ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B 00120 HELSINKI 12 FINLAND TEL. 601 322

Keskusteluaiheita
Discussion papers

Kari Alho

THE ECONOMETRIC FOUNDATION OF OPTIMAL ECONOMIC POLICY: AN ANALYSIS OF THE STATIC CASE

No 21 12.9.1978

THE ECONOMETRIC FOUNDATION OF OPTIMAL ECONOMIC POLICY: AN ANALYSIS OF THE STATIC CASE

by

Kari Alho

The Research Institute
of the Finnish Economy

Abstract

This paper begins by presenting a general framework for linking the various stages of econometric analysis and utilization: specification, estimation and selection of optimal policy. It is then assumed that least squares methods are used to estimate the model, and the influence of the explanatory power of the model on the optimal decision and on the effectiveness of economic policy is assessed. It is concluded that that the statistical criteria for the exercise of optimal policy seem to justify the standard certainty equivalence use of econometric models where the residual, and not the coefficients, is the source of uncertainty.

1. INTRODUCTION

The quantitative relationships produced through econometric analysis are widely used in economic forecasting and policy-making. This varies from the short-cut utilization of impact coefficients to large-scale simulations of the outcome of various policy alternatives. Recently it has become common to assume that a natural extension of the application of statistical methods in economic analysis is to use estimated models in the explicit solution of policy selection problems.

The successive stages of econometric model utilization are naturally linked together. The specification of the model influences the estimation results, which in turn influence the reliability of the forecasts obtained from the model. Policy recommendations depend on both model specification and estimation results.

At the outset the econometric theory of optimal economic policy neglected the vital element of "second order" uncertainty relating to the full chain of econometric analysis. In this paper we first present an outline of the stages of econometric analysis needed for the selection of optimal economic policy. Stress is given to the fact that an essential feature of the optimal procedure is to consider the effects of economic policy on the variability of the target variables around their means and not just on the means as noted by Brainard (1967). After considering the general case we proceed to the linear model and least-squares estimation, and study the properties of optimal policy. Only the static case of policy selection is considered in this paper.

In the latter part of the paper we consider the much discussed problem whether the policy variables (e.g. the money supply) should be used in an active way to finetune the economy or set passively at some constant value. This gives us insight into the influence of the explanatory power of the model on optimal policy and the effectiveness of policy. We can then draw some conclusions about the necessity of treating the impact coefficients as stochastic instead of nonstochastic, which is done in the certainty equivalence procedure.

The following framework can be used in the analysis of optimal economic policy. Assume that there is a relationship between the variable y and variables $x_1,\ldots,x_k$, denoted in short by x. In a policy context we call y the goal or target variable and x the policy or control variables. The basis for statistical model-building is the well-known regression problem: minimize $E_{y,x}(y-f(x))^2$ with respect to f,[1] the solution of which is $\hat{f}$, $\hat{f}(x) = E(y|x)$.

The econometric problem is to find or approximate the generally unknown regression function $E(y|x)$. Let the outcome of the econometric analysis be the (reduced form) function $y(x)$. It is also assumed that this gives the solution to the prediction problem. The prediction of y, conditional upon variables x, is denoted by $y^e(x)$.

---

1) $E_{y,x}$ denotes expectations with respect to variables y and x.

The economic policy problem is defined as follows. In analogy with the previous approaches, the quadratic criterion is used so that deviations of y from its target value y^*, which is assigned by the economic decision-maker, are minimized. We thus have the following formulation of the policy problem:

(1) Minimize $L(x) = \underset{y}{E}[(y-y^*)^2 | x]$ with respect to x subject to the constraint $y = y^e(x) + e$, where e is the residual term[1].

Before proceeding to the general case where, in addition to the residual, the model $y^e(x)$ is considered to be partly unknown and stochastic, let us examine briefly the ideal case where the regression function (the true model) is known to the econometrician. Because the residual is uncorrelated with the regression function, we have

$L(x) = (E(y|x) - y^*)^2 + Ee^2$, hence the optimal $\hat{x}$ satisfies

(2) $y^* = E(y|\hat{x})$.

This is the classical certainty equivalence result of optimal economic policy presented in a general form. In his original derivation, Theil (1958) analyzed the linear model case. When the model is known to the decision-maker and only the residual is unknown, it is optimal to aim directly at the target value so that it is also the expected policy outcome.

---

1) We do not explicitly consider the case where the model also includes uncontrollable explanatory variables. They may be simply seen to be fixed parameters or noncorrelated with $y^e(x)$.

2. OPTIMAL POLICY IN THE GENERAL STATIC CASE

In general it is necessary to carry out the econometric analysis first, which means that the prediction function $y^e(x)$ is stochastic. Two important assumptions underlying the following analysis (and also much of econometric inference) are that $y^e(x)$ is an unbiased prediction function and that the residual e is uncorrelated with x. First, we consider the case of only one target and one policy variable.

The objective function can now be written

$$L(x) = y^{*2} + E(y^e)^2 + \sigma_e^2 - 2y^*(Ey^e).$$

Replacing $E(y^e)^2 = \sigma^2(y^e) + (Ey^e)^2$ and taking the derivative with respect to policy variable x, we get the following basic equation for optimal economic policy

$$(3) \qquad y^* - Ey^e(\hat{x}) = \frac{1}{2}\,\frac{\partial\sigma^2(y^e)/\partial x}{\partial(Ey^e)/\partial x}\ .$$

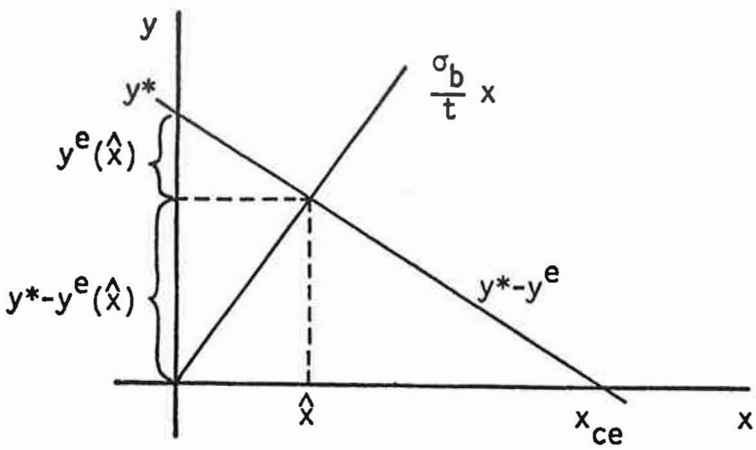
The denominator on the right hand side is assumed to be different from zero, i.e. it is assumed that policy variable x is not inefficient with respect to goal variable y. Equation (3) shows that it is not optimal to aim directly at the target if the variance of the forecast depends on the policy decision, cf. Brainard (1967).

Let us now consider the linear case. The variables are measured as deviations from their means so that $y^e = bx$. Substituting y^e for the unknown expectation $Ey^e = Ey$ allows us to express equation (3) as

$$(4) \qquad y^* - b\hat{x} = \frac{\sigma_b^2}{b}\hat{x} = \frac{\sigma_b}{t}\hat{x}, \text{ where } t = b/\sigma_b$$

using standard regression theory as an analogy. Equation (4) can be illustrated as follows.

Figure 1. Optimal policy in the linear case[1)]



It can be seen that the slope of the right hand side of (4) determines the solution. If variable x is a poor explanatory variable for y, i.e. if t is small, it is optimal to use a cautious policy so that the target y* is not entirely reached. The expected outcome of the optimal policy always lies in the range (0,y*) (or $(\bar{y}, \bar{y}+y^*)$, where $\bar{y}$ is the mean of y). From formula (4) we see that the "better" the explanatory variable (in the t-statistic sense), the more the optimal policy resembles the certainty equivalence policy where the condition $y^* = Ey^e(x)$ is fulfilled. If the variability of the target variable increases (if σ_b increases) and the ratio b/σ_b remains the same, the optimal policy becomes more cautious.

---

1) The certainty equivalence decision is denoted by x_{ce} in the figure.

The analysis can be easily generalized to the case of multiple goals and multiple policy variables. Let $y = (y_1,\dots,y_g)'$ and $y^* = (y_1^*,\dots,y_g^*)'$ be the corresponding vectors of the target variables and their target values. The criterion function is now specified to be $L = E(y-y^*)'\ W(y-y^*)$, where W is a symmetric positive definite matrix of weights. Let $\partial z/\partial x$ be the column vector of partial derivatives $\partial z/\partial x_i$, $i = 1,\dots,k$, and Σ the covariance matrix of the forecasts, i.e. $\Sigma_{ij} = E(y_i^e - Ey_i^e)(y_j^e - Ey_j^e)$. Taking the derivative of L with respect to x gives us the generalization of (3)

$$(5) \qquad \sum_i (y_i^* - Ey_i^e(\hat{x})) \sum_j w_{ij} \frac{\partial[Ey_j^e(\hat{x})]}{\partial x} = \frac{1}{2}\frac{\partial tr\ W\Sigma}{\partial x} .$$

The weights $\sum_j w_{ij}\ \partial(Ey_j^e)/\partial x$ on the left hand side of (5) denote the marginal expected impacts of the policy variables weighted over various goal variables j using intergoal weights w_{ij}. On the right hand side, the marginal impacts of the policy variables on the forecast error variances and covariances are weighted similarly. Suppose now that the model is linear, $y = \Pi x + e$. When calculating Σ we generally need all the covariances between the elements in Π. However, as Johansen (1973) has pointed out, if the matrix W is diagonal, only the covariances between the coefficients in the same equation are needed.

3. THE LINEAR CASE WITH SEVERAL POLICY VARIABLES

Let us now derive the optimal policy in the case where the policy-maker has one goal and several policy variables at his disposal. First, we may note that equation (3) applies for each policy variable x_i, $i=1,\ldots,k$, separately, but that the derivative of the variance of the forecast error with respect to x_i also depends on policy variables other than x_i. We thus have a system of simultaneous equations (5) from which the optimal policy is derived.

The vector of policy variable coefficients is now denoted by b. The unbiased prediction function is thus $y^e = b'x$ and the variance of the prediction (error) is $\sigma^2(y^e) = x'\text{cov}(b)x + \sigma_e^2$, where cov(b) is the covariance matrix of coefficients b. The equation system corresponding to (5) is

$$(6) \qquad b(y^* - y^e) = \text{cov}(b)\hat{x},$$

from which we solve the optimal policy

$$(7) \qquad \hat{x} = (bb' + \text{cov}(b))^{-1}by^*.$$

The matrix inverse in (7) exists generally because the matrix bb' + cov(b) is positive definite. Thus, the optimal solution is unique, as is well-known from Brainard (1967).

Using a suitable matrix inverse formula, we get the result presented by Zellner (1971)

$$(8) \qquad \hat{x} = \frac{\Sigma^{-1}b}{1+b'\Sigma^{-1}b}\, y^*, \text{ where } \Sigma = \text{cov}(b).$$

Let us turn to the linear case and suppose that b has been estimated by least-squares from the sample $(y_t, x_{t1},\ldots,x_{tk})$, $t = 1,\ldots,T$, which we denote by (y,X). Let s^2 be the estimated residual variance of the model. Since $b = (X'X)^{-1}X'y$, the estimate of cov(b) is $s^2(X'X)^{-1}$. In this case the true regression coefficients of the model, which are unknown, are approximated by their estimate b and the covariance matrix Σ by its estimate. We may also give a Bayesian interpretation to this case and consider b and $s^2(X'X)^{-1}$ to be the expectation and covariance matrix of the a posteriori distribution of the true coefficient vector. The optimal policy rule (7) is interpreted to be the best possible policy with the given data (y,X). Expression (8) can now be reduced to

$$(9) \qquad \hat{x} = \frac{X'y}{s^2+y'X(X'X)^{-1}X'y}\, y^*.$$

The numerator consists of the covariances in (y,X) between y and all the explanatory variables x except factor T^{-1}. The denominator is the sum of the residual variance and the sum of squares explained by the model. From (9) we note that the intensities with which policy variables are used in the optimal solution relative to their standard deviations $s(x_i)$ in the sample depend only on the pairwise correlations between the target variable and the policy variables. From (9) we get

$$(10) \qquad \frac{\hat{x}_i}{s(x_i)} = r_{y,x_i}\, s(y)A, \quad A = \frac{y^*}{s^2+y'X'(X'X)^{-1}X'y}\,.$$

The direction in which variable x_i changes relative to its mean in the optimal solution depends on the pairwise correlation of the variable in question and the relation of the target value to its mean.

Formula (10) also shows that the degree to which optimal policy is subject to model misspecification depends on the denominator in A. Generally if we incorrectly omit an explanatory variable, the denominator in A is too small and thus the remaining policy variables have too much to do and deviate more from their means than they would if all relevant variables were included. The smaller the increment in the R^2 of the model when the omitted variable is added, the smaller the bias is.

4. COMPARISON OF THE OUTCOME OF DIFFERENT POLICIES

We now analyze the outcome of the optimal policy conditional on sample (y,X). We make the following substitutions $s^2 = (1-R^2)y'y/(T-k)$, $y'X(X'X)^{-1}X'y = R^2y'y$ and obtain

$$L_o = \underset{y}{E}(y-y^*)^2 = \underset{y}{E}(y^* - b'\hat{X}-e)^2$$

$$(11) \qquad = y^{*2} \frac{1-R^2}{1+(T-k-1)R^2} + s^2 .$$

It can be seen that the value of the optimal policy is an increasing function of the explanatory power of the model, i.e. the loss L_o decreases as R^2 increases. We may also note that the loss is at a minimum when a passive goal-setting policy is pursued, when $y^* = 0$. When T approaches infinity, L_0 approaches the (lower bound) value σ_e^2, the value of L_0 in the certainty equivalence case, which is asymptotically reached in the more general framework as T approaches infinity. The total loss in (11) consists of two parts, the first depending on the multiplicative uncertainty related to the coefficients of the policy variables and the second, s^2, being the additive uncertainty related to the residual of the model. As R^2 approaches one, both components go to zero, but the multiplicative factor goes much more rapidly to zero than the additive one, as we shall see below.

We define passive policy x^P as that which always makes the policy variables identical to their sample means, $x_i^P = \bar{x}_i = 0$. In this case the objective function L takes the value L_p

$$(12) \qquad L_p = E(y-y^*)^2 = E(e-y^*)^2 = y^{*2} + s^2 .$$

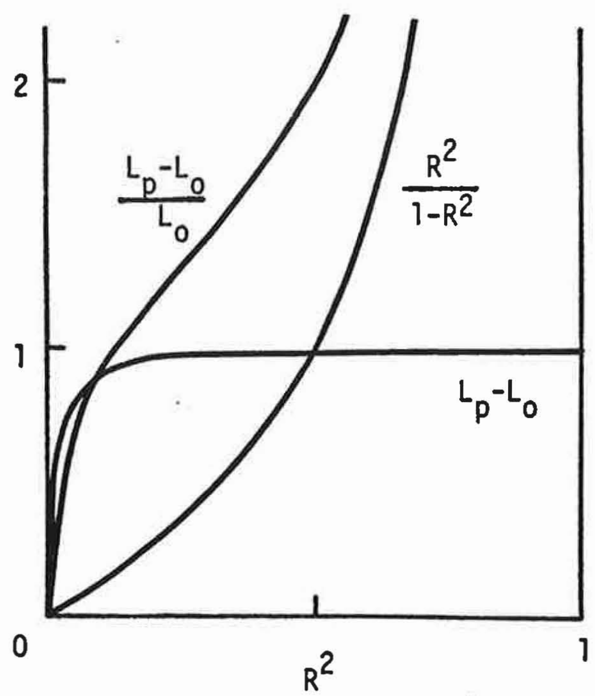
We now compare the two policies. In analogy with the usual F-statistic criterion used to testing whether the coefficient vector differs from zero, we now consider the function $(L_p-L_o)/L_o$. It can be further written

$$(13) \qquad \frac{L_p-L_o}{L_o} = \frac{1-h(R^2)^2}{\frac{s^2}{y^{*2}} + h(R^2)^2} ,$$

where $h(R^2) = (1-R^2)(1+(T-k-1)R^2)^{-1}$. As T approaches infinity, $h(R^2)$ goes to zero if $R^2 > 0$ and $(L_p-L_o)/L_o$ approaches the asymptotic upper limit y^{*2}/σ_e^2 . The gain is thus inversely related to residual variance of the model and directly related to the deviation of the target from the mean of the target variable. Thus, the gain in policy-making from econometric analysis depends partly on the intentions of the policy-maker and is not solely a theoretical problem.

Figure 2 shows (13) as a function of R^2 when T=25, k=4, and $y^{*2} = (T-k)^{-1}y'y$. In the picture the relative increase in the usual sum of squares, $R^2(1-R^2)^{-1}$, is also presented. The gain in optimal policy is uniformly greater than the corresponding gain in estimation if $y^{*2} \geq (T-k)^{-1}y'y$. If $y^{*2} < (T-k)^{-1}y'y$ the curve $(L_p-L_o)/L_o$ in figure 2 rotates downwards and intersects the curve $R^2(1-R^2)^{-1}$ approximately in the point where $R^2 = y^{*2}/(T-k)^{-1}y'y$. So, the gain increases more rapidly as a function of R^2 when the goal-setting is more active, i.e. when y^{*2} increases. On the other hand, the higher y^{*2} is, the less explanatory power is needed to attain a required relative increase in the gain of optimal policy-making over the passive one.

Figure 2. Comparison of optimal and passive policies.



If, instead of the relative gain in (13), we consider the absolute gain $L_p - L_o$ which depends only on the multiplicative factor in (11), we observe that this function very rapidly approaches its maximum value y^{*2} (which is one in figure 2) as R^2 increases from zero. In the range where econometric applications usually are located, say $R^2 > 0.4$, there is virtually no absolute gain in optimal policy over the passive one from extra explanatory power which improves both policies much in the same way. These results show that it is not essential for policy-making to reduce the multiplicative uncertainty related to the coefficients of the policy variables but the overall additive uncertainty, i.e. the residual variance of the model, even though these two factors cannot be separated from each other in econometric analysis.

The previous results reflect the fact that the optimal decision (9) itself does not depend much on R^2 of the model, except when R^2 is small. The reaction of $|\hat{x}_i|$ to an increase in R^2 depends on whether this increase can be attributed to variable x_i or not. If $k=1$ and $|r_{y,x_i}|$ increases from zero, then $|\hat{x}_i|$ increases very rapidly like $L_p - L_o$ from zero to $|y^*/b|$. If $k > 1$ and $|r_{y,x_i}|$ remains the same even though R^2 increases, then generally $|\hat{x}_i|$ decreases. For instance, a change in R^2 from 0.5 to 0.6 would reduce $|\hat{x}_i|$ by 15 per cent.

We should consider carefully the previous results and the assumptions underlying them. In fact, the above L_o is the expected outcome of the optimal policy when the observations (y,X) are given. In order to make a fair comparison, we should also let the observed y vary. Thus the curve in figure 2 should be weighted by the density of the observed R^2 around its "true" value ρ^2. This of course does not alter the above advantage of an active policy over the passive one.

Formula (11) can be written

$$L_o = y^{*2} \left[\frac{s^2}{s^2+ESS}\right]^2 + s^2,$$

where ESS is the sum of squares explained by the model. Approximating the term in parentheses simply by s^2/ESS allows us to apply standard distribution theory. If the true coefficient vector were zero, the ratio $z = ks^2/ESS$ would obey the F-distribution with parameters (T-k,k) when the residuals are normal and independent. We now have

$$E\left(\frac{s^2}{ESS}\right) = \frac{1}{k-2} \text{ and } \sigma^2 \left(\frac{s^2}{ESS}\right) = \frac{2(T-2)}{(T-k)(k-2)^2(k-4)} \quad (k > 4)$$

If T = 20 and k = 5, the expectation of the objective function is $EL_o \sim 0.378y^{*2} + \sigma_e^2$. Generally, the outcome is much better with optimal policy, and the value of the objective function hardly exceeds σ_e^2, which is the value of the objective function in the certainty equivalence case.

5. CONCLUDING REMARKS

The analysis of optimal policy with a model estimated using least squares method leads to some interesting results. The determination of optimal values for policy variables is closely linked to the partial correlations between the instrumental and goal variables, and uncertainty in the estimation of the parameters influences the use of all policy variables in the same way.

We were also able to draw inferences about the relative advantage of an active or optimal policy over a passive one. The criteria for the explanatory power of the model applied in decision-making depend on the intentions of the policy-maker. The more the target deviates from the mean of the target variable, the less explanatory power in the model is needed to reach a similar relative outcome in optimal policy over the passive one. The results also suggest that the standard certainty equivalence use of econometric models is justified because over the relevant range of uncertainty attached to coefficients of the policy variables, the multiplicative uncertainty does not have a significant role in optimal decision-making.

It should, however, be borne in mind that the same model is applied in two different ways: to derive the optimal policy and to calculate the outcome of different policies. The "objective" uncertainty in the "true" model which the statistical analysis considers is not the only source of error which confronts the decision-maker. He also has to choose between the conflicting recommendations of different schools of economic thought. In this wider context, a passice policy may seem desirable because it corresponds to cautious goalsetting, and caution may seem advisable when there are a number of competing models and economic philosophies.

REFERENCES:

BRAINARD, William C., 1967, Uncertainty and the Effectiveness of Policy, American Economic Review, 57, 411-25.

JOHANSEN, Leif, 1973, Targets and instruments under uncertainty, in Economic Structure and Development, Essays in Honour of Jan Tinbergen, (North-Holland, Amsterdam).

THEIL, H., 1958, Economic Forecasts and Policy (North-Holland, Amsterdam).

ZELLNER, Arnold, 1971, An Introduction to Bayesian Inference in Econometrics, (Wiley, New York).