Journal of Econometrics 33 (1986) 387-397. North-Holland

A STANDARD ERROR FOR THE ESTIMATED STATE VECTOR OF A STATE-SPACE MODEL

James D. HAMILTON

University of Virginia, Charlottesville, VA 22901, USA

Received July 1985, final version received June 1986

This paper motivates an estimate of the variance of the estimated state vector $\hat{\boldsymbol{x}}_t$ in a state-space model when the vector of parameters characterizing system dynamics (8) must be estimated from the data.

1. Introduction

A state-space model describes the dynamics of an output vector $\boldsymbol{y}_t$ in terms of observed inputs $\boldsymbol{u}_t$ and an unobserved state vector $\boldsymbol{x}_t$:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{F}(\boldsymbol{\theta})\boldsymbol{x}_t + \boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{u}_t + \boldsymbol{w}_{t+1}, \\ \boldsymbol{y}_t &= \boldsymbol{H}(\boldsymbol{\theta})\boldsymbol{x}_t + \boldsymbol{D}(\boldsymbol{\theta})\boldsymbol{u}_{t-1} + \boldsymbol{v}_t. \end{aligned} \tag{1}$$

Here the matrices $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$, and $\boldsymbol{D}$ are functions of an unknown vector of parameters $\boldsymbol{\theta}$, as are the variances of the white-noise disturbance terms $\boldsymbol{w}_{t+1}$ and $\boldsymbol{v}_t$:

$$\begin{aligned} \mathrm{E}\begin{bmatrix} \boldsymbol{w}_t\boldsymbol{w}_s' & \boldsymbol{w}_t\boldsymbol{v}_s' \\ \boldsymbol{v}_t\boldsymbol{w}_s' & \boldsymbol{v}_t\boldsymbol{v}_s' \end{bmatrix} &= \begin{bmatrix} \boldsymbol{Q}(\boldsymbol{\theta}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}(\boldsymbol{\theta}) \end{bmatrix} \quad \text{if } t = s, \\ &= \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \quad \text{otherwise.} \end{aligned}$$

The usefulness of this framework for analyzing system dynamics, long appreciated by engineers, has increasingly become recognized by economists. The original economic applications were the time-varying coefficient models introduced by Cooley and Prescott (1973) and Rosenberg (1973), given new life recently in the provocative work of Sims (1982). Subsequent researchers have illustrated that a wide range of economic models in which unobserved components are important may also fit into this framework. Wall (1980) and

Watson (1985) suggested that the unobserved expectations of agents, which play such a crucial role in modern economic theory, might be modelled as acting the part of the state vector $\boldsymbol{x}_t$. Recent studies by Fama and Gibbons (1982), Burmeister and Wall (1982), and Hamilton (1985) generated empirical estimates of agents' expectations based on this principle. Engle and Watson (1981) and Engle, Lilien and Watson (1985) considered the unobserved factors underlying metropolitan wage rates and housing demand as other economic applications of the state-space framework. Howrey (1984) illustrated the usefulness of the Kalman filter in forecasting inventory series in the presence of measurement error and data revisions. An excellent survey of these and other uses of the Kalman filter was provided by Engle and Watson (1985).

This paper presents an algorithm for generating a standard error for the econometric estimate of the state vector $\hat{\boldsymbol{x}}_t$, when the true value of the parameter vector $\boldsymbol{\theta}$ is unknown.[1] For clarity of exposition, I first review the case when $\boldsymbol{\theta}$ is presumed known (section 2) and then discuss in section 3 the case when $\boldsymbol{\theta}$ must be estimated from the data. An application is presented and discussed in section 4.

2. The standard error of the estimated state vector when $\boldsymbol{\theta}$ is known

In the economic applications of the Kalman filter discussed above, a classical statistician would want to think of the state vector $\boldsymbol{x}_t$, though unobserved, as having some particular 'true' value; for example, in the first quarter of 1980, the American public may have expected inflation of, say $x_t = 8.4\%$. The same classical statistician would likewise speak of the variance of the sequence $\{x_t\}$ as the average squared deviations of such values from their population mean. In this paper, by contrast, I find it more convenient to adopt the Bayesian perspective [e.g., Meinhold and Singpurwalla (1983)] that the 'true' value $\boldsymbol{x}_t$ is itself a random variable; any prior information we have about $\boldsymbol{x}_t$ or any inference drawn from observation of y and $\boldsymbol{u}$ is then reflected in the probability distribution of x_t. Let

$$\boldsymbol{z}_t \equiv (\boldsymbol{y}_t', \boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_1', \boldsymbol{u}_{t-1}', \boldsymbol{u}_{t-2}', \ldots, \boldsymbol{u}_0')'$$

denote the vector of observations through date t. Our knowledge about $\boldsymbol{x}_t$ based on observation of $\boldsymbol{z}_t$ (and based on the certain knowledge that $\boldsymbol{\theta} = \boldsymbol{\theta}_0$) can be summarized by the probability statement*

$$\boldsymbol{x}_t \sim \mathrm{N}(\hat{\boldsymbol{x}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0), \hat{\boldsymbol{P}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)).$$

[1] I introduced the proposed technique in Hamilton (1985); the present paper elaborates on the Bayesian motivation underlying this algorithm.

*Here we are assuming Gaussian distributions for $\boldsymbol{w}, \boldsymbol{v}$ and our prior for $\boldsymbol{x}_0$.

Thus, at any date $t, \hat{\boldsymbol{x}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)$ is the econometrician's best guess[3] as to the value of $\boldsymbol{x}_t$, and $\hat{\boldsymbol{P}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)$ characterizes our uncertainty about this guess:

$$\hat{\boldsymbol{P}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0) = \mathrm{E}\{[\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)]' | \boldsymbol{z}_t, \boldsymbol{\theta}_0\},$$

where this expectation is with respect to our Bayesian posterior probability distribution.

Meinhold and Singpurwalla (1983) noted that in this case when $\boldsymbol{\theta}$ is known, the Kalman filter can usefully be thought of as a recursive algorithm for updating these Bayesian distributions. Specifically, suppose we characterized our uncertainty about the initial state of the system through the prior $\boldsymbol{x}_0 \sim \mathrm{N}(\hat{\boldsymbol{x}}_0, \hat{\boldsymbol{P}}_0)$. Then for $t = 0, 1, 2, \ldots, T-1$, the moments of our posterior distribution for $\boldsymbol{x}_{t+1}$ can be updated (based on observation of $\boldsymbol{z}_{t+1}$) by the Kalman filter:

$$\begin{aligned} K_{t+1} = {} & \{[\boldsymbol{F}(\boldsymbol{\theta}_0)]\hat{\boldsymbol{P}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)[\boldsymbol{F}(\boldsymbol{\theta}_0)]' + \boldsymbol{Q}(\boldsymbol{\theta}_0)\}[\boldsymbol{H}(\boldsymbol{\theta}_0)]' \\ & \times\{[\boldsymbol{H}(\boldsymbol{\theta}_0)]([\boldsymbol{F}(\boldsymbol{\theta}_0)]\hat{\boldsymbol{P}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)[\boldsymbol{F}(\boldsymbol{\theta}_0)]' + \boldsymbol{Q}(\boldsymbol{\theta}_0)) \\ & \times[\boldsymbol{H}(\boldsymbol{\theta}_0)]' + \boldsymbol{R}(\boldsymbol{\theta}_0)\}^{-1}, \end{aligned} \tag{2a}$$

$$\begin{aligned} \hat{\boldsymbol{x}}_{t+1}(\boldsymbol{z}_{t+1}, \boldsymbol{\theta}_0) = {} & \boldsymbol{F}(\boldsymbol{\theta}_0)\hat{\boldsymbol{x}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0) + \boldsymbol{G}(\boldsymbol{\theta}_0)\boldsymbol{u}_t \\ & + \boldsymbol{K}_{t+1}\{\boldsymbol{y}_{t+1} - \boldsymbol{H}(\boldsymbol{\theta}_0) \\ & \times[\boldsymbol{F}(\boldsymbol{\theta}_0)\hat{\boldsymbol{x}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0) + \boldsymbol{G}(\boldsymbol{\theta}_0)\boldsymbol{u}_t] - \boldsymbol{D}(\boldsymbol{\theta}_0)\boldsymbol{u}_t\}, \end{aligned} \tag{2b}$$

$$\begin{aligned} \hat{\boldsymbol{P}}_{t+1}(\boldsymbol{z}_{t+1}, \boldsymbol{\theta}_0) = {} & \{\boldsymbol{I} - \boldsymbol{K}_{t+1}\boldsymbol{H}(\boldsymbol{\theta}_0)\} \\ & \times\{[\boldsymbol{F}(\boldsymbol{\theta}_0)]\hat{\boldsymbol{P}}_t(\boldsymbol{z}_t, \boldsymbol{\theta}_0)[\boldsymbol{F}(\boldsymbol{\theta}_0)]' + \boldsymbol{Q}(\boldsymbol{\theta}_0)\}. \end{aligned} \tag{2c}$$

Smoothing algorithms described by Cooley, Rosenberg and Wall (1977) and Anderson and Moore (1979) exist for adjusting these estimates-to reflect data for the entire sample, i.e., for generating $\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)$ and $\boldsymbol{P}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)$. If a Bayesian were then asked how accurately the state vector $\boldsymbol{x}_t$ has been estimated, the answer would be

$$\mathrm{E}\{[\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)]' | \boldsymbol{z}_T, \boldsymbol{\theta}_0\} = \hat{\boldsymbol{P}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0). \tag{3}$$

[3] That is, best in the sense of minimizing a quadratic loss function and using information up through date t.

3. The standard error of the estimated state vector when θ must he estimated from the data

I now investigate how this same question should be answered when the econometrician does not know the true value of the parameter vector $\boldsymbol{\theta}$, i.e., I seek to evaluate

$$\mathrm{E}\{[\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(z_T)][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(z_T)]'|z_T\}. \tag{4}$$

The optimal estimate of the state vector when $\boldsymbol{\theta}$ must be estimated from the data is well known [see, for example, Anderson and Moore (1979, ch. 10)]. Briefly, given a posterior probability distribution for the parameter vector $\boldsymbol{\theta}$, $(f(\boldsymbol{\theta}; z_T))$, we calculate

$$\hat{\boldsymbol{x}}_t(z_T) = \int \hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}) f(\boldsymbol{\theta}; z_T)\,\mathrm{d}\boldsymbol{\theta}. \tag{5}$$

For a large sample, $f(\boldsymbol{\theta}; z_T)$ approaches the asymptotic distribution of the M.L.E.:[4]

$$\boldsymbol{\theta} \sim \mathrm{N}(\hat{\boldsymbol{\theta}}(z_T), \hat{\Lambda}(z_T)),$$

i.e., our best guess of the value of $\boldsymbol{\theta}$ is the M.L.E. $\hat{\boldsymbol{\theta}}(z_T)$ and the variance associated with this guess is the M.L.E.'s asymptotic variance-covariance matrix $\hat{\Lambda}(z_T)$, the same matrix that the classical statistician would use to construct confidence intervals.

The approximation to eq. (5) that is most commonly adopted is simply to replace $\boldsymbol{\theta}_0$ in eq. (2b) with the maximum likelihood estimate $\hat{\boldsymbol{\theta}}(z_T)$. An alternative procedure, which I have found in my applications to produce virtually identical results,[5] is to generate a series of, say, 200 Monte Carlo draws from a $\mathrm{N}(\hat{\boldsymbol{\theta}}(z_T), \hat{\Lambda}(z_T))$ distribution, yielding a sample $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_{200})$. For each $\boldsymbol{\theta}_i$, run through the Kalman filter iteration (2), with smoother if desired. This will generate a time series.of estimated state vectors $\{\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}_i)\}$ and associated conditional variances $\{\boldsymbol{P}_t(z_T, \boldsymbol{\theta}_i)\}$. Now, for each t, average the

[4]See Zellner (1971, pp. 31-34). Specific algorithms for implementing maximum likelihood estimation were described bv Burmeister and Wall (1982), Watson and Engle (1983) and Hamilton (1985). More work needs to be done on the asymptotic properties of the M.L.E. in this context; see Ljung (1978), Ljung and Caines (1979), Nichols and Quinn (1980), Pagan (1980) and Tanaka (1983).

[5]If $\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta})$ were linear in 8, the two methods would produce exactly identical results. Since the asymptotic distribution theory is based on a local linearization around 8, there seems little basis for choosing $\hat{\boldsymbol{x}}_t(z_T)$ over $\hat{\boldsymbol{x}}_t(z_T, \hat{\boldsymbol{\theta}}(z_T))$. However, I show below that the contribution of parameter uncertainty to the variance of these estimators is not second order, and so should not be ignored in calculating the standard error regardless of which estimate is adopted.

$\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_i)$ across Monte Carlo draws

$$\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T) = \frac{1}{200} \sum_{i=1}^{200} \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_i),$$

which affords a simple and natural means of evaluating expression (5).

To arrive at our expression for (4), the variance of this estimate $\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)$, let us begin by considering a slightly different magnitude:

$$\mathrm{E}\{ [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)]' | \boldsymbol{z}_T, \boldsymbol{\theta}_0 \}. \tag{6}$$

Thus, we tell our Bayesian econometrician the true value of 8, but ask him nonetheless to imagine estimating $\boldsymbol{x}_t$ without using this knowledge, and tell *us* how far off he would expect to be under this hypothetical exercise. Note that the average size of this error will in general be a function of the parameter vector $\boldsymbol{\theta}_0$. Next, rewrite (6) as

$$\begin{aligned} &\mathrm{E}\{ [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) + \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)] \\ &\quad \times [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) + \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)]' | \boldsymbol{z}_T, \boldsymbol{\theta}_0 \} \\ &= \mathrm{E}\{ [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)]' | \boldsymbol{z}_T, \boldsymbol{\theta}_0 \} \\ &\quad + \{ [\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)][\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)]' \}, \end{aligned} \tag{7}$$

where cross-product terms have disappeared because conditional on $\boldsymbol{z}_T$ and $\boldsymbol{\theta}_0$, the expression $[\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)]$ is non-stochastic and because

$$\mathrm{E}\{ [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}_0)] | \boldsymbol{z}_T, \boldsymbol{\theta}_0 \} = \mathbf{0}.$$

The desired answer is obtained by now integrating (6) and (7) with respect to the posterior density of $\boldsymbol{\theta}$:

$$\begin{aligned} &\mathrm{E}\{ [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)]' | \boldsymbol{z}_T \} \\ &= \mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{z}_T}\{ [\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta})][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta})]' | \boldsymbol{z}_T \} \\ &\quad + E_{\boldsymbol{\theta}|\boldsymbol{z}_T}\{ [\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)][\hat{\boldsymbol{x}}_t(\boldsymbol{z}_T, \boldsymbol{\theta}) - \hat{\boldsymbol{x}}_t(\boldsymbol{z}_T)]' | \boldsymbol{z}_T \}. \end{aligned} \tag{8}$$

The reader may recognize (8) as a special case of the familiar decomposition

$$\mathrm{var}(X) = \mathrm{E}\{\mathrm{var}(X|Y)\} + \mathrm{var}\{\mathrm{E}(X|Y)\}.$$

[See, for example, Lindgren (1976, p. 130).]

The sum of the two terms on the right-hand side of (8) gives us the total uncertainty associated with our estimate of the state vector x,. The first term,

$$\mathrm{E}_{\boldsymbol{\theta}|z_T}\{[\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta})][\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta})]' | z_T\}, \tag{9}$$

indicates how far off we would be on average even if we knew the true value of 8. I have elsewhere referred to this as the 'filter uncertainty' because it represents the uncertainty in separating signal and noise that is inherent in using the Kalman filter itself. This is closely related to the variance that is typically reported, which is obtained by simply replacing the true value $\boldsymbol{\theta}$ with the M.L.E. $\hat{\boldsymbol{\theta}}(z_T)$ in expression (3): $\hat{\boldsymbol{P}}_t(z_T, \hat{\boldsymbol{\theta}}(z_T))$. Comparing eqs. (3) and (8), it is clear that this conventional method will systematically understate the magnitude of the true error because it ignores the second term in eq. (8):

$$\mathrm{E}_{\boldsymbol{\theta}|z_T}\{[\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}) - \hat{\boldsymbol{x}}_t(z_T)][\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}) - \hat{\boldsymbol{x}}_t(z_T)]' | z_T\}. \tag{10}$$

I have referred to this second expression (10) as 'parameter uncertainty', because it summarizes how much better we could do by knowing the true parameter vector, $(\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}))$, than if we did not know $(\hat{\boldsymbol{x}}_t(z_T))$. Expression (8) thus shows that the total uncertainty can be neatly decomposed into the sum of the filter uncertainty and parameter uncertainty.

Here, then, is how we can construct the subjective variance of the estimated state vector $\hat{\boldsymbol{x}}_t(z_T)$. Generate the Monte Carlo series $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_{200})$ drawn from a $\mathrm{N}(\hat{\boldsymbol{\theta}}(z_T), \hat{\Lambda}(z_T))$ distribution and for each $\boldsymbol{\theta}_i$ calculate the sequences $\{\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}_i)\}$ and $\{\hat{\boldsymbol{P}}_t(z_T, \boldsymbol{\theta}_i)\}$ in the manner described earlier. Next, for each t, average the $\boldsymbol{P}_t(z_T, \boldsymbol{\theta}_i)$ across Monte Carlo draws

$$\frac{1}{200} \sum_{i=1}^{200} \hat{\boldsymbol{P}}_t(z_T, \boldsymbol{\theta}_i).$$

Recalling (3), this clearly is a natural Monte Carlo estimate of the contribution of the filter uncertainty, expression (9). Likewise, for each t calculate the variance of $\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}_i)$ across Monte Carlo draws

$$\frac{1}{200} \sum_{i=1}^{200} [\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}_i) - \hat{\boldsymbol{x}}_t(z_T)][\hat{\boldsymbol{x}}_t(z_T, \boldsymbol{\theta}_i) - \hat{\boldsymbol{x}}_t(z_T)]',$$

this being the logical estimate of the contribution of parameter uncertainty (10). Summing these two terms gives an expression for the total variance, which a Bayesian would use to characterize the posterior distribution for $\boldsymbol{x}_t$ and which a classical statistician could use to construct confidence intervals for the econometric estimate $\hat{\boldsymbol{x}}_t(z_T)$.

Table 1

Estimates of expected inflation, filter uncertainty, parameter uncertainty, and econometric standard error.[a]

Quarter	Expected inflation (1)	Filter uncertainty (2)	Parameter uncertainty (3)	Econometric standard error (4)
1951: I	1.6	20.00	0.00	4.47
1951: II	0.2	2.13	0.04	1.47
1951: III	4.4	1.24	0.25	1.22
1951: IV	1.8	0.66	0.23	0.94
1952: I	-0.7	0.36	0.25	0.78
1952: II	2.5	0.29	0.22	0.71
1952: III	2.1	0.23	0.15	0.62
1952: IV	2.3	0.20	0.25	0.67
1953: I	2.7	0.19	0.14	0.57
1953: II	2.0	0.18	0.10	0.53
1953: III	0.7	0.18	0.10	0.53
1953: IV	-0.4	0.17	0.14	0.56
1954: I	2.9	0.17	0.17	0.58
1954: II	1.6	0.17	0.20	0.61
1954: III	0.8	0.17	0.17	0.58
1954: IV	1.5	0.17	0.18	0.59
1955: I	2.1	0.17	0.16	0.57
1055: II	2.9	0.17	0.10	0.52
1955: III	3.0	0.17	0.10	0.52
1955: IV	2.2	0.17	0.05	0.47
1956: I	2.8	0.17	0.04	0.46
1956: II	3.5	0.17	0.04	0.46
1956: III	3.7	0.17	0.05	0.47
1956: IV	3.6	0.17	0.03	0.45
1957: I	3.8	0.17	0.03	0.44
1957: II	2.6	0.17	0.04	0.46
1957: III	3.5	0.17	0.08	0.49
1957: IV	1.6	0.17	0.09	0.51
1958: I	1.5	0.17	0.14	0.55
1958: II	1.2	0.17	0.08	0.50
1958: III	2.1	0.17	0.13	0.54
1958: IV	2.6	0.17	0.06	0.48
1959: I	2.9	0.17	0.09	0.51
1959: II	3.1	0.17	0.08	0.49
1959: III	2.3	0.17	0.03	0.44
1959: IV	2.6	0.17	0.05	0.47
1960: I	2.2	0.17	0.05	0.46
1960: II	1.6	0.17	0.06	0.48
1960: III	1.9	0.17	0.04	0.46
1960: IV	1.3	0.17	0.09	0.50
1961: I	0.8	0.17	0.04	0.46
1961: II	1.8	0.17	0.03	0.45
1961: III	1.9	0.17	0.07	0.49
1961: IV	1.7	0.17	0.05	0.47

Table 1 (continued)

Quarter	Expected inflation (1)	Filter uncertainty (2)	Parameter uncertainty (3)	Econometric standard error (4)
1962: I	3.0	0.17	0.05	0.46
1962: II	1.9	0.17	0.06	0.47
1962: III	2.0	0.17	0.06	0.47
1962: IV	2.2	0.17	0.04	0.46
1963: I	2.1	0.17	0.07	0.49
1963: II	1.7	0.17	0.04	0.45
1963: III	1.9	0.17	0.05	0.46
1963: IV	2.4	0.17	0.06	0.47
1964: I	2.3	0.17	0.04	0.45
1964: II	2.3	0.17	0.04	0.46
1964: III	2.1	0.17	0.04	0.45
1964: IV	1.9	0.17	0.03	0.44
1965: I	3.4	0.17	0.03	0.45
1965: II	2.6	0.17	0.05	0.46
1965: III	2.8	0.17	0.04	0.46
1965: IV	2.9	0.17	0.04	0.45
1966: I	3.8	0.17	0.04	0.46
1966: II	4.3	0.17	0.04	0.45
1966: III	3.6	0.17	0.05	0.47
1966: IV	4.0	0.17	0.04	0.45
1967: I	2.6	0.17	0.05	0.47
1967: II	2.5	0.17	0.07	0.49
1967: III	3.7	0.17	0.06	0.48
1967: IV	3.8	0.17	0.05	0.47
1968: I	5.0	0.17	0.05	0.47
1968: II	5.2	0.17	0.07	0.49
1968: III	3.6	0.17	0.08	0.49
1968: IV	5.2	0.17	0.06	0.48
1969: I	5.0	0.17	0.07	0.49
1969: II	5.0	0.17	0.14	0.55
1969: III	6.6	0.17	0.13	0.55
1969: IV	5.2	0.17	0.09	0.51
1970: I	6.5	0.17	0.07	0.49
1970: II	4.8	0.17	0.09	0.51
1970: III	4.4	0.17	0.06	0.48
1970: IV	4.8	0.17	0.05	0.47
1971: I	4.4	0.17	0.13	0.55
1971: II	5.4	0.17	0.10	0.52
1971: III	4.7	0.17	0.08	0.50
1971: IV	3.0	0.17	0.33	0.70
1972: I	4.2	0.17	0.09	0.51
1972: II	3.7	0.17	0.19	0.60
1972: III	3.5	0.17	0.22	0.63
1972: IV	4.7	0.17	0.10	0.52

Table 1 (continued)

Quarter	Expected inflation (1)	Filter uncertainty (2)	Parameter uncertainty (3)	Econometric standard error (4)
1973: I	5.1	0.17	0.17	0.57
1973: II	6.6	0.17	0.07	0.48
1973: III	7.4	0.17	0.13	0.54
1973: IV	6.5	0.17	0.30	0.68
1974: I	7.6	0.17	0.11	0.53
1974: II	8.8	0.17	0.19	0.60
1974: III	8.4	0.17	0.38	0.74
1974: IV	10.2	0.17	0.10	0.52
1975: I	8.1	0.17	0.21	0.61
1975: II	6.3	0.17	0.19	0.60
1975: III	6.3	0.17	0.14	0.55
1975: IV	5.6	0.17	0.10	0.52
1976: I	4.7	0.17	0.11	0.53
1976: II	4.4	0.17	0.08	0.49
1976: III	3.9	0.17	0.13	0.55
1976: IV	5.3	0.17	0.09	0.52
1977: I	5.8	0.17	0.09	0.50
1977: II	6.1	0.17	0.05	0.47
1977: III	5.4	0.17	0.06	0.48
1977: IV	5.7	0.17	0.05	0.47
1978: I	5.6	0.17	0.10	0.52
1978: II	7.7	0.17	0.11	0.53
1978: III	7.5	0.17	0.10	0.51
1978: IV	8.8	0.17	0.11	0.52
1979: I	7.4	0.17	0.15	0.57
1979: II	8.2	0.17	0.12	0.54
1979: III	7.9	0.17	0.06	0.48
1979: IV	9.2	0.17	0.33	0.70
1980: I	8.6	0.17	0.35	0.72
1980: II	9.1	0.17	0.21	0.62
1980: III	8.6	0.17	0.21	0.61
1980: IV	11.5	0.17	0.98	1.07
1981: I	10.4	0.17	0.38	0.74
1981: II	9.5	0.17	0.56	0.85
1981: III	9.2	0.17	0.52	0.83
1981: IV	7.7	0.17	0.41	0.76
1982: I	7.7	0.17	0.36	0.73

‘Source: Hamilton (1985).

4. Application

Table 1 presents the results of Hamilton (1985) in applying this technique to estimating the expectations of inflation on the part of the American public. These expectations were modelled as part of an unobserved state vector accounting for the dynamics of nominal interest rates. A very diffuse prior of $\hat{P}_0 = 20$ was used, and no smoother was employed. The filter uncertainty converges very quickly to a constant 0.17, corresponding to the steady-state values reached by $\hat{P}_t(z_t, \theta_i)$. By contrast, the parameter uncertainty varies considerably with the data. For some observations, such as the first quarter of 1965, the filter produces virtually the same guess $\hat{x}_t(z_t, \theta_i)$ for all plausible values for θ_i. For others, such as the fourth quarter of 1980, different parameter estimates lead to quite different inferences about the public's expectations. The square root of the sum of these two components, reported in column (4), averages around 50 basis points ($\pm 0.5\%$); that is, the econometrician might hope that the series in column (1) was typically within 100 basis points ($\pm$ 1%) of the true unobserved expectations of the public.

Summary statistics that ignored uncertainty about parameters would clearly have resulted in seriously misleading values for some quarters. Having a sensible measure of the total uncertainty with which the state vector is measured, as well as a break-down of where this uncertainty comes from, may prove a useful tool in future applications of the Kalman filter.

References

Anderson, Brian D.O. and John B. Moore, 1979, Optimal filtering (Prentice-Hall, Englewood Cliffs, NJ).

Burmeister, Edwin and Kent D. Wall, 1982, Kalman filtering estimation of unobserved rational expectations with an application to the German hyperinflation, Journal of Econometrics 20, 255-284.

Cooley, Thomas F. and Edward C. Prescott, 1973, Tests of an adaptive regression model, Review of Economics and Statistics 55, 248-256.

Cooley, Thomas F., Barr Rosenberg and Kent D. Wall, 1977, A note on optimal smoothing for time varying coefficient problems, Annals of Economic and Social Measurement 6, 453-456.

Engle, Robert F. and Mark W. Watson, 1981, A one-factor multivariate time series model of metropolitan wage rates, Journal of the American Statistical Association 76, 774-781.

Engle, Robert F. and Mark W. Watson, 1985, Applications of Kalman filtering in economics Mimeo., Oct. (University of California, San Diego, CA).

Engle, Robert F., David L. Lilien and Mark W. Watson, 1985, A DYMIMIC model of housing price determination, Journal of Econometrics 28, 307–326.

Fama, Eugene F. and Michael R. Gibbons, 1982, Inflation, real returns, and capital investment, Journal of Monetary Economics 9, 297–323.

Hamilton, James D., 1985, Uncovering financial market expectations of inflation, Journal of Political Economy 93, 1224–1241.

Howrey, E. Philip, 1984, Data revision, reconstruction, and prediction: An application to inventory investment, Review of Economics and Statistics 66, 386-393.

Lindgren, Bernard W.. 1976. Statistical theory, 3rd ed. (Macmillan, New York).

Ljung, Lennart, 1978, 'Convergence analysis of parametric identification methods, IEEE Transactions on Automatic Control AC-23, 770–783.

Ljung, Lennart and Peter E. Caines, 1979, Asymptotic normality of prediction error estimates for approximate system models, Stochastics 3, 29–46.
Meinhold, Richard J. and Nozer D. Singpurwalla, 1983, Understanding the Kalman filter, American Statistician 37, 123–127.
Nicholls, D.F. and B.G. Quinn, 1980, The estimation of random coefficient models: I, Journal of Time Series Analysis 1, 37–46.
Pagan, Adrian, 1980, Some identification and estimation results for regression models with stochastically varying coefficients, Journal of Econometrics 13, 341–363.
Rosenberg, Barr, 1973, The analysis of a cross section of time series by stochastically convergent parameter regression, Annals of Economic and Social Measurement 2, 399–428.
Sims, Christopher A., 1982, Policy analysis with econometric models, Brookings Papers on Economic Activity, 107–152.
Tanaka, Katsuto, 1983, Non-normality of the Lagrange multiplier statistic for testing the constancy of regression coefficients, Econometrica 51, 1577–1582.
Wall, Kent D., 1980, Generalized expectations modeling in macroeconometrics, Journal of Economic Dynamics and Control 2, 161–184.
Watson, Mark W., 1985, Recursive solution methods for dynamic linear rational expectations models, Discussion paper no. 1142, March (Harvard University, Cambridge, MA).
Watson, Mark W. and Robert F. Engle, 1983, Alternative algorithms for the estimation of dynamic factor, MIMIC, and varying coefficient regression models, Journal of Econometrics 23, 385–400.
Zellner, Arnold, 1971, An introduction to Bayesian inference in econometrics (Wiley, New York).