_t,\lambda,\eta_t), \tag{16}$$

where G is known, λ is a fixed parameter vector, and η_t a vector of disturbances. Then the remainder of the economy follows

$$y_{t+1} = F(y_t,x_t,\theta(\lambda),\epsilon_t), \tag{17}$$

where, as indicated, the behavioral parameters θ vary systematically with the parameters λ governing policy and other "shocks." The econometric problem in this context is that of estimating the function $\theta(\lambda)$.

In a model of this sort, a *policy* is viewed as a change in the parameters λ, or in the function generating the values of policy variables at particular times. A change in policy (in λ) affects the behavior of the system in two ways: first by altering the time series behavior of x_t; second by leading to modification of the behavioral parameters $\theta(\lambda)$ governing the rest of the system. Evidently, the way this latter modification can be expected to occur depends crucially on the way the policy change is carried out. If the policy change occurs by a sequence of decisions following no discussed or pre-announced pattern, it will become known to agents only gradually, and then perhaps largely as higher variance of "noise." In this case, the movement to a new $\theta(\lambda)$, if it occurs in a stable way at all, will be unsystematic, and econometrically unpredictable. If, on the other hand, policy changes occur as fully discussed and understood changes in *rules*, there is some hope that the resulting structural changes can be forecast on the basis of estimation from past data of $\theta(\lambda)$.

It is perhaps necessary to emphasize that this point of view towards conditonal forecasting, due originally to Knight and, in modern form, to Muth, does not attribute to agents unnatural powers of instantly divining the true structure of policies affecting them. More modestly, it asserts that agents' responses become predictable to outside observers only when there can be some confidence that agents and observers share a common view of the nature of the shocks which must be forecast by both.

The preference for "rules versus authority" in economic policy making suggested by this point of view, is not, as I hope is clear, based on any demonstrable optimality properties of rules-in-general (whatever that might mean). There seems to be no theoretical argument ruling out the possibility that (for example) delegating economic decision-making authority to some individual or group might not lead to superior (by some criterion) economic performance than is attainable under some, or all, hypothetical rules in the sense of (16). The point is rather that this possibility cannot *in principle* be substantiated empirically. The only *scientific* quantitative policy evaluations avail-

able to us are comparisons of the consequences of alternative policy rules.

7 Concluding Remarks

This essay has been devoted to an exposition and elaboration of a single syllogism: given that the structure of an econometric model consists of optimal decision rules of economic agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any change in policy will systematically alter the structure of econometric models.

For the question of the short-term forecasting, or tracking ability of econometric models, we have seen that this conclusion is of only occasional significance. For issues involving policy evaluation, in contrast, it is fundamental; for it implies that comparisons of the effects of alternative policy rules using current macroeconometric models are invalid regardless of the performance of these models over the sample period or in ex ante short-term forecasting.

The argument is, in part, destructive: the ability to forecast the consequences of "arbitrary," unannounced sequences of policy decisions, currently claimed (at least implicitly) by the theory of economic policy, appears to be beyond the capability not only of the current-generation models, but of conceivable future models as well. On the other hand, as the consumption example shows, conditional forecasting under the alternative structure (16) and (17) is, while scientifically more demanding, entirely operational.

In short, it appears that policy makers, if they wish to forecast the response of citizens, must take the latter into their confidence. This conclusion, if ill-suited to current econometric practice, seems to accord well with a preference for democratic decision making.

Notes

1. See Phelps et al. [31], Phelps' earlier [30] and Friedman [13].

2. The earliest wage-price sector embodying the "trade-off" is (as far as I know) in the 1955 version of the Klein-Goldberger model [19]. It has persisted, with minimal conceptual change, into all current generation forecasting models. The subsequent shift of the "trade-off" relationship to center stage in policy discussions appears due primarily to Phillips [32] and Samuelson and Solow [33].

3. See in particular Marschak's discussion in [25] (helpfully recalled to me by T. D. Wallace) and Tinbergen's in [36], especially his discussion of "qualitative policy" in ch. 5, pp. 149–185.

4. See, for example, de Menil and Enzler [6], Hirsch [16] and Hymans [17].

5. The ill-fated Brookings model project was probably the ultimate expression of this view.

6. See, for example, Gordon [14].

7. A good account of this and other aspects of forecasting in theory and practice is provided by Klein [20]. A fuller treatment is available in Evans and Klein [9].

8. See Klein [20].

9. This is not to suggest that all parameter drift is due to this source. For example, shifts in production functions due to technological change are probably well described by a random walk scheme.

10. Of course, the hypothesis continues to be tested as new data sources become available, and anomalies continue to arise. (For a recent example, see Mayer [26]). Thus one may expect that, as with most "confirmed" hypotheses, it will someday be subsumed in some more general formulation.

11. In [12], Friedman proposes an alternative view to Muth's, namely that the weight used in averaging past incomes (λ) is the same as the discount factor used in averaging future incomes (β). It is Muth's theory, rather than Friedman's of [12], which is consistent with the cross-section tests based on relative variances mentioned above.

12. Let σ_v^2 be the variance of v_t and $\sigma_{\Delta w}^2$ be the variance of the increments of w_t, then the relationship is

$$\lambda = 1 + \frac{1}{2}\frac{\sigma_{\Delta w}^2}{\sigma_v^2} - \frac{\sigma_{\Delta w}}{\sigma_v}\sqrt{1 + \frac{1}{4}\frac{\sigma_{\Delta w}^2}{\sigma_v^2}}\,.$$

13. [10], p. 113.

14. In particular, the low estimates of 'α' (see [15], Table 2, p. 400), which should equal capital's share in value added, are probably due to a sizeable transitory component in a variable which is treated theoretically as though it were subject to permanent changes only.

15. A tax credit designed for stabilization would, of course, need to respond to projected movements in the shift variable a_t. In this case, the transition probabilities p and q would vary with indicators (say current and lagged a_t values) of future economic activity. Since my aim here is only to get an idea of the quantitative importance of a correct treatment of expectations, I will not pursue this design problem further.

16. The cost of capital of .14 and the depreciation rate of .15 (for manufacturing equipment) are annual rates from [15]. Since the ratio $(2 + r - \delta)/(r + \delta)$ is

not time-unit free, the assumption that all movement toward the new desired stock of capital takes place in one year is crucial at this point: by defining a *period* as shorter than one year this ratio will increase, and conversely for a longer period.

17. For the reason given in note 16.

18. It should be noted that this conclusion *reinforces* the qualitative conclusion reached by Hall and Jorgenson [15], p. 413.

19. Sargent [34] and I [23] have developed this conclusion earlier in similar contexts.

20. This model is taken, with a few changes, from my earlier [24].

21. This supply function for goods should be thought of as drawn up given a cleared labor market in i. See Lucas and Rapping [22] for an analysis of the factors underlying this function.

22. This implication that the variability in demand affects the slope of the "trade-off" is the basis for the tests of the natural rate hypothesis reported in [24], as well as those by Adie [1] and B. Klein [18].

23. This observation has been made earlier, for exactly the reasons set out in section 5.1, by Eisner [8] and Dolde [7], p. 15.

References

1. Adie, Douglas K., "The Importance of Expectations for the Phillips Curve Relation," Research Paper No. 133, Department of Economics, Ohio University (undated).

2. Ando, Albert and Franco Modigliani, "The Life Cycle Hypothesis of Saving; Aggregate Implications and Tests," *American Economic Review*, v. 53 (1963), pp. 55–84.

3. Cooley, Thomas F. and Edward C. Prescott, "An Adaptive Regression Model," International Economic Review, (June 1973), 364–371.

4. Cooley, Thomas F. and Edward C. Prescott, "Tests of the Adaptive Regression Model," *Review of Economics and Statistics*, (April 1973), 248–256.

5. Cooley, Thomas F. and Edward C. Prescott, "Estimation in the Presence of Sequential Parameter Variation," *Econometrica*, forthcoming.

6. de Menil, George and Jared J. Enzler, "Prices and Wages in the FRB-MIT-Penn Econometric Model," in Otto Eckstein, ed., *The Econometrics of Price Determination Conference* (Washington: Board of Governors of the Federal Reserve System and Social Science Research Council), 1972, pp. 277–308.

7. Dolde, Walter, "Capital Markets and the Relevant Horizon for Consumption Planning," Yale doctoral dissertation, 1973.

8. Eisner, Robert, "Fiscal and Monetary Policy Reconsidered," *American Economic Review*, v. 59 (1969), pp. 897–905.

9. Evans, Michael K. and Lawrence R. Klein, *The Wharton Econometric Forecasting Model*. 2nd, Enlarged Edition (Philadelphia: University of Pennsylvania Economics Research Unit), 1968.

10. Fisher, Franklin M., "Discussion" in Otto Eckstein, ed., *op. cit.* (reference [6]), pp. 113–115.

11. Friedman, Milton, *A Theory of the Consumption Function*. (Princeton: Princeton University Press), 1957.

12. Friedman, Milton, "Windfalls, the 'Horizon', and Related Concepts in the Permanent Income Hypothesis," in Carl F. Christ, *et. al.*, eds., *Measurement in Economics* (Stanford: Stanford University Press), 1963, pp. 3–28.

13. Friedman, Milton, "The Role of Monetary Policy," *American Economic Review*, v. 58 (1968), pp. 1–17.

14. Gordon, Robert J., "Wage-Price Controls and the Shifting Phillips Curve," *Brookings Papers on Economic Activity*, 1972, no. 2, pp. 385–421.

15. Hall, Robert E. and Dale W. Jorgenson, "Tax Policy and Investment Behavior," *American Economic Review*, v. 57 (1967), pp. 391–414.

16. Hirsch, Albert A., "Price Simulations with the OBE Econometric Model," in Otto Eckstein, ed., *op. cit.* (reference [6]), pp. 237–276.

17. Hymans, Saul H., "Prices and Price Behavior in Three U.S. Econometric Models," in Otto Eckstein, ed., *op. cit.* (reference [6]), pp. 309–322.

18. Klein, Benjamin, "The Effect of Price Level Unpredictability on the Composition of Income Change," unpublished working paper, April, 1973.

19. Klein, Lawrence R. and Arthur S. Goldberger, *An Econometric Model of the United States, 1929–1952*. (Amsterdam: North Holland), 1955.

20. Klein, Lawrence R., *An Essay on the Theory of Economic Prediction*. (Helsinki: Yrjo Jahnsson Lectures), 1968.

21. Knight, Frank H., *Risk, Uncertainty and Profit*. (Boston: Houghton-Mifflin), 1921.

22. Lucas, Robert E., Jr. and Leonard A. Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy*, v. 77 (1969), pp. 721–754.

23. Lucas, Robert E., Jr., "Econometric Testing of the Natural Rate Hypothesis," in Otto Eckstein, ed., *op. cit.* (reference [6]), pp. 50–59.

24. Lucas, Robert E., Jr., "Some International Evidence on Output-Inflation Trade-Offs," *American Economic Review*, v. 63 (1973).

25. Marschak, Jacob, "Economic Measurements for Policy and Prediction," in William C. Hood and Tjalling G. Koopmans, eds., *Studies in Econometric Method*, Cowles Commission Monograph 14 (New York: Wiley), 1953, pp. 1–26.

26. Mayer, Thomas, "Tests of the Permanent Income Theory with Continuous Budgets," *Journal of Money, Credit, and Banking*, v. 4 (1972) pp. 757–778.

27. Modigliani, Franco and Richard Brumberg, "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in K. K. Kurihara, ed., *Post-Keynesian Economics*. (New Brunswick: Rutgers University Press), 1954.

28. Muth, John F., "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, v. 55 (1960), pp. 299–306.

29. Muth, John F., "Rational Expectations and the Theory of Price Movements," *Econometrica*, v. 29 (1961), pp. 315–335.

30. Phelps, Edmund S., "Money Wage Dynamics and Labor Market Equilibrium," *Journal of Political Economy*, v. 76 (1968), pp. 687–711.

31. Phelps, Edmund S., *et al.*, *The New Microeconomics in Employment and Inflation Theory*. (New York: Norton), 1970.

32. Phillips, A. W., "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," *Economica*, v. 25 (1958), pp. 283–299.

33. Samuelson, Paul A. and Robert M. Solow, "Analytical Aspects of Anti-Inflation Policy," *American Economic Review*, v. 50 (1960), pp. 177–194.

34. Sargent, Thomas J., "A Note on the 'Accelerationist' Controversy," *Journal of Money, Credit, and Banking*, v. 3 (1971), pp. 721–725.

35. Tinbergen, Jan, *On the Theory of Economic Policy*. (Amsterdam: North Holland), 1952.

36. Tinbergen, Jan, *Economic Policy: Principles and Design*. (Amsterdam: North Holland), 1956.

Some International Evidence on Output-Inflation Tradeoffs

This paper reports the results of an empirical study of real output-inflation tradeoffs, based on annual time-series from eighteen countries over the years 1951–67. These data are examined from the point of view of the hypothesis that average real output levels are invariant under changes in the time pattern of the rate of inflation, or that there exists a "natural rate" of real output. That is, we are concerned with the questions (i) does the natural rate theory lead to expressions of the output-inflation relationship which perform satisfactorily in an econometric sense for all, or most, of the countries in the sample, (ii) what testable restrictions does the theory impose on this relationship, and (iii) are these restrictions consistent with recent experience?

Since the term "natural rate theory" refers to varied aggregation of models and verbal developments,[1] it may be helpful to sketch the key elements of the particular version used in this paper. The first essential presumption is that *nominal* output is determined on the aggregate demand side of the economy, with the division into real output and the price level largely dependent on the behavior of *suppliers* of labor and goods. The second is that the partial "rigidities" which dominate short-run supply behavior result from suppliers' lack of information on some of the prices relevant to their decisions. The third presumption is that inferences on these relevant, unobserved prices are made optimally (or "rationally") in light of the stochastic character of the economy.

Reprinted from *American Economic Review* 63 (June 1973):326–334 by permission.