Federal Reserve Bank of Minneapolis
Research Department Working Paper

TEMPORAL AGGREGATION BIAS AND GOVERNMENT POLICY EVALUATION*

Lawrence J. Christiano

Federal Reserve Bank of Minneapolis

Working Paper 302

April 1986

*I have benefitted from discussions with Albert Marcet and Tom Sargent.

1. Introduction

This note illustrates some of the effects of temporal aggregation bias on government policy evaluation. Following Lucas [1976], a change in government policy is modeled as a change in the value of a structural parameter. The objective of the econometrician is to determine the effect of the policy change on the conditional expectation of a variable of interest. In the example, the data are generated by a continuous time model, while the econometrician proceeds as though the underlying model is discrete, with a timing interval that coincides with the data sampling interval. The example provides an indication about the empirical circumstances in which the consequences of this specification error will be severe.

In the example, agents set employment, $y(t)$, to maximize discounted profits subject to a given real wage rate process, $\{-z(t)\}$, and productivity shock process, $\{x(t)\}$. Temporal aggregation produces greater distortion, in the example, the greater the costs of rapidly adjusting $y(t)$ and the greater the proportion of variation in $\{z(t)\}$ that is concentrated at high frequencies. In both of these cases agents' decisions are relatively sensitive to contemporaneous and near-term expected movements in $\{z(t)\}$ and $\{x(t)\}$. These are dynamic relations that are not well captured by the econometrician's discrete time version of the model.

The simplicity of the example not only makes analytical results possible, it also permits abstracting from other temporal aggregation effects that have been studied extensively in the literature. 1/ In the example, temporal aggregation does not distort the predictive performance of the econometrician's model in the absence of structural change. This is because both the misspecified discrete approximate model and the true continuous time model imply the same second moment restrictions on the sampled data available to the econometrician. Models that are likely to be used in practical situations are unlikely to share this characteristic. This is fortunate since differences between overidentifying restrictions are a potentially valuable source of information about which model--discrete or continuous--is better suited to the data. 2/

The next section presents the model and our results. Conclusions appear in section 3.

2. The Illustration

Suppose agents choose $y(t)$, $t \geq 0$ to maximize

$$\lim_{N\to\infty} E_0 \frac{1}{N} \int_0^N \{(x(t)+z(t))y(t) - \frac{1}{2} H_1 y(t)^2 - \frac{1}{2} H_2 [Dy(t)]^2\}dt \tag{1a}$$

subject to

$$x(t) = Dv_t^x \tag{1b}$$

$$(\sigma+D)z(t) = v_t^z. \tag{1c}$$

Here, (v_t^x, v_t^z) is a continuous time vector white noise and D is the time derivative operator. 3/ (See Bergstrom [1976] for a discussion of the technical details associated with the D operator and a continuous time white noise.) Also, σ, H_1, and H_2 are positive constants. The model is a continuous time version of the model of employment studied in Sargent [1978; 1979, Chapter XIV, section 1].

Exploiting certainty equivalence and applying standard calculus of variations techniques, the solution to (1) is given by [4/]

$$\text{(2)}\qquad \begin{pmatrix} Dy(t) \\ Dz(t) \end{pmatrix} = \begin{bmatrix} \rho & \frac{1}{H_2(\sigma-\rho)} \\ 0 & -\sigma \end{bmatrix} \begin{pmatrix} y(t) \\ z(t) \end{pmatrix} + u(t),$$

where $\rho = -(H_1/H_2)^{1/2}$ and $u(t) = (-\frac{1}{H_2} v_t^x, v_t^z)$. The representation of $\{y(t), z(t)\}$ sampled at the integers is

$$\text{(3)}\qquad \begin{pmatrix} y(t) \\ z(t) \end{pmatrix} = \begin{bmatrix} e^{\rho} & \frac{e^{\rho} - e^{-\sigma}}{H_2(\sigma-\rho)(\sigma+\rho)} \\ 0 & e^{-\sigma} \end{bmatrix} \begin{pmatrix} y(t-1) \\ z(t-1) \end{pmatrix} + \xi(t),$$

where $\xi(t)$ is a discrete time, bivariate white noise.

We suppose that the econometrician has observations on $\{y_t, z_t, t=T, T-1, T-2, \ldots\}$. He/she incorrectly believes the data are generated by the following discrete problem. Maximize over contingency plans for $\{y_0, y_1, \ldots\}$

$$\text{(4a)}\qquad \lim_{N\to\infty} E_0 \frac{1}{N} \sum_{t=0}^{N} \{(x(t)+z(t))y(t) - \frac{1}{2} h_1 y(t)^2 - \frac{1}{2} h_2 (y(t)-y(t-1))^2\}$$

subject to

$$\text{(4b)}\qquad x(t) = \varepsilon_t$$

$$\text{(4c)}\qquad z(t) = \phi z(t-1) + \delta_t,$$

where ε_t and δ_t are white noise processes, uncorrelated at nonzero lags. Here $|\phi| < 1$, and h_1, h_2 are positive constants. The reduced form representation corresponding to this problem is

$$\text{(4d)} \qquad \begin{pmatrix} y(t) \\ z(t) \end{pmatrix} = \begin{bmatrix} \lambda & \frac{\lambda\phi}{h_2(1-\lambda\phi)} \\ 0 & \phi \end{bmatrix} \begin{pmatrix} y(t-1) \\ z(t-1) \end{pmatrix} + w_t$$

$$\text{(5)} \qquad = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{pmatrix} y(t-1) \\ z(t-1) \end{pmatrix} + w_t,$$

say. (See Sargent [1979] for details about the derivation of (4).) Here λ is the unique solution to

$$\lambda + \frac{1}{\lambda} = \frac{h_1}{h_2} + 2, \ |\lambda| < 1,$$

and $\{w_t\}$ is a discrete time 2 x 1 vector white noise.

The structural parameters of the discrete time problem, (3), are exactly identified from the reduced form parameters in (5) as follows:

$$\text{(6)} \qquad \begin{aligned} \lambda &= b_{11} \\ \phi &= b_{22} \\ h_2 &= \frac{b_{11}b_{22}}{b_{12}(1-b_{11}b_{22})} \\ h_1 &= h_2(\lambda + \frac{1}{\lambda} - 2). \end{aligned}$$

The reduced form restrictions implied by the econometrician's approximate model are

$$
\begin{aligned}
& b_{21} = 0 \\
(7) \qquad & b_{11}, b_{22} \in (-1,1) \\
& \text{sign}\,(b_{11}b_{22}) = \text{sign}\,(b_{12}),
\end{aligned}
$$

where sign (x) denotes the sign of the variable x. It can be verified that these restrictions are satisfied regardless of values of the structural parameters of the underlying parent model (e.g., (3)). Consequently, in population the econometrician will obtain the true values of the reduced form parameters, $[b_{ij}]$. It follows that from the point of view of the data, models (1) and (4) are observationally equivalent. On the other hand, we shall see that (1) and (4) differ in their implications for the consequences of a policy intervention to be described below.

Substituting from (3) into (6) we obtain the econometrician's population estimates of the structural parameters of his/her misspecified model as a function of the structural parameters of the true model, (1):

$$
\begin{aligned}
& \lambda = e^{\rho} \\
& \phi = e^{-\sigma} \\
(8) \qquad & h_2 = \frac{e^{\rho-\sigma} c_1^2 (\sigma-\rho)(\sigma+\rho)}{(e^{\rho}-e^{-\sigma})(1-e^{(\rho-\sigma)})} \\
& h_1 = h_2(\lambda + \frac{1}{\lambda} - 2).
\end{aligned}
$$

We consider a government policy intervention at time T, which has the effect of altering the value of σ. The econometrician is informed of the resulting effect on the parameter of the discrete time representation of $z(t)$; i.e., the effect on $\phi = e^{-\sigma}$.

The mean of the conditional distribution of $y(T+1)$, which we denote by $E_T(y(T+1);\sigma)$, in the absence of a government policy change is

$$(9) \qquad E_T(y(T+1);\sigma) = b_{11}y(T) + b_{12}z(T).$$

With the change in government policy, the conditional mean of $y(T+1)$ is

$$(10) \qquad E_T(y(T+1);\bar{\sigma}) = b_{11}y(T) + \bar{b}_{12}z(T).$$

In (10), account has been taken of the fact that a change in σ has no effect on b_{11} (see (3) and (5)). Here,

$$(11) \qquad \bar{b}_{12} = \frac{e^{\rho} - e^{-\bar{\sigma}}}{H_2(\bar{\sigma}-\rho)(\bar{\sigma}+\rho)}$$

where $\bar{\sigma}$ is the new value of σ. The variables ρ and H_2 are left at their pre-intervention values because they are not functions of σ.

The econometrician will produce the following conditional forecast of $y(T+1)$, after being informed of the change in ϕ to $\bar{\phi} = e^{-\bar{\sigma}}$:

$$(12) \qquad E_T^e(y(T+1);\bar{\sigma}) = b_{11}y(T) + \tilde{b}_{12}z(T).$$

Here b_{11} is left unchanged because, by (6), the econometrician believes b_{11} to be a function of λ which is unrelated to ϕ. By (6) the econometrician sets

$$\tilde{b}_{12} = \frac{b_{11}\bar{\phi}}{h_2(1-b_{11}\bar{\phi})}$$

$$= \frac{(e^{\rho}-e^{-\sigma})(1-e^{(\rho-\sigma)})}{H_2(\sigma-\rho)(\sigma+\rho)(1-e^{(\rho-\bar{\sigma})})} e^{(\sigma-\bar{\sigma})}.$$

Evidently, the extent to which the econometrician's conditional one-step-ahead forecast of y(T+1) is misleading depends on the discrepancy between $\tilde{b}_{12}$ and $\bar{b}_{12}$. It is therefore of interest to consider the following expression,

$$(13) \qquad \frac{\tilde{b}_{12}}{\bar{b}_{12}} = \frac{(e^{\rho}-e^{-\sigma})(1-e^{(\rho-\sigma)})(\bar{\sigma}-\rho)(\bar{\sigma}+\rho)}{(e^{\rho}-e^{-\bar{\sigma}})(1-e^{(\rho-\bar{\sigma})})(\sigma-\rho)(\sigma+\rho)} e^{(\sigma-\bar{\sigma})}.$$

Now, the variance of z(t) is inversely proportional to σ. Consider an intervention which has the effect of increasing the variance of z(t) by 10 percent, so that $\bar{\sigma} = .9 \times \sigma$. Then (13) may be written,

$$(14) \qquad \psi(\rho,\sigma) = \frac{\tilde{b}_{12}}{\bar{b}_{12}} = \frac{(e^{\rho}-e^{-\sigma})(1-e^{(\rho-\sigma)})(.9\sigma-\rho)(.9\sigma+\rho)}{(e^{\rho}-e^{-.9\sigma})(1-e^{\rho-.9\sigma})(\sigma-\rho)(\sigma+\rho)} e^{.1\sigma}.$$

As was stated in the introduction, one would expect the consequence of timing misspecification (here measured by ψ) to be more severe the greater the importance of adjustment costs in the firm's objective function. This is because the presence of adjustment costs creates an incentive on the part of firms to condition current decisions on forecasts of the future. In addition, the presence of adjustment costs also has the effect of creating a dependency between current decisions and decisions made in the recent past. The parameter H_2 measures the relative weight of costs of adjustment in the example of this paper. It appears in ψ via $\rho = -\sqrt{H_1/H_2}$. The reasoning in this paragraph would therefore suggest that ψ departs from 1 with a decrease in $|\rho|$. It is supported by the fact, easily verified, that $\psi(\rho,\sigma) \rightarrow 1$ as $\rho \rightarrow -\infty$ for any value of σ. The case $\rho = -\infty$ corresponds to $H_2 = 0$, when there are no adjustment costs and the model is static. 5/ Evidence on the behavior of ψ for less extreme values of ρ is provided in Table 1, which will be discussed below.

It was also mentioned in the introduction that one would expect the consequence of specification error to be more severe the greater the proportion of the variation in $\{z(t)\}$ that is concentrated at high frequencies. Below, we provide a measure, P, of the fraction of the total variation in $\{z(t)\}$ concentrated at high frequencies. We show that P increases with σ, so that we expect ψ to depart from 1 as σ increases.

The spectral density of $\{z(t)\}$ in (1c) is

$$(15) \qquad S_z(\omega) = \left(\frac{c}{\sigma^2+\omega^2}\right),$$

where c is a constant. As is well known, the variance of $\{z(t)\}$ can be expressed as the integral of the spectrum of $\{z(t)\}$ over all frequencies as follows:

$$\text{(16)} \qquad \text{Var}\,(z(t)) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_z(\omega)d\omega = \frac{1}{\pi}\int_{0}^{\infty} S_z(\omega)d\omega,$$

since $S_z(\omega) = S_z(-\omega)$. The fraction of the variance of $\{z(t)\}$ that is concentrated in high (i.e., above π) frequencies, $P(\sigma)$, is therefore

$$\text{(17)} \qquad P(\sigma) = 1 - \frac{\int_0^{\pi} S_z(\omega)d\omega}{\int_0^{\infty} S_z(\omega)d\omega}.$$

Carrying out the integration, one gets [6]

$$\text{(18)} \qquad P(\sigma) = 1 - \frac{2}{\pi}\tan^{-1}\left(\frac{\pi}{\sigma}\right),$$

where the range of $\tan^{-1}$ is restricted to $[0,\pi)$. Evidently, $P(\sigma) \to 1$ as $\sigma \to \infty$ and $P(\sigma) \to 0$ as $\sigma \to 0$. Therefore, as σ increases, the fraction of total variation in $\{z(t)\}$ that is concentrated in high frequencies increases.

The numerical results reported in Table 1 are consistent with the arguments of the preceding two paragraphs. Table 1 indicates that, over the range of parameter values considered, the econometrician is led to overpredict the effect on b_{12} of the government policy change. The range of error for the parameter values considered is from 0 to 23 percent.

Table 1

$\psi(\rho,\sigma)$

$\sigma\downarrow$ $\rho\rightarrow$	-.5	-1.0	-1.5	-2.0	-2.5	-3.0	-3.5	-4.0
.5		1.004	1.004	1.003	1.003	1.003	1.002	1.002
1.0	1.016		1.014	1.013	1.012	1.011	1.010	1.008
1.5	1.035	1.034		1.030	1.027	1.024	1.022	1.019
2.0	1.061	1.059	1.056		1.048	1.044	1.039	1.035
2.5	1.094	1.092	1.087	1.082		1.069	1.062	1.055
3.0	1.133	1.130	1.125	1.118	1.110		1.091	1.082
3.5	1.179	1.175	1.169	1.160	1.150	1.139		1.114
4.0	1.229	1.225	1.218	1.208	1.196	1.183	1.116	

3. Conclusion

An example has been presented in which the reduced form distortions usually associated with temporal aggregation are not present. These are the distortions studied by Sims [1971], Geweke [1978], Marcet [1985], Hansen and Sargent [1983], Christiano and Eichenbaum [1986], and Christiano [forthcoming] and--if present in our example--would result in models (1) and (4) being observationally distinguishable. The two features of the setup of this paper that rule out these distortions are: (a) that models (1) and (4) place no overidentifying restrictions on their respective reduced forms, other than that $\{z(t)\}$ fails to Granger-cause $\{y(t)\}$, and (b) that the continuous time representation is specified to have a first order vector autoregressive representation in continuous time.

Despite the absence of the usual temporal aggregation effects, the example of the paper nevertheless exhibits a potential for bias due to model timing misspecification. The bias arises when the model is put to use evaluating the consequence of a government policy intervention. The source of bias is that the econometrician who mistakenly proceeds as though the true model is discrete misspecifies the mapping from the government policy parameter to the parameter of the private decision rule. In the present context, the policy parameter is ϕ, and the relevant decision rule parameter is b_{12}. Equation (6) indicates that the econometrician believes that the value of b_{12} in the post-intervention period $(\bar{b}_{12})$ is related to ϕ and the pre-intervention reduced form parameters (b_{11}, b_{12}, b_{22}) as follows:

$$\bar{b}_{12} = b_{12}\left(\frac{\phi}{b_{22}}\right)\left(\frac{1-b_{11}b_{22}}{1-b_{11}\phi}\right).$$

On the other hand, the link between the actual post-intervention value of b_{12} (denoted by $\tilde{b}_{12}$) and $(\phi, b_{11}, b_{12}, b_{22})$ is

$$\tilde{b}_{12} = b_{12}\left[\frac{(\log b_{22}+\log b_{11})(\log b_{22}-\log b_{11})(b_{11}-\phi)}{(\log\phi+\log b_{11})(\log\phi-\log b_{11})(b_{11}-b_{22})}\right].$$

Footnotes

1/See, for example, several of the papers in Bergstrom [1976], Sims [1971], Geweke [1978], Marcet [1978], Hansen and Sargent [1983], Christiano and Eichenbaum [1986], and Christiano [forthcoming].

2/Christiano [forthcoming] provides a formal technique for determining which model timing interval seems best suited to the data.

3/A discussion of (1b) appears in Appendix A.

4/See Hansen and Sargent [1980] for a detailed discussion of the solution to (1).

5/When $H_2 = 0$ and $\rho = -\infty$, $\{y(t)\}$ is not a well-defined stochastic process. This is because I specified the productivity shock, $\{x(t)\}$, (see (1b)) to be a generalized stochastic process (see Appendix A for a discussion). When $H_2 = 0$, then $\{y(t)\}$ is a generalized stochastic process too since in this case, the solution to (1) is just $y(t) = (x(t)+z(t))/H_1$. I specified $\{x(t)\}$ the way I did in (1b) in order to guarantee that the solution to (1) be the vector first order differential equation given in (2). This in turn was desired in order to simplify the analysis. As is easily verified, the autoregressive matrix in (3) possesses a well-defined limit as $H_2 \to 0$. On the other hand, the variance of the first element in $\{\xi(t)\}$ explodes to infinity as $H_2 \to 0$.

6/Here, we use the facts $\int_a^b (\frac{1}{\sigma^2+\omega^2})d\omega = \frac{1}{\sigma} \tan^{-1} (\frac{\omega}{\sigma})|_a^b$ and $\tan^{-1} (0) = 0$.

References

Bergstrom, A. R., 1976, editor, Statistical Inference in Continuous Time Economic Models, North-Holland, Amsterdam.

____________, 1983, "Gaussian Estimation of Structural Parameters in Higher Order Continuous Time Dynamic Models," Econometrica 51: 117-152.

Christiano, L. J., forthcoming, "A Method for Estimating the Timing Interval in a Linear Econometric Model, With an Application to Taylor's Model of Staggered Contracts," Journal of Economic Dynamics and Control.

Christiano, L. J. and M. Eichenbaum, 1986, "Temporal Aggregation and Structural Inference in Macroeconomics," presented to Carnegie-Rochester Conference on Public Policy, April 11-12, 1986.

Geweke, J. B., 1978, "Temporal Aggregation in the Multivariate Regression Model," Econometrician 46: 643-662.

Hansen, L. P. and T. J. Sargent, 1980, "Methods for Estimating Continuous Time Rational Expectations Models From Discrete Data," Federal Reserve Bank of Minneapolis Staff Report 59, April.

Hansen, L. P. and T. J. Sargent, 1984, "Two Difficulties in Interpreting Vector Autoregressions," Federal Reserve Bank of Minneapolis Working Paper 227, September.

Lucas, R. E., Jr., 1976, "Econometric Policy Evaluation: A Critique," The Phillips Curve and Labor Markets, edited by K. Brunner and A. H. Meltzer, Carnegie-Rochester Conferences on Public Policy 1, Amsterdam: North-Holland.

Marcet, A., 1985, "Temporal Aggregation of Economic Time Series," unpublished manuscript.

Sargent, T. J., 1978, "Estimation of Dynamic Labor Demand Schedules Under Rational Expectations," Journal of Political Economy 86, no. 6.

____________, 1979, Macroeconomic Theory, Academic Press.

Sims, C. A., 1971, "Discrete Approximations to Continuous Time Lag Distributions in Econometrics," Econometrica 38: 545-564.

Appendix A

An Interpretation of the Productivity Shock Specification, (1b)

The specification of the technology shock process in (1b) requires elaboration since, taken literally, it implies a contradiction. The model assumes that firms observe and react to a continuous record on $\{Dv_t^x\}$. On the other hand, $\{Dv_t^x\}$ is not a realizable stochastic process and so cannot literally be "observed."

One way to interpret (1b) is as the limit of the following sequence of representations as $\gamma \downarrow 0$:

$$\text{(A1)} \qquad x(t) = \frac{D}{(\gamma D+1)^2} v_t^x .$$

For each $\gamma > 0$, (A1) is an ordinary stochastic process, and therefore, realizable. In addition, for each $\gamma > 0$, the optimal rule for setting $Dy(t)$ is

$$\text{(A2)} \qquad Dy(t) = \rho y(t) + \left[\frac{1}{H_2(\sigma-\rho)}\right] z(t) + \frac{1}{H_2} E_t \int_0^\infty e^{\rho\tau} x(t+\tau) d\tau .$$

Making use of results in Hansen and Sargent [1980], (A1) and (A2) imply

$$\text{(A3)} \qquad Dy(t) = \rho y(t) + \left[\frac{1}{H_2(\sigma-\rho)}\right] z(t) + \left[\frac{-D(-\gamma\rho+1)^2-\rho(\gamma D+1)^2}{H_2(\gamma D+1)^2(-\gamma\rho+1)^2(D+\rho)}\right] v_t^x .$$

Note that (A3) becomes the first equation in (2) as $\gamma \downarrow 0$.