Volume Title: Annals of Economic and Social Measurement, Volume 5, number 3

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm76-3

Publication Date: July 1976

Chapter Title: Methods of Efficient Parameter Estimation in Control Problems

Chapter Author: John B. Taylor

Chapter URL: http://www.nber.org/chapters/c10484

Chapter pages in book: (p. 339 - 347)

Annals of Economic and Social Measurement, 5/3, 1976

METHODS OF EFFICIENT PARAMETER ESTIMATION IN CONTROL PROBLEMS

BY JOHN B. TAYLOR

This paper examines the asymptotic efficiency of parameter estimates generated by single period optimal control rules in the multiple regression model. Several well known rules are shown to generate parameter estimates with unacceptably large variances which decline extremely slowly as the sample size grows. An alternative class of certainty equivalence rules is suggested in order to improve the efficiency of the parameter estimates with relatively little deterioration in control performance. For this class the larger is the number of parameters to be estimated, the poorer is control performance, thus indicating a trade-off between estimation and control.

Recent research on multiperiod control theory in models with unknown parameters has primarily focused on optimization of target variable performance with little emphasis on properties of the parameter estimates. Consequently the parameter estimates which evolve when such control theory is applied, frequently have undesirable statistical properties such as large variances or fail to converge. Such poor parameter estimates do not indicate that the suggested control rules are defective since their stated purpose is to improve system performance rather than estimate parameters which are irrelevant for control. In fact in some cases improving parameter estimates can only be accomplished by sacrificing target variable performance.

However, in many practical applications of control theory it is necessary to obtain good parameter estimates. For example, a policy maker using an econometric model for control purposes might be interested in estimating structural relationships even if these are not immediately used for control. Knowledge of these structural relationships would be useful should the loss function change in some unpredictable way in the future.

The purpose of this paper is to examine the efficiency of parameter estimates generated by control rules in the multivariate regression model,

$$y_t = \beta' x_t + u_t, \qquad t = 1, 2, \ldots \tag{1}$$

where y_t is a scalar target variable with target y^*, β is a vector of k unknown parameters, x_t is a control vector[1] with k components, and u_t is a scalar random variable which is independently and identically distributed with zero mean and finite variance σ^2. The special case where $k = 1$ was discussed by Prescott (1972) from a Bayesian viewpoint and Taylor (1974) from a non-Bayesian viewpoint. In Section 1 we show that single period optimal control rules which have long been suggested for this model (sometimes as approximations to the multiperiod problem) yield parameter estimates which are extremely inefficient. The direction of the vector x_t for these rules is nearly orthogonal to the vector which minimizes the

[1] This model assumes that all k exogenous variables are subject to control. Several different issues arise when not all variables are subject to control. Experimental results on this more general problem are reported in Anderson and Taylor (1975).

generalized variance of the estimate of the vector β. Experimental results indicate that the variance of the estimates of the elements of β decline imperceptively if at all over sample sizes as large as 3,000. In Section 2 an alternative control rule is proposed and is shown to generate more efficient parameter estimates with relatively little deterioration in control performance. The primary measure of efficiency throughout this paper is the variance of the asymptotic distribution of the individual parameter estimates, though selected experimental results are reported to indicate the rate of convergence to these asymptotic distributions. As is usual with statistical results based on large sample theory, they should be used with caution in small samples.

1. Single Period Optimal Control Rules

Many of the control rules which have been suggested for this problem have been derived using Bayesian methods and are either optimal Bayes rules or approximations to these rules. In the most common formulation of the problem, the criterion of performance is the sum $\sum_{t=1}^{T}(y_t - y^*)^2$ with u_t normally distributed. A control rule is then chosen so as to minimize the expectation of this sum with respect to the prior distribution of β and the distribution of u_t. Since the dynamic programming solution is intractable for large T, an approximate solution is suggested by truncating the problem at small T. The most common approximation is to set $T=1$. This single period solution has been referred to as the sequential updating solution by Zellner (1971) and the myopic solution by Prescott (1972), and is one form of open loop strategy suggested in th engineering literature. This is also the rule considered by Brainard (1967) where emphasis was on macroeconomic policy implications of uncertain β. We refer to this type of rule as a *single period optimal Bayes rule.*

When σ^2 is known and the prior distribution of β is uninformative, the single period optimal Bayes rule is given by

$$x_{t+1} = \frac{A_t\hat{\beta}_t y^*}{\hat{\beta}'_t A_t \hat{\beta}_t + \sigma^2}, \qquad t = k, k+1, \dots, \tag{2}$$

where $A_t = \sum_{i=1}^{t} x_i x'_i$ and where

$$\hat{\beta}_t = A_t^{-1} \sum_{i=1}^{t} x_i y_i \tag{3}$$

and where we assume that $x_1, x_2, \dots, x_k$ are given such that $\sum_{i=1}^{k} x_i x'_i$ is nonsingular. It is well known that the optimal portfolio of instruments in the vector x_{t+1} is such that those instruments corresponding to elements of β with relatively large variances will be relatively small in absolute value; and that negative covariances can be exploited by applying appropriate weights to some instruments.

The sequential updating rule in (2) is not a certainty equivalence rule because it does not satisfy the equation $y^* = \hat{\beta}'_t x_{t+1}$. Basu (1974) has suggested a modification to (2) which will be referred to as the *single period optimal certainty equivalence rule* and is defined by

$$x_{t+1} = \frac{A_t\hat{\beta}_t}{\hat{\beta}'_t A_t \hat{\beta}_t} y^*. \tag{4}$$

Although both (2) and (4) are designed for the purpose of stabilizing y_t about y^*, one might hope that the sequence of parameter estimates that evolve when these rules are used have desirable statistical properties. This however is not the case. The resulting estimates of β have extremely large variances even for very large samples as is illustrated experimentally below.

The reason for these poor parameter estimates can be seen intuitively by examining the single period Bayes risk

$$(5) \qquad E[(y_{t+1}-y^*)^2|x_t, y_t, \ldots, x_1, y_1]=\sigma^2(1+x'_{t+1}A_tx_{t+1})+(\hat{\beta}'_tx_{t+1}-y^*)^2.$$

The single period optimal Bayes rule is chosen to minimize (5). The single period optimal certainty equivalence rule minimizes the first term on the right hand side of (5) constraining the second term to equal zero as has been pointed out by Basu (1974). The equality

$$(6) \qquad (1+x'_{t+1}A_tx_{t+1})=\frac{|A_t+x_{t+1}x'_{t+1}|}{|A_t|}$$

indicates however that the first term on the right hand side of (5) is inversely proportional to the determinant of the conditional covariance matrix of $\hat{\beta}_{t+1}$ (that is the generalized conditional variance) given observations through time t. Thus the single period optimal Bayes rule maximizes the generalized conditional variance of $\hat{\beta}_{t+1}$ added to $(\hat{\beta}'x_{t+1}-y^*)^2$, and the single period optimal certainty equivalence rule maximizes the generalized conditional variance of $\hat{\beta}_{t+1}$ subject to $\hat{\beta}'_tx_{t+1}-y^*=0$. This suggests why the rules result in poor parameter estimates. The direction of x_{t+1} is almost orthogonal to the direction which will give most information about β. This also suggests that obtaining better parameter estimates might necessarily result in a deterioration of the performance of the target variable.

The experimental results reported in Tables 1 and 2 illustrate the estimation problem in a model with two unknown parameters and two controls. The parameter values for the Monte Carlo experiment reported here were $\beta_1=1$, $\beta_2=2$, $\sigma^2=1$, $y^*=1$, and the initial values for the two control variables x_{1t} and x_{2t} were $x_{11}=1$, $x_{12}=3$, $x_{21}=1$, $x_{22}=2$. The initial variances are therefore var$(\hat{\beta}_{1t})=5$, var $(\hat{\beta}_{2t})=10$, and cov $(\hat{\beta}_{1t}, \hat{\beta}_{2t})=-7$ when $t=2$. Several models with different initial conditions (including smaller initial variances) were also examined with similar results. The estimated moments for this rule are calculated on the bases of 100 replications of a 3,000 period time horizon. The term $E(y^*-\beta'x_{t+1})^2$ represents the expected one period loss minus σ^2, and declines as t^{-1}. For $t>100$, the estimate of $E(y^*-\beta'x_{t+1})^2$ is within the 95% probability interval (computed using the χ^2 distribution) of t^{-1}. Thus the asymptotic distribution of $y^*-\beta'x_t$ in this model with more than one unknown parameter and more than one control is the same as that derived by Taylor (1974) in the case of a single control and a single unknown parameter.

Note however that the variance of the parameter estimates generated by both rules has a very small decline after $t=10$ and shows no tendency to converge to zero. Plots of the observed values of $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ are also illuminating and are found in Figure 1 at $t=100$ (the small ellipse should be ignored for now) for rule (2). Rule (4) results in almost exactly the same picture. Although the variances of

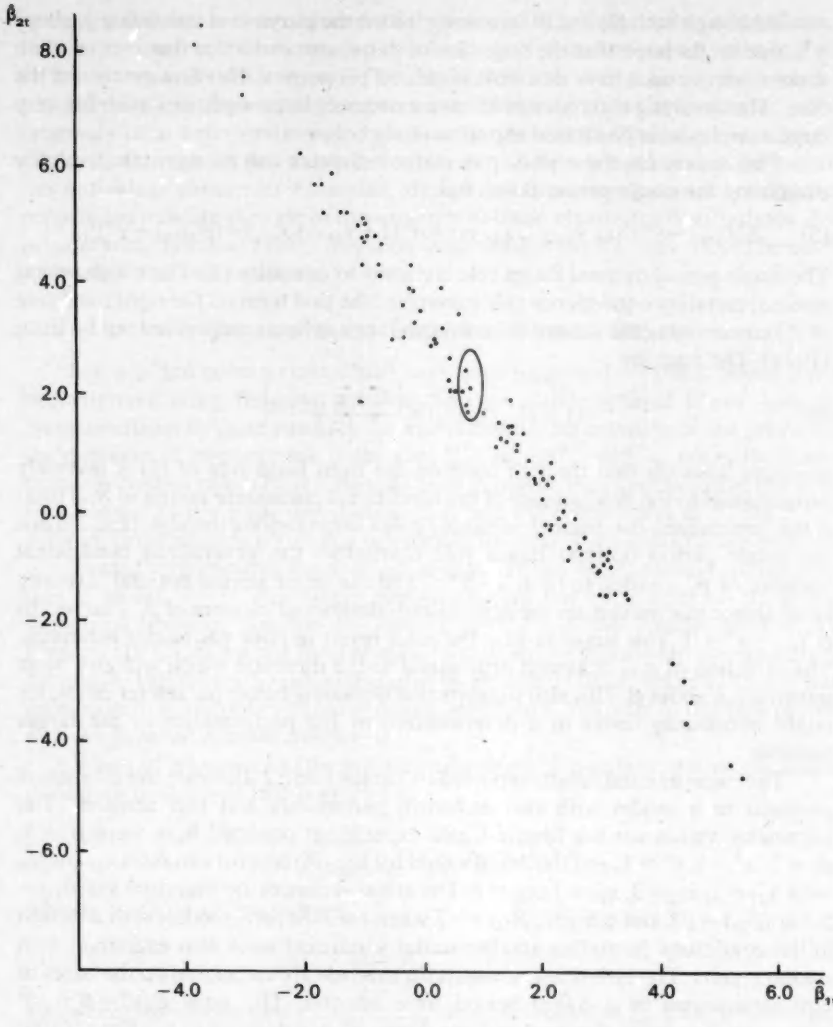

Figure 1 Scatter of Parameter Estimates at $t = 100$ for Single Period Bayes Optimal Control Rule.

$\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ do not seem to converge to zero, the variance of a linear combination of $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ does converge to zero. This linear combination is determined by the initial conditions of the experiment. Its slope in the $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ plane is approximately equal to the slope of the characteristic vector corresponding to the smallest root of A_2. As our primary purpose is to illustrate the inefficiency of these parameter estimates, and to suggest a more efficient method, we will not pursue the properties of these limiting distributions.

TABLE 1
ESTIMATED MOMENTS FOR THE SINGLE PERIOD OPTIMAL BAYES RULE

t	$E(y^*-\beta' x_{t+1})$	$E(y^*-\beta' x_{t+1})^2$	$E\mu_t$	$E\nu_t$
10	1.4246(−2)	1.1871(−2)	1.6947(1)	6.7079(−2)
50	1.2452(−2)	9.2865(−3)	2.7110(1)	6.7401(−2)
100	1.2673(−2)	5.9180(−3)	3.9692(1)	6.7576(−2)
500	2.7140(−3)	1.5648(−3)	1.4170(2)	6.7843(−2)
1000	4.6860(−4)	8.7835(−4)	2.7042(2)	6.7893(−2)
2000	−3.3138(−3)	5.2660(−4)	5.2917(2)	6.7920(−2)
3000	−6.7492(−4)	3.5588(−4)	7.8791(2)	6.7929(−2)

t	$E(\hat\beta_{1t})$	var $(\hat\beta_{1t})$	$E(\hat\beta_{2t})$	var $(\hat\beta_{2t})$	cov $(\hat\beta_{1t}, \hat\beta_{2t})$
10	1.2245	3.9665	1.6729	7.9716	−5.5723
50	1.2353	3.8011	1.6627	7.5589	−5.3205
100	1.2525	3.8571	1.6380	7.4839	−5.3474
500	1.2213	3.8040	1.6490	7.5462	−5.3500
1000	1.2174	3.7571	1.6477	7.4684	−5.2918
2000	1.2189	3.7299	1.6332	7.3630	−5.2368
3000	1.2223	3.7606	1.6372	7.4221	−5.2801

TABLE 2
ESTIMATED MOMENTS FOR THE SINGLE PERIOD OPTIMAL CERTAINTY EQUIVALENCE RULE

t	$E(y^*-\beta' x_{t+1})$	$E(y^*-\beta' x_{t+1})^2$	$E\mu_t$	$E\nu_t$
10	−1.2828(−3)	1.2658(−2)	1.9623(1)	6.7193(−2)
50	3.0818(−3)	9.4183(−3)	2.7433(1)	6.7408(−2)
100	6.3162(−3)	5.8517(−3)	4.0211(1)	6.7581(−2)
500	8.8642(−4)	1.5525(−3)	1.4288(2)	6.7844(−2)
1000	−4.8982(−4)	8.7559(−4)	2.7193(2)	6.7893(−2)
2000	−3.7995(−3)	5.2874(−4)	5.3102(2)	6.7920(−2)
3000	−1.0042(−3)	3.5575(−4)	7.8997(2)	6.7929(−2)

t	$E(\hat\beta_{1t})$	var $(\hat\beta_{1t})$	$E(\hat\beta_{2t})$	var $(\hat\beta_{2t})$	cov $(\hat\beta_{1t}, \hat\beta_{2t})$
10	1.2246	3.9675	1.6729	7.9731	−5.5738
50	1.2354	3.8011	1.6627	7.5581	−5.3207
100	1.2525	3.8568	1.6380	7.4837	−5.3474
500	1.2213	3.8037	1.6489	7.5454	−5.3496
1000	1.2174	3.7571	1.6476	7.4682	−5.2918
2000	1.2189	3.7300	1.6332	7.3632	−5.2369
3000	1.2222	3.7606	1.6372	7.4221	−5.2801

Tables 1 and 2 also present the estimated mean of the smaller root ν_t and the larger root μ_t of A_t. The smaller root increases extremely slowly as suggested by the direction of each x_t, while the larger root increases on the order of t for large t.

2. PARAMETER ESTIMATING CERTAINTY EQUIVALENCE RULES

In this section we consider some certainty equivalence rules which generate estimates of specified linear combinations of the vector β and which are consistent asymptotically normal (using the $\sqrt{t}$ norm). Therefore the variance of these estimates will tend to converge to zero as t^{-1}, which is certainly an improvement over the estimates just considered whose variances do not converge at all. For all these rules $\beta' x_t$ is also consistent asymptotically normal (viewed as an estimate of y^* and again using the $\sqrt{t}$ norm), and therefore the mean square of $y^* - \beta' x_t$ tends to converge to zero as t^{-1}, the same rate as the rules considered in the previous section. However, if more than one linear combination of parameters is to be estimated, then the variance of the limiting distribution of $\sqrt{t}(y^* - \beta' x_t)$ is larger than for the single period optimal rules.

Consider first the case where interest is in estimating a single preassigned linear combination $\phi = c'\beta$ of the elements of β, where c is a specified vector of k elements with at least one element (assumed without loss of generality to be the first, c_1) not equal to zero. Then a certainty equivalence rule for which this linear combination is consistent asymptotically normal is defined such that x_1 is an arbitrary but nonzero multiple of c and

(7) $$x_{t+1} = \frac{cy^*}{\hat{\phi}_t}, \qquad t = 1, 2, \ldots,$$

where $\hat{\phi}_t$ is the least squares estimate of ϕ. Note that because $y_t = y^* \phi / \hat{\phi}_{t-1} + u_t$ and $x_{1t} = y^* c_1 / \hat{\phi}_{t-1}$ we have

(8) $$\hat{\phi}_t = c_1 \frac{\sum_{i=1}^t x_{1i} y_i}{\sum_{i=1}^t x_{1i}^2}, \qquad t = 1, 2, \ldots.$$

The following two theorems give the asymptotic properties of $\hat{\phi}$ and $y^* - \beta' x_t$.

Theorem 1. In model (1) *with* $c'\beta \neq 0$ *and the sequence* x_t *defined by* (7), $\hat{\phi}_t$ *converges to* ϕ *with probability one and* $\beta' x_t$ *converges to* y^* *with probability one.*

Proof. From equation (8)

(9) $$\hat{\phi}_t = \phi + c_1 \frac{\sum_{i=1}^t x_{1i} u_i}{\sum_{i=1}^t x_{1i}^2}.$$

Using the results of Theorem 1 of Taylor (1974) the second term on the right hand side of (9) converges with probability one. This implies from (7) that x_{1t} does not converge to zero and therefore that $\sum_{i=1}^t x_{1i}^2$ diverges with probability one. This implies that the second term on the right hand side of (9) coverges to zero with probability one, which completes the proof of the theorem.

Theorem 2. Under the assumptions of Theorem 1, $\sqrt{t}(\hat{\phi}_t - \phi)$ *has a limiting normal distribution with mean* 0 *and variance* $\sigma^2 \phi^2 / (y^*)^2$, *and* $\sqrt{t}(y^* - \beta' x_t)$ *has a limiting normal distribution with mean zero and variance* σ^2.

Proof. From (9) we have that

(10) $$\sqrt{t}(\hat{\phi}_t - \phi) = c_1 \sqrt{t} \frac{\sum_{i=1}^t x_{1i} u_i}{\sum_{i=1}^t x_{1i}^2}.$$

Using the result of Theorem 1 above and Theorem 2 of Taylor (1974), the right hand side of (10) converges in distribution to the normal with mean zero and variance equal to $c_1^2\sigma^2(\operatorname{plim}_{t\to\infty}\sum_{i=1}^t x_{1i}^2/t)^{-1}=\sigma^2\phi^2/(y^*)^2$. The second result follows from the equality

$$\sqrt{t}(y^*-\beta' x_{t+1})=\sqrt{t}y^*(\hat{\phi}_t-\phi)/\hat{\phi}_t.$$

Theorem 2 indicates that the estimate of ϕ generated by this rule has acceptable statistical properties, its variance declining approximately as t^{-1}. Note also that the asymptotic variance of $y^*-\beta' x_{t+1}$ is equal to that of the single period optimal rules examined experimentally in Section 1. In this case no sacrifice in asymptotic control performance is necessary to get this efficient parameter estimate. Note also the special case of this rule where $c'\beta=\beta_1$. Then only β_1, the first element of β, is estimated.

To obtain estimates of more than one linear combination of parameter estimates, the following procedure can be used. Suppose $H\leq k$ linear combinations of the elements of β are to be estimated. Let these be $\phi_h=c_h'\beta$, $h=1,\ldots,H$, where each of the c_h are vectors with k elements with at least one element of each (assumed to be the hth element of c_h, labeled c_{hh}) not equal to zero. Partition the set of integers into H sets $I_1,\ldots,I_H$ such that I_h contains the integers $jH+h$, $j=0,1,2,\ldots$. Let $I_h(t)$ be the intersection of each of these sets with all integers less than or equal to t. Then a control rule for which the estimates of ϕ_h are consistent asymptotically normal is given by $x_1,\ldots,x_H$ each an arbitrary nonzero multiple of $c_1,\ldots,c_H$ and

$$x_{t+1}=\frac{c_h y^*}{\hat{\phi}_{ht}},\qquad \text{if } (t+1)\varepsilon I_h,\qquad t=H,H+1,\ldots,\tag{11}$$

where $\hat{\phi}_{ht}$ is the least squares estimate of ϕ_h. Because $y_t=\phi_h y^*/\hat{\phi}_{ht-1}+u_t$ and $x_{ht}=c_{hh}y^*/\hat{\phi}_{ht-1}$ for $t\in I_h$ we can write this least squares estimate as

$$\hat{\phi}_{ht}=c_{hh}\frac{\sum_{i\in I_h(t)}x_{hi}y_i}{\sum_{i\in I_h(t)}x_{hi}^2},\qquad h=1,\ldots H.\tag{12}$$

Theorem 3. In model (1) *with the sequence x_t defined by* (7) *if $c_h'\beta\neq 0$ for $h=1,\ldots,H$, then $\hat{\phi}_{ht}$ each converge to zero with probability one and $\beta' x_t$ converges to y^* with probability one.*

Proof. From equation (12) we have that

$$\hat{\phi}_{ht}=\phi_h+c_{hh}\frac{\sum_{i\in I_h(t)}x_{hi}u_i}{\sum_{i\in I_h(t)}x_{hi}^2}\tag{13}$$

Since the number of elements in each $I_h(t)$ diverges to infinity as $t\to\infty$, the arguments of Theorem 1 can be used to show that the second term on the right hand side of (13) converges to zero with probability one. Thus $\hat{\phi}_{ht}\to\phi_h$ and from (11) $\beta' x_t\to y^*$ with probability one.

Theorem 4. Under the assumptions of Theorem 3, $\sqrt{t}(\hat{\phi}_{1t}-\phi_1),\ldots,\sqrt{t}(\hat{\phi}_{Ht}-\phi_H)$ *have a limiting normal distribution with means* 0, *variances $\sigma^2H\phi_1^2/(y^*)^2,\ldots,\sigma^2H\phi_H^2/(y^*)^2$, and covariances* 0. *In addition $\sqrt{t}(y^*-\beta' x_t)$ has a limiting normal distribution with mean zero and variance $H\sigma^2$.*

Proof. To establish the limiting distribution we must show that $z_t = \sqrt{t}\sum_{h=1}^{H} a_h(\hat{\phi}_{ht} - \phi_h)$ has a limiting normal distribution with mean zero and variance

$$V = \frac{H\sigma^2}{(y^*)^2} \sum_{h=1}^{H} \phi_h^2 a_h^2 \tag{14}$$

for any set of real numbers $a_1, \ldots, a_h$. Using equation (13) and the method of Theorem 2 we find that the difference between z_t and

$$\frac{H}{y^*} \sum_{h=1}^{H} \left(a_h \phi_h \frac{\sum_{i \in I_h(t)} u_i}{\sqrt{t}} \right) \tag{15}$$

has a zero probability limit. Each of the H terms in (15) converge in distribution to the normal by the central limit theorem, and each of these are independent since they contain no common random elements. Thus (15) converges to the normal distribution with mean 0 and variance V given in (14). The limiting distribution of $\sqrt{t}(y^* - \beta' x_t)$ then follows using equation (11).

The implication of Theorem 4 is that when the certainty equivalence rules given here are used, the estimates of several linear combinations of the elements of β can have variances which decline approximately as t^{-1}. Further each of the linear combinations have equal asymptotic "t-ratios," since the asymptotic standard deviation of each estimate is the same proportion of its true value. However, the larger the number of linear combinations that are estimated the larger is the asymptotic variance of $y^* - \beta' x_t$ which is our measure of control performance. Thus, using this method of analysis, there is a tradeoff between parameter estimation and control. Only one parameter combination can be estimated without cost.

A special case of the parameter estimating control rule in (1) is when each of the k parameters of β are to be estimated. In that case the procedure is simply to use only one instrument at a time, setting the other to zero, and switching to a different instrument each time period.

Table 3 presents the results of the use of such a rule[2] in the model introduced in Section 1. Comparing Table 3 with Tables 1 and 2 we can see that there is an enormous improvement in the efficiency of the parameter estimates. As predicted by the above asymptotic theory the variances of $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ are approximated by $2/t$ and $4/t$ respectively so that the ratio of the estimated coefficient to the standard error is close to $(2t)^{-1/2}$ for both. In addition the mean square of $y^* - \beta' x_t$ is approximated by $2/t$ as is predicted by Theorem 4. Since two parameters are being estimated this value is larger than that in Tables 1 and 2. Finally in Figure 1 the ellipse centered at (1, 2) contains all the 100 estimated values of $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ at $T = 100$.

3. Concluding Remarks

This paper has been concerned with the problem of efficient parameter estimation in a regression model where the dependent variable is being controlled

[2] Since the initial conditions of this model do not conform to those of rule (11), the experiment is run as a sensitivity test of the asymptotic theory to more general initial conditions. All estimated second moments reported in Table 3 for $t > 100$ are within the 95% probability interval of the asymptotic approximation.

TABLE 3
ESTIMATED MOMENTS FOR THE PARAMETER ESTIMATING CERTAINTY EQUIVALENCE RULE

t	$E(y^*-\beta' x_{t+1})$	$E(y^*-\beta' x_{t+1})^2$	$E\mu_t$	$E\nu_t$
10	−7.0609(−2)	9.9945(−2)	3.1970(1)	4.6497
50	1.2063(−2)	1.5823(−2)	5.1968(1)	1.2744(1)
100	1.5570(−2)	1.1417(−2)	7.6562(1)	2.0422(1)
500	2.2346(−4)	3.1722(−3)	2.7543(2)	7.2874(1)
1000	−2.1217(−3)	2.0152(−3)	5.2571(2)	1.3575(2)
2000	−4.3591(−3)	1.0202(−3)	1.0276(3)	2.6198(2)
3000	−2.4752(−3)	6.2541(−4)	1.5290(3)	3.8809(2)

t	$E(\hat{\beta}_{1t})$	var $(\hat{\beta}_{1t})$	$E(\hat{\beta}_{2t})$	var $(\hat{\beta}_{2t})$	cov $(\hat{\beta}_{1t}, \hat{\beta}_{2t})$
10	1.0460	1.3183(−1)	1.9909	2.1271(−1)	−1.1975(−1)
50	1.0085	3.1645(−2)	2.0577	7.1733(−2)	−1.5749(−2)
100	1.0035	1.2683(−2)	2.0544	4.5837(−2)	−3.3406(−3)
500	1.0067	3.6111(−3)	2.0068	1.2760(−2)	−6.4295(−4)
1000	1.0034	1.7229(−3)	1.9997	7.8874(−3)	−2.6985(−4)
2000	0.9851	9.8709(−4)	1.9933	3.8351(−3)	−4.0770(−5)
3000	1.0016	6.7184(−4)	1.9963	2.4927(−3)	8.1619(−5)

by all the independent variables. Some well known single period optimal rules were shown to give unacceptable parameter estimates. As an alternative to these rules, a class of parameter estimating certainty equivalence rules were proposed and shown to have much greater efficiency as parameter estimators with relative small reduction in control efficiency. For these rules the more parameters that are estimated the lower is the control efficiency. This suggests a tradeoff between estimation and control.

Columbia University

REFERENCES

[1] Anderson, T. W. and J. B. Taylor (1975), "Some Experimental Results on the Statistical Properties of Least Squares Estimates in Control Problems," forthcoming *Econometrica.*

[2] Basu, A. (1974), "Control Level Determination in Regression Models," Technical Report No. 139, IMSSS, Stanford University.

[3] Brainard, W. (1967), "Uncertainty and the Effectiveness of Policy," *American Economic Review*, 62, 411–425.

[4] Prescott, E. C. (1972), "The Multiperiod Control Problem Under Uncertainty," *Econometrica*, 40, 1043–1058.

[5] Taylor, J. B. (1974), "Asymptotic Properties of Multiperiod Control Rules in the Linear Regression Model," *International Economic Review*, 15, 472–484.

[6] Tse, E. (1974), "Adaptive Dual Control Methods," *Annals of Economic and Social Measurement*, 3, 65–84.

[7] Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.