International Journal of Economics and Finance; Vol. 15, No. 2; 2023
ISSN 1916-971X E-ISSN 1916-9728
Published by Canadian Center of Science and Education

# Approximating an Infinite Horizon Model in the Presence of Optimal Experimentation

H. M. Amman[1] & M. P. Tucci[2]

[1] Amsterdam School of Economics, University of Amsterdam, the Netherlands

[2] Dipartimento di Economia Politica e Statistica, University of Siena, Italy

Correspondence: M. P. Tucci, Dipartimento di Economia Politica e Statistica, University of Siena, Piazza San Francesco 7, 53100 Siena, Italy.

Received: October 4, 2022 Accepted: November 30, 2022 Online Published: January 31, 2023

doi:10.5539/ijef.v15n2p70 URL: https://doi.org/10.5539/ijef.v15n2p70

**Abstract**

In an recent article Amman and Tucci (2020) make a comparison of the two dominant approaches for solving models with optimal experimentation in economics; the value function approach and an approximation approach. The approximation approach goes back to engineering literature in the 1970ties (cf. Tse & Bar-Shalom, 1973). Kendrick (1981) introduces this approach in economics. By using the same model and dataset as in Beck and Wieland (2002), Amman and Tucci conclude that differences may be small between the both approaches. In the previous paper we did not present the derivation of the approximation approach for this class of models. Hence, here we will present all derivations of the approximation approach for the case where there is an infinite horizon as is most common in economic models. By presenting the derivations, a better understanding and insight is obtained by the reader on how the value function is adequately approximated.

**Keywords:** optimal experimentation, approximation method, adaptive control, active learning, time-varying parameters, numerical experiments

## 1. Introduction

Recently there has been a renewed interest in *optimal experimentation*. In the engineering literature referred to as active learning, see e.g. Amman and Tucci (2020), Buera et al. (2011), Savin and Blueschke (2016). There are two dominant approaches for solving this class of models. The first method is based on the value function approach and the second on an approximation method. The former uses Bellman's (1957) dynamic programming approach for the closed loop (value function) form of the problem, which is used in studies by Prescott (1972), Taylor (1974), Easley and Kiefer (1988), Kiefer (1989), Kiefer and Nyarko (1989), Aghion et al. (1991), Beck and Wieland (2002), Coenen et al. (2005), Levin et al. (2003) and Wieland (2000) and many more.

In principle, the value function approach is theoretically the preferred method as it derives the optimal values for de policy variables. Unfortunately, it suffers from the *curse of dimensionality* and is only applicable to problems of low dimensionality due to the fact that the solution space needs to be discretized. The approximation methods as described in Cosimano (2008) and Cosimano and Gapen (2005), Kendrick (1981) and Hansen and Sargent (2007) use approaches, that are applied in the neighborhood of the linear regulator problems (Note 1). Because of this *local* nature with respect to the random elements of the approach, the method allows for models of larger dimension.

In Amman and Tucci (2020) both the value function approach and the approximation method are used to solve the same problem and their solutions are compared. For this purpose we used a common testbed model as presented in MacRae (1975) and Beck and Wieland (2002) (Note 2). In that paper the focus is on comparing the policy function results reported in Beck and Wieland (2002), through the value function, to those obtained through an approximation method. In this paper we present the full derivation of testbed model. In this way providing insight into the nature of the approximation approach.

## 2. Statement of the Problem

Tucci et al. (2010) consider a simple control problem with one state, one control and a time horizon of $T$ periods in which the policy maker wants to find $u_0, u_1, \ldots, u_{T-1}$ to minimize

$$J = E_0 \left\{ \frac{1}{2} w_T (x_T - \tilde{x}_T)^2 + \frac{1}{2} \sum_{t=0}^{T-1} \left[ w_t (x_t - \tilde{x}_t)^2 + \lambda_t (u_t - \tilde{u}_t)^2 \right] \right\} \quad (1)$$

where $E_0$ is the expectation operator conditional on the information available at time 0, subject to

$$x_{t+1} = \alpha x_t + \beta u_t + \gamma + \varepsilon_{t+1} \text{ for } t = 0, 1, ..., T-1 \quad (2)$$

with $x_t$ and $u_t$ the state and control variables, respectively, and the tilde indicating the desired path of the specified variable. Also $\alpha$, $\beta$ and $\gamma$ are the parameters of the system equation and $\varepsilon_{t+1}$ is an error term identically and independently distributed (i.i.d.) normal with mean zero and variance $q$. Finally, the initial state $x0$ and the penalty weights $w$'s and $\lambda$'s are given constants. The parameter associated with the control is assumed constant but unknown with mean, at time $t$, $b_t$ and variance $\sigma_{t|t}^{\beta\beta}$. Also, the state is measured without error (Note 3).

Following Tse and Bar-Shalom (1973) methods for solving active learning stochastic control problem, Tucci et al. (2010) compute, for each time period, the approximate cost-to-go at different values of the control and then choose that value which yields the minimum approximate cost (Note 4). This approximate cost-to-go is decomposed into three terms and, for the present problem, written as

$$J_N = J_{D,N} + J_{C,N} + J_{P,N} \quad (3)$$

where $J_N$ is the total cost-to-go with $N$ periods remaining and $J_{D,N}$, $J_{C,N}$ and $J_{P,N}$ are the deterministic, cautionary and probing component, respectively. The deterministic component includes only terms which are not stochastic. The cautionary one includes uncertainty only in the next time period and the probing term contains uncertainty in all future time periods. Thus the probing term includes the motivation to perturb the controls in the present time period in order to reduce future uncertainty about parameter values (Note 5).

In the following pages, this model is rewritten as an infinite horizon model and the associated formulae for the approximate cost-to-go are derived. The problem now is to find the set of controls $u_t$ for $t = 0, 1, ..., \infty$, where $t = 0$ denotes the current period, which minimizes the linear functional

$$J = E_0 \left\{ \frac{1}{2} \sum_{t=0}^{\infty} \left( x_t^2 w_t + u_t^2 \lambda_t \right) \right\} \quad (4)$$

with the desired path for the state and control set equal to 0, $x_t$ subject to the system equation (2) and $\lambda_t = \rho^t \lambda$ and $w_t = \rho^t w$ where ρ is the discount factor between 0 and 1.

The control problem (2) and (4) is solved treating the stochastic parameters as additional state variables as in Kendrick (1981; 2002, Chapter 10) and restating it in terms of an augmented state vector $z_t$ as: find the controls $u_t$ for $t = 0, 1, ..., \infty$ minimizing

$$J = E_0 \left\{ \frac{1}{2} \sum_{t=0}^{\infty} \left( z_t' W_t^* z_t + u_t^2 \lambda_t \right) \right\} \quad (5)$$

with $W_t^*$ having $w_t$ on the top left corner and zeros elsewhere. subject to the discrete-time system equations, with no measurement equation,

$$z_{t+1} = A^z z_t + \beta_t^z u_t + \gamma^z + \varepsilon_t^z \quad (6)$$

with the arrays defined as

$$z_t = \begin{bmatrix} x_t \\ \beta_t \end{bmatrix}, A^z = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}, \beta_t^z = \begin{bmatrix} \beta_t \\ 0 \end{bmatrix}, \gamma^z = \begin{bmatrix} \gamma \\ 0 \end{bmatrix}, \varepsilon_t^z = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix} \quad (7)$$

Problems (2) and (4) and (5)-(7) are equivalent "however the first is described as a linear quadratic problem with random coefficients and the second as a nonlinear (in $x$, $u$ and $\beta$) stochastic control problem" as noted in Kendrick (1981; 2002, p. 94).

## 3. One-Period Ahead Projection of the Mean and Variance of the Augmented State Vector Z

For this simple model the one-period ahead projection of the mean of the augmented state vector $z_t$, after control at time zero is applied, is

$$\hat{x}_{1|0} = \alpha x_0 + b_0 u_o^\tau + \gamma \tag{8}$$

$$b_{1|0} = b_0 \tag{9}$$

where $b_0$ is the estimate of the unknown parameter at time 0, with estimated variance $\sigma_{0|0}^{\beta\beta} \equiv \sigma_b^2$ to save on notation, $x_0$ is the initial condition for the state and $u_o^\tau$ being the search control at iteration $\tau$, with the Certainty Equivalence (*CE*) solution being the first search control, i.e. $u_o^1 \equiv u_o^{CE}$. The projected mean of the parameter is equal to its current estimate because the unknown parameter is assumed constant.

For the model presented in Beck and Wieland (2002) and MacRae (1975) (BWM) with no measurement error, the projected variances look like (Note 6)

$$\begin{aligned} \sigma_{1|0}^{xx} &= \left(u_0^\tau\right)^2 \sigma_{0|0}^{\beta\beta} + q \\ \sigma_{1|0}^{\beta x} &= \sigma_{0|0}^{\beta\beta} u_0^\tau \\ \sigma_{1|0}^{\beta\beta} &= \sigma_{0|0}^{\beta\beta} \equiv \sigma_b^2 \end{aligned} \tag{10}$$

## 4. The Nominal Path for the State and Control

At this point the nominal, or *CE*, path for state and control are needed. This is done by solving the *CE* problem for the un-augmented system from time 1 on, using $\hat{x}_{1|0}$ as initial condition and the nominal path for the parameters. Given that in the present case all of them are assumed constant, at this stage the estimate $b_0$ is treated as the true parameter for all future periods. Then the nominal control for a generic period $j$ in the time-horizon can be expressed as, in the present case,

$$u_{0,j} = G_j x_{0,j} + g_j \quad \text{for } j = 1, ..., \infty$$

When the conditions for the existence of an infinite horizon solution are satisfied, see e.g. De Koning (1982), Hansen and Sargent (2007, section 4.2.1), with $\lambda_j = \rho^j \lambda$ and $w_j = \rho^j w$, the optimal control law is time invariant, i.e.

$$G = -\left(\lambda + \rho k^{CE} b_0^2\right)^{-1} \alpha \rho k^{CE} b_0 \tag{11}$$

$$g = -\left(\lambda + \rho k^{CE} b_0^2\right)^{-1} b_0 \left(\rho k^{CE} \gamma + \rho p^{CE}\right) \tag{12}$$

with $k_{j+1}^{CE} = \rho k_j^{CE}$ and $p_{j+1}^{CE} = \rho p_j^{CE} \forall_j$, where $k^{CE}$ and $p^{CE}$ are the fixed point solutions to the usual Riccati recursions (Note 7)

$$k^{CE} = w + \alpha^2 \rho k^{CE} - \left(\alpha \rho k^{CE} b_0\right)^2 \left(\lambda + \rho k^{CE} b_0^2\right)^{-1} \tag{13}$$

and

$$p^{CE} = \alpha\left(\rho k^{CE}\gamma + \rho p^{CE}\right) - \alpha \rho k^{CE} b_0^2 \left(\lambda + \rho k^{CE} b_0^2\right)^{-1} \left(\rho k^{CE}\gamma + \rho p^{CE}\right) \tag{14}$$

respectively. Then $g$ can be rewritten as

$$g = G\alpha^{-1}\gamma(1 + \rho p^*) \tag{15}$$

with $p^* = [1 - \rho\,(\alpha + b_0 G)]^{-1}(\alpha + b_0 G)$. Generalizing the results in Tucci et al. (2010) it can be shown, by repeated substitutions, that in the infinite horizon problem the $j$-th nominal control can be written as the sum of two components (Appendix A). One associated with $\hat{x}_{1|0}$ depending upon the control applied at time 0, $u_0$, and the other due solely to the system parameters and exogenous forces, in this case the constant term $\gamma$. Namely

$$\begin{aligned} u_{0,j} &= G_{0,j} x_{0,j} + g_{0,j} \\ u_{0,j} &= G_{0,j} \hat{x}_{1|0} + g_{0,j} \end{aligned} \tag{16}$$

with

$$G_{0,j} = G(\alpha + b_0 G)^{j-1} \tag{17}$$

$$g_{0,j} = G\alpha^{-1}\gamma(\alpha + b_0 G + b_0 G\rho p^*)\sum_{i=1}^{j-1}(\alpha + b_0 G)^{i-1} + g \quad \text{for } j = 2,3,\ldots \tag{18}$$

and the nominal control at time *j* can be rewritten as

$$x_{0,j} = (\alpha + b_0 G)^{j-1}\hat{x}_{1|0} + \alpha^{-1}\gamma(\alpha + b_0 G + b_0 G\rho p^*)\sum_{i=1}^{j-1}(\alpha + b_0 G)^{i-1} \tag{19}$$

In the special case where γ = 0, the nominal state and control are simply

$$u_{0,j} = G_{0,j}x_{0,j} = G_{0,j}\hat{x}_{1|0} \tag{20}$$

and

$$x_{0,j} = (\alpha + b_0 G)^{j-1}\hat{x}_{1|0} \tag{21}$$

**5. Riccati Equations for the Arrays of the Augmented System**

The K and p Riccati arrays of the augmented system are partitioned as

$$K_j = \begin{bmatrix} k_j^{xx} & k_j^{x\beta} \\ k_j^{\beta x} & k_j^{\beta\beta} \end{bmatrix}, \qquad p_j = \begin{bmatrix} p_j^x \\ p_j^\beta \end{bmatrix} \tag{22}$$

In the former array, $k^{xx}$ matrix corresponds to the quantity $k^{CE}$ discussed in the previous section and when the condition for stabilization holds, i.e. $a + b_0 G$ is stable, and γ = 0 the quantities $k^{x\beta} = k^{\beta x}$ and $k^{\beta\beta}$ reduce to

$$k_j^{\beta x} = [\rho\,(\alpha + b_0 G)]^{j-1}\,k_1^{\beta x} \tag{23}$$

with

$$\begin{aligned} k_1^{\beta x} &= \rho k_1^{xx}(\alpha + b_0 G)\left[1 - \rho\,(\alpha + b_0 G)^2\right]^{-1} G x_{0,1} \\ &= \tilde{k}_1^{\beta x} x_{0,1} \end{aligned} \tag{24}$$

as shown in Appendix B and Appendix F, and

$$\begin{aligned} k_j^{\beta\beta} &= \rho\,(\alpha + b_0 G)^2 k_{j-1}^{\beta\beta} = \left[\rho\,(\alpha + b_0 G)^2\right]^{j-1} k_1^{\beta\beta} \\ &= \left[\rho\,(\alpha + b_0 G)^2\right]^{j-1} \tilde{k}_1^{\beta\beta} x_{0,1}^2 \end{aligned} \tag{25}$$

with

$$\begin{aligned} \tilde{k}_1^{\beta\beta} &= \rho k_1^{xx}\left[1 + \rho\,(\alpha + b_0 G)^2\right]\left[1 - \rho\,(\alpha + b_0 G)^2\right]^{-2} G^2 \\ &-(\rho k_1^{xx})^2\left[1 - \rho\,(\alpha + b_0 G)^2\right]^{-3}(\lambda_1 + \rho k_1^{xx} b_0^2)^{-1} b_0^2 G^2 \end{aligned} \tag{26}$$

as shown in Appendix C and Appendix F (Note 8). The elements of the p Riccati vector are defined as

$$p_j^x = k_j^{CE} x_{oj} + p_j^{CE} \tag{27}$$

and

$$\begin{aligned} p^\beta &= u_o p_{j+1}^x + p_{j+1}^\beta - \left[p_{j+1}^x + u_o k_{j+1}^{xx} b_0 + k_{j+1}^{\beta x} b_0\right] \\ &\times \left(\lambda + k_{j+1}^{xx} b_0^2\right)^{-1}\left(\lambda u_o + p_{j+1}^x b_0\right) \end{aligned} \tag{28}$$

with $k_j^{CE} = \rho^j k^{CE}$ and $p_j^{CE} = \rho^j p^{CE}$.

## 6. Updating the Covariances of the Augmented System

For the BWM problem the updating equations for the covariances of the augmented system look like (Note 9)

$$\Sigma_{j|j} = \begin{bmatrix} O & O \\ -\sigma^{\beta x}_{j|j-1}\left(\sigma^{xx}_{j|j-1}\right)^{-1} & 1 \end{bmatrix} \Sigma_{j|j-1} \tag{29}$$

then the elements of the updated covariance matrix are defined as

$$\sigma^{xx}_{j|j} = 0, \sigma^{x\beta}_{j|j} \equiv \sigma^{\beta x}_{j|j} = 0, \sigma^{\beta\beta}_{j|j} = \sigma^{\beta\beta}_{j|j-1} - \sigma^{\beta x}_{j|j-1}\left(\sigma^{xx}_{j|j-1}\right)^{-1}\sigma^{x\beta}_{j|j-1} \tag{30}$$

where the projected covariances take the form in (10) when *j* and *j-1* replace 1 and 0, respectively. Combining (30) and (10), it yields, for *j* = 1,

$$\sigma^{\beta\beta}_{1|1} = \sigma^{\beta\beta}_{1|0} - \sigma^{\beta x}_{1|0}\left(\sigma^{xx}_{1|0}\right)^{-1}\sigma^{x\beta}_{1|0} = \sigma^2_b q\left(u_0^2\sigma_b^2 + q\right)^{-1} \tag{31}$$

and in general it can be shown that (Appendix D).

$$\begin{aligned} \sigma^{\beta\beta}_{j|j} &= \sigma_b^2 q\left(\sigma_b^2 \sum_{i=0}^{j-1} u_{0,i}^2 + q\right)^{-1} \\ &= \sigma_b^2 q\left(\sigma_b^2 u_0^2 + q\right)^{-1}\left[1 + S\sum_{\substack{l=2 \\ for\, j\geq 2}}^{j}(\alpha + b_0 G)^{2(l-2)}\right]^{-1} \end{aligned} \tag{32}$$

with

$$S = G^2\left(\alpha x_0 + b_0 u_0\right)^2 \sigma^{\beta\beta}_{0|0}\left(u_0^2\sigma^{\beta\beta}_{0|0} + q\right)^{-1} \tag{33}$$

and

$$u_{0,0} \equiv u_0 \tag{34}$$

## 7. The Approximate Cost-to-Go

As in Kendrick (1981; 2002, Chapter 10) the approximate cost-to-go associated with the 'search' control $u^{\tau}$ is decomposed into three parts: deterministic $J_D$, cautionary $J_C$ and probing $J_P$. The deterministic component for the control at time 0 is, see, e.g., equation (10.49) in Kendrick (1981; 2002),

$$J_{D,T-1} = \frac{1}{2}\lambda_t u_t^2 + \frac{1}{2}k_T^{CE}x_{0,T}^2 + \frac{1}{2}\sum_{j=t+1}^{T-1}\left(k_j^{CE}x_{0,j}^2 + 2p_j^{CE}x_{0,j} + \lambda_j u_{0,j}^2\right) \tag{35}$$

For the model at hand, equation (35) can be rewritten as

$$J_{D,\infty} = \frac{1}{2}\lambda_0 u_0^2 + \frac{1}{2}\sum_{j=1}^{\infty}\left(k_j^{CE}x_{0,j}^2 + \lambda_j u_{0,j}^2\right) \tag{36}$$

where $k_j^{CE} \equiv k_j^{xx} = \rho^j k^{xx}$ and $\lambda_0 = \lambda$. Equation (36) can be written more compactly as

$$J_{D,\infty} = \psi_1 u_0^2 + \psi_2 u_0 + \psi_3 \tag{37}$$

The parameters in equation (37) simplify to

$$\begin{aligned} \psi_1 &= \frac{1}{2}\left\{\lambda + \rho b_0^2(k^{CE} + \lambda G^2)\left[1 - \rho(\alpha + b_0 G)^2\right]^{-1}\right\} \\ \psi_2 &= \rho(k^{CE} + \lambda G^2)\left[1 - \rho(\alpha + b_0 G)^2\right]^{-1} b_0 \alpha x_0 \\ \psi_3 &= \frac{1}{2}\rho(k^{CE} + \lambda G^2)\left[1 - \rho(\alpha + b_0 G)^2\right]^{-1}(\alpha x_0)^2 \end{aligned} \tag{38}$$

when there is no constant term and zero desired path for the state and control (Appendix E). The cautionary component looks like

$$J_{C,\infty} = \frac{1}{2}\left[k_1^{xx}\left(\sigma_b^2 u_0^2 + q\right) + k_1^{\beta\beta}\sigma_b^2\right] + k_1^{x\beta}\sigma_b^2 u_0 + \frac{1}{2}\sum_{j=1}^{\infty}\left(\rho^j k_1^{xx} q\right) \tag{39}$$

By using the definitions of the $k$'s and rearranging the terms it yields

$$J_{C,\infty} = \delta_1 u_0^2 + \delta_2 u_0 + \delta_3 \tag{40}$$

with

$$\begin{aligned} \delta_1 &= \frac{1}{2}\sigma_b^2 \left(k_1^{xx} + \tilde{k}_1^{\beta\beta} b_0^2 + 2\tilde{k}_1^{\beta x} b_0\right) \\ \delta_2 &= \sigma_b^2 \left(\tilde{k}_1^{\beta\beta} b_0 + \tilde{k}_1^{\beta x}\right) \alpha x_0 \\ \delta_3 &= \frac{1}{2} k_1^{xx} q (1-\rho)^{-1} + \frac{1}{2}\sigma_b^2 \tilde{k}_1^{\beta\beta} \alpha^2 x_0^2 \end{aligned} \tag{41}$$

as apparent from Appendix F, when the identity $\sigma_{0|0}^{\beta\beta} \equiv \sigma_b^2$ is used. Finally, the probing component takes the form

$$J_{P,\infty} = \frac{1}{2} \sum_{j=1}^{\infty} \left[p_{j+1}^x + u_o \rho^j k_1^{xx} b_0 + k_{j+1}^{\beta x} b_0\right]^2 \left[\rho^j \left(\lambda_0 + k_1^{xx} b_0^2\right)\right]^{-1} \sigma_{j|j}^{\beta\beta} \tag{42}$$

Similarly to Amman and Kendrick (1995) and Tucci et al. (2010), equation (42) can be rewritten as

$$J_{P,\infty} = \frac{1}{2} \frac{g(u_0)}{h(u_0)} \tag{43}$$

with

$$h(u_0) = \left(u_0^2 \sigma_b^2 + q\right) \left(\sigma_b^2 q\right)^{-1} \tag{44}$$

and

$$g(u_0) = \phi_1 \left(\phi_2 u_0 + \phi_3\right)^2 \tag{45}$$

with

$$\begin{aligned} \phi_1 &= \left[(\rho k_1^{xx})^2 \left(\lambda + \rho k_1^{xx} b_0^2\right)^{-1}\right] \left[1 - \rho \left(\alpha + b_0 G\right)^2\right]^{-4} b_0^2 G^2 \\ \phi_2 &= b_0 \\ \phi_3 &= \alpha x_0 \end{aligned} \tag{46}$$

as shown in Appendix G. At this point by substituting (37), (40) and (43) into (35) yields

$$\begin{aligned} J_\infty &= (\psi_1 + \delta_1) u_0^2 + (\psi_2 + \delta_2) u_0 + (\psi_3 + \delta_3) + \\ &\quad \left(\frac{\sigma_b^2 q}{2}\right) \frac{\phi_1 (\phi_2 u_0 + \phi_3)^2}{(\sigma_b^2 u_0^2 + q)} \end{aligned} \tag{47}$$

with the parameters defined as in (38), (41) and (46). As shown in Appendix H through Appendix J, these new definitions are perfectly consistent with those associated to the two-period finite horizon model reported in Amman and Kendrick (1995) and Tucci et al. (2010).

**8. Numerical Example**

In this section the DUAL infinite horizon control is computed using the parameter set in Beck and Wieland (2002, Figure 1, p. 1367) which translates to

$$\alpha = 1, b_0 = -0.5, \gamma = 0, q = 1, \sigma_{0|0}^{\beta\beta} = \sigma_b^2 = 0.25, w = 1, \lambda = 0, \rho = 0.95 \tag{48}$$

in the present context. With this parameter set, the fixed point solution to the usual Riccati recursions for the unaugmented system is

$$\begin{aligned} k^{CE} &= 1 + \rho k^{CE} - 0.25 \left(\rho k^{CE}\right)^2 \left(0.25 \rho k^{CE}\right)^{-1} \\ &= 1 + \rho k^{CE} - \rho k^{CE} = 1 \end{aligned} \tag{49}$$

with $\rho k^{CE} \equiv \rho k^{xx} \equiv k_1^{xx} = 0.95$ and the time invariant optimal control law simplifies to

$$G = -\left(0.25 \rho k^{CE}\right)^{-1} \rho k^{CE} (-0.5) = 2 \tag{50}$$

It follows that the relevant terms for the computation of the approximate cost-to-go described in the previous section 9 specialize to

$$(\alpha + b_0 G) = 1 + 2(-0.5) = 0 \tag{51}$$

$$
\begin{aligned}
\rho k_1^{xx} &= \rho(\rho k^{xx}) = \rho^2 k^{xx} = (0.95)^2 \\
\tilde{k}_1^{\beta x} &= (0.95)^2(0)[1-(0.95)(0)^2]^{-1}2 = 0 \\
\tilde{k}_1^{\beta\beta} &= (0.95)^2(1)(1)^{-2}2^2 - 0.95(0.95)^2(1)^{-3}[0.25(0.95)]^{-1}(0.25)\,2^2 = 0
\end{aligned}
\tag{52}
$$

Then the coefficients characterizing the deterministic, cautionary and probing component are, respectively,

$$
\begin{aligned}
\psi_1 &= \frac{1}{2}(0.25)0.95 = 0.119 \\
\psi_2 &= 0.95\,(-0.5)\,x_0 = -0.475x_0 \\
\psi_3 &= \frac{1}{2}(0.95)x_0^2 = 0.475x_0^2
\end{aligned}
\tag{53}
$$

$$
\begin{aligned}
\delta_1 &= \frac{1}{2}0.25(0.95) = 0.119 \\
\delta_2 &= 0 \\
\delta_3 &= \frac{1}{2}(0.95)(1)(0.05)^{-1} = 9.5
\end{aligned}
\tag{54}
$$

and

$$
\begin{aligned}
\phi_1 &= 0.95\,(0.95)^2\,(0.25*0.95)^{-1}\,(1)^{-4}(0.25*4) = 0.95^2*4 \\
\phi_2 &= -0.5 \\
\phi_3 &= x_0
\end{aligned}
\tag{55}
$$

By comparing the new results with those associated with a two-period model reported in Tucci et al. (2010, equations 34-39) some interesting features emerge. First of all the ψ's in the deterministic component are the same both in the finite and infinite model except for the fact that the former uses undiscounted penalty weights on the state, i.e. $w_1 = w_2 = 1$, and the latter assumes $w_t = \rho^t w$ with $w = 1$. The same consideration explains the slight difference existing between the new and old coefficient $\delta_1$ in the cautionary component and $\varphi_1$ in the probing one. It is noteworthy that the coefficient $\delta_2$ in the cautionary component and $\varphi_2$ and $\varphi_3$, in absolute value, in the probing one are identical in the finite and infinite model. This means that these coefficients are not affected by the penalty weight on the state. The main difference between the finite and infinite model lies in $\delta_3$, the constant term in the cautionary component, which jumps from 1, the variance of the system disturbance, to 9.5 which is, approximately, half the inverse of the discount rate, i.e. $\frac{1}{2}(1-\rho)^{-1}$. Therefore this coefficient reflects the infinite sum of the discount factor $\rho$.

The results for the $J_D$, $J_C$, $J_P$ and $J_\infty$, using the above parameters, are plotted in figure 1, which is for certain levels of the parameters not globally convex. As mentioned earlier, Amman and Kendrick (1995), the solution of the may suffere from non-convexities and can have several (local) minima for certain parameter sets. Figure 2 shows clearly that when the uncertainty of the parameter $\beta$, $\sigma_b^2$ increase the chance of hitting a multiple minima increases. Hence, when doing a numerical optimization with the model, caution is required.

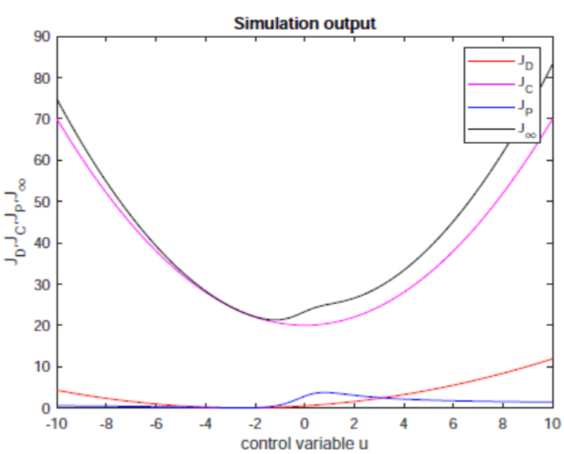


Figure 1. Plotting $J_D$, $J_C$, $J_P$, $J_\infty$, $\sigma_b^2 = 1.00$

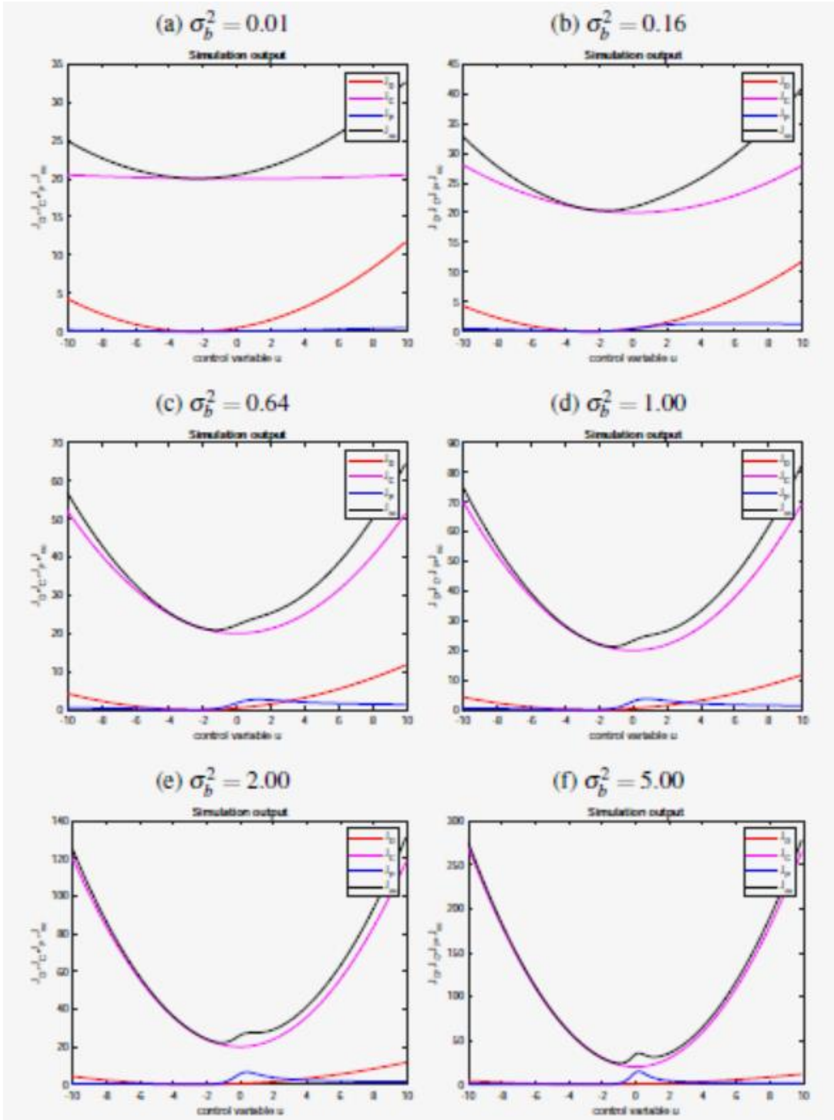

Figure 2. Plotting $J_D$, $J_C$, $J_P$, $J_\infty$ for various $\sigma_b^2$

## 9. Conclusion

By applying a well-known testbed model, we presented the full derivation of an (value function) approximation approach, in this way providing insight into the nature of the approximation. The appropriate Riccati quantities for the augmented system have been derived and the time-invariant feedback rule defined. The resulting formulas are easy to compute and allow for problems of higher dimensions that can be solved in feasable time. Due to the local nature of the approximation, caution is required when the model suffers from a high degree of stochasticity as defined by the various (co)variance of in the model. With high levels of randomness, the solution may produce multiple local optima.

### Acknowledgement

We greatfully acknowledge the open-access publication fee support for this article by the University of Siena (Università di Siena), Siena, Italy.

## References

Aghion, P. et al. (1991). Optimal Learning by Experimentation. *Review of Economic Studies*, *58*, 621-654. https://doi.org/10.2307/2297825

Amman, H. M. (1996). Numerical methods for linear-quadratic models. In H. M. Amman, D. A. Kendrick, & J. Rust (Eds.), *Handbook of Computational Economics* (Vol. 13, pp. 579-618). Amsterdam, the Netherlands: North-Holland Publishers (Elsevier). https://doi.org/10.1016/S1574-0021(96)01015-5

Amman, H. M., & Kendrick, D. A. (1995). Nonconvexities in Stochastic Control Models. *International Economic Review, 36*, 455-475. https://doi.org/10.2307/2527206

Amman, H. M., & Neudecker, H. (1997). Numerical solution methods of the algebraic matrix Riccati equation. *Journal of Economic Dynamics and Control, 21*, 363-370. https://doi.org/10.1016/S0165-1889(96)00936-0

Amman, H. M., Kendrick, D. A., & Tucci, M. P. (2020). Approximating the Value Function for Optimal Eexperimentation. *Macroeconomic Dynamics, 24*(5). https://doi.org/10.1017/S1365100518000664

Bar-Shalom, Y., & Sivan, R. (1969). On the Optimal Control of Discrete-Time Linear Systems with Random Parameters. *IEEE Transactions on Automatic Control, 14*(1), 3-8. https://doi.org/10.1109/TAC.1969.1099087

Beck, G. W., & Wieland, V. (2002). Learning and control in a changing economic environment. *Journal of Economic Dynamics and Control, 26*(9), 1359-1377. https://doi.org/10.1016/S0165-1889(01)00075-6

Bellman, R. E. (1957). *Dynamic Programming*. Princeton, New Jersey: Princeton University Press.

Buera, F. J., Monge-Naranjo, A., & Primiceri, G. E. (2011). Learning the wealth of nations. *Econometrica, 79*(1), 1-45. https://doi.org/10.3982/ECTA8299

Coenen, G., Levin, A., & Wieland, V. (2005). Data Uncertainty and the Role of Money as an Information Variable for Monetary Policy. *European Economic Review, 49*, 975-1006. https://doi.org/10.1016/j.euroecorev.2003.08.005

Cosimano, T. F. (2008). Optimal experimentation and the perturbation method in the neighborhood of the augmented linear regulator problem. *Journal of Economics, Dynamics and Control, 32*, 1857-1894. https://doi.org/10.1016/j.jedc.2007.03.009

Cosimano, T. F., & Gapen, M. T. (2005a). *Program Notes for Optimal Experimentation and the Perturbation Method in the Neighborhood of the Augmented Linear Regulator Problem*. Working Paper. Notre Dame, Indiana, USA: Department of Finance, University of Notre Dame.

Cosimano, T. F., & Gapen, M. T. (2005b). *Recursive Methods of Dynamic Linear Economics and Optimal Experimentation using the Perturbation Method*. Working Paper, Notre Dame, Indiana, USA: Department of Finance, University of Notre Dame.

Easley, D., & Kiefer, N. M. (Sept. 1988). Controlling a Stochastic Process with Unknown Parameters. *Econometrica, 56*(5), 1045-1064. https://doi.org/10.2307/1911358

Hansen, L. P., & Sargent, T. J. (2007). *Robustness*. Princeton University Press. https://doi.org/10.1515/9781400829385

Kendrick, D. A. (1981). *Stochastic Control for Economic Models*. New York, New York: McGraw-Hill.

Kendrick, D. A. (2002). *Stochastic control for economic models* (online edition). Retrieved from http://www.utexas.edu/cola/depts/economics/faculty/dak2

Kiefer, N. (1989). A Value Function Arising in the Economics of Information. *Journal of Economic Dynamics and Control, 13*, 201-223. https://doi.org/10.1016/0165-1889(89)90018-3

Kiefer, N., & Nyarko, Y. (1989). Optimal Control of an Unknown Linear Process with Learning. *International Economic Review, 30*, 571-586. https://doi.org/10.2307/2526776

Koning, W. L. (1982). Infinite horizon optimal control of linear discrete time systems with stochastic parameters. *Automatica,* 18(4), 443-453. https://doi.org/10.1016/0005-1098(82)90072-3

Levin, A., Wieland, V., & Williams, J. C. (June 2003). The Performance of Forecast-Based Monetary Policy Rules Under Model Uncertainty. *American Economic Review, 93*(3), 622-645. https://doi.org/10.1257/000282803322157016

MacRae, E. C. (1975). An Adaptive Learning Role for Multi-period Decision Problems. *Econometrica, 43*, 893-906. https://doi.org/10.2307/1911332

Prescott, E. C. (1972). The Multi-period Control Problem under Uncertainty. *Econometrica, 40*, 1043-1058. https://doi.org/10.2307/1913854

Savin, I., & Blueschke, D. (2016). Lost in Translation: Explicitly Solving Nonlinear Stochastic Optimal Control Problems Using the Median Objective Value. *Computational Economics, 48*(2), 317-338. https://doi.org/10.1007/s10614-015-9526-3

Taylor, J. B. (1974). Asymptotic Properties of Multi-period Control Rules in the Linear Regression Model. *International Economic Review, 15*(2), 472-482. https://doi.org/10.2307/2525872

Tse, E. (1973). Further comments on Adaptive stochastic control for a class of linear systems. *IEEE Transactions on Automatic Control, 18*, 324-326. https://doi.org/10.1109/TAC.1973.1100320

Tse, E., & Bar-Shalom, Y. (1973). An actively adaptive control for linear systems with random parameters via the dual control approach. *IEEE Transactions on Automatic Control, 18*(2), 109-117. https://doi.org/10.1109/TAC.1973.1100242

Tucci, M. P. (2004). *The Rational Expectation Hypothesis, Time-varying Parameters and Adaptive Control*. Dordrecht, the Netherlands: Springer. https://doi.org/10.1007/978-1-4020-2874-8

Tucci, M. P., Kendrick, D. A., & Amman, H. M. (2010). The Parameter Set in an Adaptive Control Monte Carlo Experiment: Some Considerations. *Journal of Economic Dynamics and Control, 34*(9), 1531-1549. https://doi.org/10.1016/j.jedc.2010.06.014

Wieland, V. (2000a). Learning by Doing and the Value of Optimal Experimentation. *Journal of Economic Dynamics and Control, 24*, 501-543. https://doi.org/10.1016/S0165-1889(99)00015-9

Wieland, V. (2000b). Monetary Policy, Parameter Uncertainty and Optimal Learning. *Journal of Monetary Economics, 46*, 199-228. https://doi.org/10.1016/S0304-3932(00)00023-4

## Notes

Note 1. For consistency and clarity in the main text, we used the term approximation method instead of adaptive or dual control. The adaptive or dual control approach in MacRae (1975), see Kendrick (1981), Amman (1996) and Tucci (2004), uses methods that draw on earlier work in the engineering literature by Bar-Shalom and Sivan (1969) and Tse (1973).There are differences between this approach and the approximation approaches in Cosimano (2008) and Savin and Blueschke (2016) which we will not discuss in detail here. Through out the paper we will use the approach in Kendrick (1981).

Note 2. Throughout the paper we will use the abbreviation BWM for the testbed model

Note 3. This is equivalent to setting H=I and R=O in Kendrick (1981; 2002, Chapter 10 -11) or Tucci (2004, chapter 2-5).

Note 4. See Kendrick (1981; 2002, Chapter 9-10) or Tucci (2004, chapter 2) for details.

Note 5. See Kendrick (1981; 2002, pp. 97-98) for an introduction to this decomposition.

Note 6. See, e.g., Kendrick (1981; 2002, Chapter 10, p. 102) or Tucci (2004, chapter 2, pp. 21-22) for details.

Note 7. In this case the Riccati equation is scalar function and can easily be solved. The multi-dimensional case can be more complicated to solve. See Amman and Neudecker (1997).

Note 8. This compares with $k_1^{\beta x} = 2w_2(\alpha + bG_1)G_1x_{0,1}$ and $k_1^{\beta\beta} = w_2G_1^2x_{0,1}^2 + w_2^2(\alpha + 2bG_1)^2\left[-\left(\lambda_1 + b^2w_2\right)\right]^{-1}x_{0,1}^2$ in the two-period finite horizon model.

Note 9. See, e.g., Kendrick (1981; 2002, Chapter 10, p. 103) or Tucci (2004, chapter 2, pp. 27-28) for details.

## Appendix A

### Deriving the nominal path for control as a function of the projected state

Given a certain control at time 0, say $u_0$, the nominal, or Certainty Equivalence (CE), value of $x_1$, denoted by $x_{0,1}$, is given by

$$x_{0,1} = \alpha x_0 + \beta u_0 + \gamma$$

when the system parameters are assumed constant and known. Then the nominal or *CE* value of $u_1$, $u_{0,1}$, in a two-period control problem is given by

$$\begin{aligned} u_{0,1} &= G_1x_{0,1} + g_1 \\ &= \left(-\frac{1}{\lambda_1 + \beta^2 w_2}\right)\left[\alpha\beta w_2 x_{0,1} + \beta w_2\left(\gamma - \tilde{x}_2\right) - \lambda_1\tilde{u}_1\right] \end{aligned} \tag{A-1}$$

where $w_2$ is the penalty on the state in the final period and the tilde stands for desired path. When the desired path for the state and control is zero, the above formula simplifies to

$$\begin{aligned} u_{0,1} &= G_1 x_{0,1} + g_1 \\ &= \left(-\frac{\alpha\beta k_2}{\lambda_1+\beta^2 k_2}\right) x_{0,1} + \left(-\frac{1}{\lambda_1+\beta^2 k_2}\right)\beta\left(k_2\gamma + p_2\right) \end{aligned} \tag{A-2}$$

with $G_1$ and $g_1$ implicitly defined, and $k_2$ and $p_2$ the appropriate Riccati quantities, for any finite period control problem. The associated nominal value of $x2$ is

$$\begin{aligned} x_{0,2} &= \alpha x_{0,1} + \beta u_{0,1} + \gamma x_{0,2} \\ &= (\alpha + \beta G_1) x_{0,1} + \beta g_1 + \gamma \end{aligned} \tag{A-3}$$

Then the nominal control for the finite horizon problem at time 2 can be written as

$$\begin{aligned} u_{0,2} &= G_2 x_{0,2} + g_2 \\ &= G_2(\alpha + \beta G_1) x_{0,1} + G_2 \alpha^{-1}(\alpha + \beta G_1 + 1)\gamma \\ &\quad + \alpha^{-1} G_2 \left(\beta G_1 k_2^{-1} p_2 + k_3^{-1} p_3\right) \end{aligned} \tag{A-4}$$

with $g_2$ defined similarly to $g_1$. By repeating this procedure, it is then apparent that the nominal control at any time $j$ in the planning horizon can be rewritten as the sum of two components. One associated with $x_{0,1}$ depending upon the control applied at time 0, $u_0$, and the other due solely to the system parameters and exogenous forces, in this case the constant term $\gamma$. Namely,

$$u_{0,j} = G_j x_{0,j} + g_j = G_{0,j} x_{0,1} + g_{0,j} \tag{A-5}$$

with

$$G_{0,j} = G_j \left[\prod_{i=1}^{j-1} (\alpha + \beta G_i)\right] \tag{A-6}$$

$$\begin{aligned} g_{0,j} &= \alpha^{-1} G_j \gamma \sum_{i=1}^{j} \left[\prod_{l=i}^{j-1} (\alpha + \beta G_l)\right] + \\ &\quad \alpha^{-1} G_j \left\{ k_{j+1}^{-1} p_{j+1} + \sum_{i=1}^{j-1} \left[\prod_{l=i+1}^{j-1} (\alpha + \beta G_l)\right] \beta G_i k_{i+1}^{-1} p_{i+1} \right\} \end{aligned} \tag{A-7}$$

where it is implied that the product term in square brackets is one when $l > j-1$ and the feedback quantities $G_j$ and $g_j$ are defined as

$$\begin{aligned} G_j &= -\left(\lambda_j + k_{j+1}\beta^2\right)^{-1} \alpha k_{j+1} \beta \\ g_j &= -\left(\lambda_j + k_{j+1}\beta^2\right)^{-1} \beta \left(k_{j+1}\gamma + p_{j+1}\right) \end{aligned} \tag{A-8}$$

The associated nominal state at time $j$ can obviously be written as

$$\begin{aligned} x_{0,j} &= \left[\prod_{i=1}^{j-1} (\alpha + \beta G_i)\right] x_{0,1} + \alpha^{-1}\gamma \sum_{i=1}^{j-1} \left[\prod_{l=i}^{j-1} (\alpha + \beta G_l)\right] \\ &\quad + \alpha^{-1} \sum_{i=1}^{j-1} \left[\prod_{l=i+1}^{j-1} (\alpha + \beta G_l)\right] \beta G_i k_{i+1}^{-1} p_{i+1} \end{aligned} \tag{A-9}$$

with all symbols as previously defined. When the conditions for the existence of an infinite horizon solution are satisfied, see e.g. De Koning (1982), Hansen and Sargent (2007), with $\lambda_j = \rho^j \lambda$ and $w_j = \rho^j w$, the optimal control law is time invariant, i.e. the quantities in (A-8) specialize to

$$G = -\left[\left(\lambda + \rho k \beta^2\right)\right]^{-1} \alpha \rho k \beta \tag{A-10}$$

$$g = -\left(\lambda + \rho k \beta^2\right)^{-1} \beta \left(\rho k \gamma + \rho p\right) \tag{A-11}$$

with $k_{j+1} = \rho k_j$ and $p_{j+1} = \rho p_j \forall_j$, where $k$ and $p$ are the fixed point solutions to the usual Riccati recursions

$$k \equiv k^{CE} = w + \alpha^2 \rho k - (\alpha \rho k \beta)^2 \left(\lambda + \rho k \beta^2\right)^{-1} \tag{A-12}$$

and

$$p \equiv p^{CE} = \alpha\left(\rho k \gamma + \rho p\right) - \beta \rho k \alpha \left(\lambda + \rho k \beta^2\right)^{-1} \beta \left(\rho k \gamma + \rho p\right) \tag{A-13}$$

respectively. Then equation (A-11) can be rewritten as

$$g = G\alpha^{-1}\gamma(1+\rho p^*) \tag{A-14}$$

with

$$p^* = [1-\rho(\alpha+\beta G)]^{-1}(\alpha+\beta G) \tag{A-15}$$

In the infinite horizon model the above formulae (A-5) and (A-9) simplify as follows

$$u_{0,j} = Gx_{0,j} + g = G_{0,j}x_{0,1} + g_{0,j} \quad \text{for} \quad j = 1,2,\ldots \tag{A-16}$$

$$x_{0,j} = G^*_{0,j}x_{0,1} + g^*_{0,j} \quad \text{for} \quad j = 2,3,\ldots \tag{A-17}$$

with

$$G_{0,j} = G(\alpha+\beta G)^{j-1} = GG^*_{0,j} \quad \text{for} \quad j = 1,2,\ldots \tag{A-18}$$

$$g_{0,j} = Gg^*_{0,j} + g \quad \text{for} \quad j = 2,3,\ldots \tag{A-19}$$

where

$$\begin{aligned} g^*_{0,j} &= \alpha^{-1}\gamma\sum_{i=1}^{j-1}(\alpha+\beta G)^i + \alpha^{-1}\gamma\sum_{i=1}^{j-1}(\alpha+\beta G)^{i-1}\beta G\rho p^* \\ &= \alpha^{-1}\gamma(\alpha+\beta G+\beta G\rho p^*)\sum_{i=1}^{j-1}(\alpha+\beta G)^{i-1} \\ &\quad \text{for} \quad j = 2,3,\ldots \end{aligned} \tag{A-20}$$

It is important to notice that when there is no exogenous variable or intercept, and the desired path for the state and control are zero as asssumed here, the $g$ terms disappear and the nominal control and state are simply

$$u_{0,j} = G(\alpha+\beta G)^{j-1}x_{0,1} \tag{A-21}$$

$$\begin{aligned} x_{0,j} &= (\alpha+\beta G)^{j-1}x_{0,1} \\ &\quad \text{for} \quad j = 2,3,\ldots \end{aligned} \tag{A-22}$$

**Appendix B**

**Deriving submatrix $k^{\beta x}$ of the augmented system in the infinite horizon model**

In the BWM model, when the unknown parameter $\beta$ is replaced by its estimate at time 0, $b_0$, the general formula for $k^{\beta x}$, see e.g. Kendrick (1981; 2002, equation (10.40)) or Tucci (2004, equation 2.56), specializes to

$$\begin{aligned} k_1^{\beta x} &= u_{0,1}k_2^{xx}\alpha + k_2^{\beta x}\alpha - \left(p_2^x + u_{0,1}k_2^{xx}b_0 + k_2^{\beta x}b_0\right) \\ &\quad \times \left(\lambda_1 + k_2^{xx}b_0^2\right)^{-1}\alpha k_2^{xx}b_0 \\ &= \rho k_1^{xx}(\alpha+b_0G)u_{0,1} + k_2^{\beta x}(\alpha+b_0G) + p_2^xG \end{aligned} \tag{B-1}$$

with

$$p_j^x = k_j^{xx}x_{0,j} + p_j^{CE} \tag{B-2}$$

In the infinite horizon model, see, e.g., equation (A-13) in Appendix A,

$$p^{CE} = [1-\rho(\alpha+Gb)]^{-1}(\alpha+Gb)\rho k^{CE}\gamma = p^*\rho k^{CE}\gamma \tag{B-3}$$

Then it follows that

$$\begin{aligned} p_2^x &= k_2^{xx}x_{0,2} + p_2^{CE} \\ &= \rho k_1^{xx}x_{0,2} + \rho p_1^{CE} \\ &= \rho k_1^{xx}(\alpha+b_0G)x_{0,1} + c_2^p \end{aligned} \tag{B-4}$$

where

$$c_2^p = \rho k_1^{xx}(\alpha+b_0G)\alpha^{-1}\gamma(1+\rho p^*) \tag{B-5}$$

Therefore

$$\begin{aligned} p_2^xG &= \left[\rho k_1^{xx}(\alpha+b_0G)x_{0,1} + \rho k_1^{xx}(\alpha+b_0G)\alpha^{-1}\gamma(1+\rho p^*)\right]G \\ &= \rho k_1^{xx}(\alpha+b_0G)(Gx_{0,1}+g) \end{aligned} \tag{B-6}$$

with $G$ and $g$ as in equations (A-10)-(A-11) in Appendix A. Then $k^{\beta x}$ can be rewritten as

$$k_1^{\beta x} = 2\rho k_1^{xx}(\alpha + b_0 G)(Gx_{0,1} + g) + k_2^{\beta x}(\alpha + b_0 G) \tag{B-7}$$

with

$$k_2^{\beta x} = 2\rho^2 k_1^{xx}(\alpha + b_0 G)(Gx_{0,2} + g) + k_3^{\beta x}(\alpha + b_0 G) \tag{B-8}$$

Then, by repeated substitution, it can be shown that

$$\begin{aligned} k_1^{\beta x} &= 2\rho k_1^{xx}(\alpha + b_0 G) u_{0,1} + \\ &\quad (\alpha + b_0 G)\left[2\rho^2 k_1^{xx}(\alpha + b_0 G) u_{0,2} + (\alpha + b_0 G) k_3^{\beta x}\right] \\ &= 2\rho k_1^{xx}(\alpha + b_0 G) u_{0,1} + 2\rho^2 k_1^{xx}(\alpha + b_0 G)^2 u_{0,2} + \dots \\ &= 2\sum_{j=1}^{\infty} \rho^j k_1^{xx}(\alpha + b_0 G)^j u_{0,j} \end{aligned} \tag{B-9}$$

By using equation (A-14) in Appendix A for the nominal control, it follows that $k^{\beta x}$ can be viewed as the sum of two components, one dependent uponthe control applied at time 0, $u_0$, and the other due solely to the system parameters and exogenous forces, in this case the constant term γ. Namely,

$$k_1^{\beta x} = k_1^{\beta x}(x_{0,1}) + c_1^{\beta x} \tag{B-10}$$

with

$$k_1^{\beta x}(x_{0,1}) = 2\sum_{j=1}^{\infty} \rho^j k_1^{xx}(\alpha + b_0 G)^j G_{0,j} x_{0,1} \tag{B-11}$$

$$c_1^{\beta x} = 2\sum_{j=1}^{\infty} \rho^j k_1^{xx}(\alpha + b_0 G)^j g_{0,j} \tag{B-12}$$

Replacing the definition of $G_{0,j}$, i.e. equation (A-18) in Appendix A, into (B-11) yields

$$\begin{aligned} k_1^{\beta x}(x_{0,1}) &= 2\sum_{j=1}^{\infty} (\alpha + b_0 G)^{j-1}(\alpha + b_0 G)^j \rho^j k_1^{xx} G x_{0,1} \\ &= 2\rho k_1^{xx}(\alpha + b_0 G)\left[1 - \rho(\alpha + b_0 G)^2\right]^{-1} G x_{0,1} \end{aligned} \tag{B-13}$$

The component associated with the constant term γ, i.e. $c_1^{\beta x}$, can be rewritten as

$$\begin{aligned} c_1^{\beta x} &= 2\rho k_1^{xx}(\alpha + b_0 G) g + \\ &\quad 2\rho k_1^{xx}(\alpha + b_0 G)\sum_{j=2}^{\infty}\left\{g + (g_{0,j} - g)\right\}\rho^{j-1}(\alpha + b_0 G)^{j-1} \end{aligned} \tag{B-14}$$

with

$$g_{0,j} - g = (g_{0,2} - g)\left[\sum_{\substack{i=1 \\ j\geq 2}}^{j-1} (\alpha + b_0 G)^{i-1}\right] \tag{B-15}$$

$$(g_{0,2} - g_{0,1}) \equiv (g_{0,2} - g) = G\alpha^{-1}\gamma(\alpha + b_0 G + b_0 G\rho p^*) \tag{B-16}$$

because

$$g_{0,i} - g_{0,i-1} = g_{0,2} - g \quad \text{for} \quad i = 1, 2, \dots, j \tag{B-17}$$

The first infinite summation on the right hand side is equal to

$$\begin{aligned} \sum_{j=2}^{\infty} \rho^{j-1}(\alpha + b_0 G)^{j-1} &= \rho(\alpha + b_0 G)\sum_{j=0}^{\infty} \rho^j (\alpha + b_0 G)^j \\ &= \rho(\alpha + b_0 G)\left[1 - \rho(\alpha + b_0 G)\right]^{-1} \end{aligned} \tag{B-18}$$

The double summation on the right hand side is equal to

$$\begin{aligned}
\Sigma_{j=2}^{\infty} \quad & \left[\sum_{i=1}^{j-1}(\alpha+b_0G)^{i-1}\right]\rho^{j-1}(\alpha+b_0G)^{j-1}\sum_{j=2}^{\infty}\rho^{j-1}(\alpha+bG)^{j-1} \\
& +(\alpha+bG)\sum_{j=3}^{\infty}\rho^{j-1}(\alpha+bG)^{j-1} \\
& +(\alpha+bG)^2\sum_{j=4}^{\infty}\rho^{j-1}(\alpha+bG)^{j-1} \\
& +(\alpha+G\beta)^3\sum_{j=5}^{\infty}\rho^{j-1}(\alpha+bG)^{j-1}+\ldots \\
= \quad & \\
& \rho(\alpha bG)\left[1+\rho(\alpha+bG)^2+\rho^2(\alpha+bG)^4+\ldots\right] \\
& \times\sum_{j=1}^{\infty}\rho^{j-1}(\alpha+bG)^{j-1} \\
= \quad & \\
& \rho(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)^2\right]^{-1}\left[1-\rho(\alpha+b_0G)\right]^{-1}
\end{aligned} \tag{B-19}$$

when the system is stable and $\rho<1$, then

$$\begin{aligned}
c_1^{\beta x} &= 2\rho k_1^{xx}(\alpha+b_0G)g+2\rho k_1^{xx}(\alpha+b_0G) \\
&\times \left\{g(\alpha+b_0G)\rho\left[1-\rho(\alpha+b_0G)\right]^{-1}+G\alpha^{-1}\gamma(\alpha+b_0G+b_0G\rho p^*)\right. \\
&\times \left.(\alpha+b_0G)\rho\left[1-(\alpha+b_0G)^2\rho\right]^{-1}\left[1-(\alpha+b_0G)\rho\right]^{-1}\right\} \\
&= 2\rho k_1^{xx}(\alpha+b_0G)g+2\rho k_1^{xx}(\alpha+b_0G)^2\rho\left[1-\rho(\alpha+b_0G)\right]^{-1} \\
&\quad\times\left\{g+(g_{0,2}-g)\left[1-\rho(\alpha+b_0G)^2\right]^{-1}\right\}
\end{aligned} \tag{B-20}$$

Therefore when the system is stable and $\rho<1$, the component $c_1^{\beta x}$ depends only upon $g_{0,1}\equiv g$ and $(g_{0,2}-g_{0,1})\equiv(g_{0,2}-g)$ and

$$\begin{aligned}
k_1^{\beta x} &= 2\rho k_1^{xx}(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)^2\right]^{-1}Gx_{0,1} \\
&+2\rho k_1^{xx}(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)\right]^{-1} \\
&\times g\left\{1+\rho(\alpha+b_0G)(g_{0,2}-g)g^{-1}\left[1-\rho(\alpha+b_0G)^2\right]^{-1}\right\}
\end{aligned} \tag{B-21}$$

With $x_{0,1}\equiv\hat{x}_{1|0}$. By repeating the same procedure for $k_2^{\beta x}$ yields

$$k_2^{\beta x}=2\sum_{j=2}^{\infty}\rho^j k_1^{xx}(\alpha+bG)^{j-1}u_{0,j} \tag{B-22}$$

and after replacing the nominal controls with equation (A-14) in Appendix A, computing the infinite summation and double summation and rearranging the terms, the quantity $k_2^{\beta x}$ can be rewritten as

$$\begin{aligned}
k_2^{\beta x} &= k_2^{\beta x}(x_{0,2})+c_2^{\beta x} \\
&= 2\rho^2 k_1^{xx}(\alpha+bG)^2\left[1-\rho(\alpha+bG)^2\right]^{-1}Gx_{0,1}+c_2^{\beta x}
\end{aligned} \tag{B-23}$$

with

$$\begin{aligned}
c_2^{\beta x} &= 2\rho^2 k_1^{xx}(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)\right]^{-1} \\
&\times g\left\{1+(g_{0,2}-g)g^{-1}+\rho(\alpha+b_0G)(g_{0,3}-g_{0,2})g^{-1}\left[1-\rho(\alpha+b_0G)^2\right]^{-1}\right\}
\end{aligned} \tag{B-24}$$

It should be noticed that

$$\begin{aligned} k_2^{\beta x}(x_{0,2}) &= 2\rho^2 k_1^{xx}(\alpha+b_0G)^2\left[1-\rho(\alpha+b_0G)^2\right]^{-1}Gx_{0,1} \\ &= \rho(\alpha+b_0G)k_1^{\beta x}(x_{0,1}) \end{aligned} \tag{B-25}$$

and

$$c_2^{\beta x} = \rho c_1^{\beta x} + 2\rho^2 k_1^{xx}(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)\right]^{-1}(g_{0,2}-g) \tag{B-26}$$

Repeating this procedure it can be shown that, in general,

$$\begin{aligned} &k_j^{\beta x} = k_j^{\beta x}(x_{0,j}) + c_j^{\beta x} = [\rho(\alpha+b_0G)]^{j-1}k_1^{\beta x}(x_{0,1}) \\ &+\rho^{j-1}c_1^{\beta x} + 2\textstyle\sum_{i=2}^{j}\rho^i k_1^{xx}(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)\right]^{-1}(g_{0,2}-g). \end{aligned} \tag{B-27}$$

Equation (B-27) simplifies to

$$\begin{aligned} k_j^{\beta x} &= [\rho(\alpha+b_0G)]^{j-1}\left\{\rho k_1^{xx}(\alpha+b_0G)\left[1-\rho(\alpha+b_0G)^2\right]^{-1}Gx_{0,1}\right\} \\ &= \tilde{k}_1^{\beta x}x_{0,1} \end{aligned} \tag{B-28}$$

when the constant term $\gamma$ is zero.

## Appendix C

### Deriving submatrix $\mathbf{k}^{\beta\beta}$ of the augmented system in the infinite horizon model

In the BWM model, when the unknown parameter $\beta$ is replaced by its estimate at time 0 $b0$, the general formula for $k^{\beta\beta}$, see e.g. Kendrick (1981; 2002, equation (10.42)) or Tucci (2004, equation 2.57), specializes to

$$\begin{aligned} k_j^{\beta\beta} &= \left(u_{0,j}^2 k_{j+1}^{xx} + u_{0,j}k_{j+1}^{\beta x}\right) + \left(u_{0,j}k_{j+1}^{x\beta} + k_{j+1}^{\beta\beta}\right) \\ &- \left[p_{j+1}^x + u_{0,j}k_{j+1}^{xx}b_0 + k_{j+1}^{\beta x}b_0\right]^2. \\ &\times\left(\lambda_j + k_{j+1}^{xx}b_0^2\right)^{-1} \end{aligned} \tag{C-1}$$

Using the results in Appendix B, when $j$ =1 this submatrix can be rewritten as

$$\begin{aligned} k_1^{\beta\beta} &= u_{0,1}\rho k_1^{xx}u_{0,1} + 2\left[k_2^{\beta x}x_{0,2} + c_2^{\beta x}\right]u_{0,1} + k_2^{\beta\beta} \\ &- \left\{\rho k_1^{xx}(\alpha+b_0G)G^{-1}u_{0,1} + u_{0,1}\rho k_1^{xx}b_0 + \left[k_2^{\beta x}(x_{0,2}) + c_2^{\beta x}\right]b_0\right\}^2 \\ &\times \left(\lambda_1 + \rho k_1^{xx}b_0^2\right)^{-1} \end{aligned} \tag{C-2}$$

with

$$\begin{aligned} k_2^{\beta\beta} &= u_{0,2}\rho k_2^{xx}u_{0,2} + 2\left[k_3^{\beta x}x_{0,3} + c_3^{\beta x}\right]u_{0,2} + k_3^{\beta\beta} \\ &- \left\{\rho k_2^{xx}(\alpha+b_0G)G^{-1}u_{0,2} + u_{0,2}\rho k_2^{xx}b_0 + \left[k_3^{\beta x}(x_{0,3}) + c_3^{\beta x}\right]b_0\right\}^2 \\ &\times \left(\lambda_2 + \rho k_2^{xx}b_0^2\right)^{-1} \end{aligned} \tag{C-3}$$

where $G$ is as in equation (A-10) in Appendix A. Then, by repeated substitution, it can be shown that

$$\begin{aligned} &k_1^{\beta\beta} = \sum_{j=1}^{\infty}\rho^j k_1^{xx}u_{0,j}^2 + 2\sum_{j=1}^{\infty}\left[k_{j+1}^{\beta x}(x_{0,j+1}) + c_{j+1}^{\beta x}\right]u_{0,j} \\ &-\sum_{j=1}^{\infty}\left\{\rho^j k_1^{xx}(\alpha+b_0G)G^{-1}u_{0,j} + u_{0,j}\rho^j k_1^{xx}b_0 + \left[k_{j+1}^{\beta x}x_{0,j+1} + c_{j+1}^{\beta x}\right]b_0\right\}^2 \\ &\times\left(\rho^j\lambda + \rho^j k_1^{xx}b_0^2\right)^{-1} \end{aligned} \tag{C-4}$$

When $\gamma$=0 and the desired paths are zero the first term reduces to

$$\begin{aligned}\sum_{j=1}^{\infty}\rho^{j}k_1^{xx}u_{0,j}^2 &= \sum_{j=1}^{\infty}\rho^{j}k_1^{xx}\left(G_{0,j}x_{0,1}\right)^2 \\ &= \rho k_1^{xx}\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-1}G^2x_{0,1}^2\end{aligned} \tag{C-5}$$

with $x_{0,1}\equiv\hat{x}_{1|0}$, the second one looks like

$$\begin{aligned}&2\sum_{j=1}^{\infty}k_{j+1}^{\beta x}x_{0,j+1}G_{0,j}x_{0,1} = \\ &2\rho k_1^{xx}\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-2}\rho\left(\alpha+b_0G\right)^2G^2x_{0,1}^2\end{aligned} \tag{C-6}$$

and the squared portion is

$$\begin{aligned}&\sum_{j=1}^{\infty}\left\{\rho k_j^{xx}b_0u_{0,j}+k_{j+1}^{\beta x}b_0\right\}^2\left(\lambda_j+\rho k_j^{xx}b_0^2\right)^{-1} = \\ &\sum_{j=1}^{\infty}\left\{\rho^{j}k_1^{xx}\left(\alpha+b_0G\right)^{j-1}\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-1}b_0Gx_{0,1}\right\}^2 \\ &\times\left[\rho^{j-1}\left(\lambda_1+\rho k_1^{xx}b_0\right)\right]^{-1}\end{aligned} \tag{C-7}$$

Then equation (C-4) specializes to

$$\begin{aligned}k_1^{\beta\beta} &= \rho k_1^{xx}\left[1+\rho\left(\alpha+b_0G\right)^2\right]\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-2}G^2x_{0,1}^2 \\ &\quad-\left(\rho k_1^{xx}\right)^2\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-3}\left(\lambda_1+\rho k_1^{xx}b_0^2\right)^{-1}b_0^2G^2x_{0,1}^2 \\ &= \bar{k}_1^{\beta\beta}x_{0,1}^2\end{aligned} \tag{C-8}$$

Similarly, when $\gamma$ = 0, the desired paths are zero and the system is stabilizable

$$\begin{aligned}k_2^{\beta\beta} &= \rho k_1^{xx}\left[1+\rho\left(\alpha+b_0G\right)^2\right]^{-2}\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-2}\rho\left(\alpha+b_0G\right)^2G^2x_{0,1}^2 \\ &- \left(\rho k_1^{xx}\right)^2\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-3}\left(\lambda_1+\rho k_1^{xx}b_0^2\right)^{-1}b_0^2\rho\left(\alpha+b_0G\right)^2G^2x_{0,1}^2\end{aligned} \tag{C-9}$$

By comparing $k_1^{\beta\beta}$ and $k_2^{\beta\beta}$ it is apparent that, in this special case,

$$k_2^{\beta\beta}=\rho\left(\alpha+b_0G\right)^2k_1^{\beta\beta} \tag{C-10}$$

and by repeating this procedure it is possible to show that in general

$$\begin{aligned}k_j^{\beta\beta} &= \rho\left(\alpha+b_0G\right)^2k_{j-1}^{\beta\beta} \\ &= \left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}k_1^{\beta\beta}\end{aligned} \tag{C-11}$$

**Appendix D**

**Deriving the updated variance of the augmented system in the infinite horizon model**

By ~~10~~ combining equations (10) and (30) in the text, it follows that the updated variance of the stochastic parameter $\beta$ in the BWM model for a generic period *j* is given by

$$\begin{aligned}\sigma_{j|j}^{\beta\beta} &= \sigma_{j-1|j-1}^{\beta\beta}-\left(\sigma_{j-1|j-1}^{\beta\beta}u_{0,j-1}\right)^2\left(u_{0,j-1}^2\sigma_{j-1|j-1}^{\beta\beta}+q\right)^{-1} \\ &= \sigma_{j-1|j-1}^{\beta\beta}q\left(u_{0,j-1}^2\sigma_{j-1|j-1}^{\beta\beta}+q\right)^{-1}\end{aligned} \tag{D-1}$$

It follows that

$$\sigma_{1|1}^{\beta\beta}=\sigma_{0|0}^{\beta\beta}q\left(u_0^2\sigma_{0|0}^{\beta\beta}+q\right)^{-1} \tag{D-2}$$

with $\sigma_{0|0}^{\beta\beta} \equiv \sigma_b^2$ as in the text and, using this result, the updated variance for $j$ = 2 can be rewritten as

$$\begin{aligned} \sigma_{2|2}^{\beta\beta} &= \sigma_{1|1}^{\beta\beta} q\left(u_{0,1}^2 \sigma_{1|1}^{\beta\beta}+q\right)^{-1} \\ &= \sigma_{0|0}^{\beta\beta} q\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1}\left(u_{0,1}^2 \sigma_{0|0}^{\beta\beta}\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1}+1\right)^{-1} \\ &= \sigma_{0|0}^{\beta\beta} q\left[\sigma_{0|0}^{\beta\beta}\left(u_{0,1}^2+u_0^2\right)+q\right]^{-1} \end{aligned} \tag{D-3}$$

By repeating this procedure it can be shown that in general

$$\begin{aligned} \sigma_{j|j}^{\beta\beta} &= \sigma_{j-1|j-1}^{\beta\beta} q\left(u_{0,j-1}^2 \sigma_{j-1|j-1}^{\beta\beta}+q\right)^{-1} \\ &= \sigma_{0|0}^{\beta\beta} q\left(\sigma_{0|0}^{\beta\beta} \sum_{i=0}^{j-1} u_{0,i}^2+q\right)^{-1} \end{aligned} \tag{D-4}$$

when $\sigma_{j-1|j-1}^{\beta\beta}$ is replaced by its definition and $u_{0,0} \equiv u_0$. From equation (A-21) in Appendix A, it is known that when there is no exogenous variable or intercept, and the desired path for the state and control are zero as asssumed here, the nominal control and state are simply

$$u_{0,j} = G(\alpha+b_0 G)^{j-1} x_{0,1} \quad \text{for} \quad j=1,2,\ldots$$

with

$$x_{0,1} \equiv \hat{x}_{1|0} = \alpha x_0 + b_0 u_0$$

and the unknown parameter $\beta$ replaced by its estimate at time 0, i.e. $b_0$. Then

$$\begin{aligned} \sigma_{2|2}^{\beta\beta} &= \sigma_{0|0}^{\beta\beta} q\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1} q\left[u_{0,1}^2 \sigma_{0|0}^{\beta\beta} q\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1}+q\right]^{-1} \\ &= \sigma_{0|0}^{\beta\beta} q\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1}\left[1+G^2\left(\alpha x_0+b_0 u_0\right)^2 \sigma_{0|0}^{\beta\beta}\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1}\right]^{-1} \\ &= \sigma_{1|1}^{\beta\beta}(1+S)^{-1} \end{aligned} \tag{D-5}$$

with

$$S = G^2\left(\alpha x_0+b_0 u_0\right)^2 \sigma_{0|0}^{\beta\beta}\left(u_0^2 \sigma_{0|0}^{\beta\beta}+q\right)^{-1} \tag{D-6}$$

The updated variance for $j$ =3 is

$$\sigma_{3|3}^{\beta\beta} = \sigma_{1|1}^{\beta\beta}(1+S)^{-1} q\left[u_{0,2}^2 \sigma_{1|1}^{\beta\beta}(1+S)^{-1}+q\right]^{-1} \tag{D-7}$$

then using the definition of the nominal control and rearranging yields

$$\begin{aligned} \sigma_{3|3}^{\beta\beta} &= \sigma_{1|1}^{\beta\beta}(1+S)^{-1}\left[(\alpha+b_0 G)^2\left(1+S^{-1}\right)^{-1}+1\right]^{-1} \\ &= \sigma_{1|1}^{\beta\beta}\left[1+S+(\alpha+b_0 G)^2 S\right]^{-1} \end{aligned} \tag{D-8}$$

By repeating this procedure it can be shown that in general

$$\sigma_{j|j}^{\beta\beta} = \sigma_{1|1}^{\beta\beta}\left[1+S \sum_{\substack{l=2 \\ for j \geq 2}}^{j}(\alpha+b_0 G)^{2(l-2)}\right]^{-1} \tag{D-9}$$

## Appendix E

### The deterministic component

The deterministic component of the approximate cost-to-go can be written as in Kendrick (1981; 2002, equation

(10.49)), i.e.

$$J_{D,T-1} = \frac{1}{2}\lambda_t u_t^2 + \frac{1}{2}k_T^{CE} x_{0,T}^2 + \frac{1}{2}\sum_{j=t+1}^{T-1}\left(k_j^{CE} x_{0,j}^2 + 2p_j^{CE} x_{0,j} + \lambda_j u_{0,j}^2\right) \tag{E-1}$$

with *CE* indicating the Certainty Equivalence value associated with the non-augmented model, and in the infinite horizon model when $t = 0$ it looks like

$$J_{D,\infty} = \frac{1}{2}\lambda_0 u_0^2 + \frac{1}{2}\sum_{j=1}^{\infty}\left(k_j^{CE} x_{0,j}^2 + \lambda_j u_{0,j}^2\right) \tag{E-2}$$

where $k^{CE}$ and $k^{CE}$ are the fixed point solutions to the usual Riccati equations, $k_j^{CE} \equiv k_j^{xx} = \rho^j k^{xx} \equiv \rho^j k^{CE}$, $\lambda_0 = \lambda$, $\rho$ is the discount factor and the unknown parameter $\beta$ is replaced by its estimate at time 0, i.e. $b_0$. Equation (E-2) can be rewritten as

$$J_{D,\infty} = \psi_1 u_0^2 + \psi_2 u_0 + \psi_3 \tag{E-3}$$

with

$$\begin{aligned}
\psi_1 &= \frac{1}{2}\left\{\lambda + \rho b_0^2 (k^{CE} + \lambda G^2)[1-\rho(\alpha + b_0 G)^2]^{-1}\right\} \\
\psi_2 &= \rho(k^{CE} + \lambda G^2)[1-\rho(\alpha + b_0 G)^2]^{-1} b_0 \alpha x_{0,1} \\
\psi_3 &= \frac{1}{2}\left\{\rho(k^{CE} + \lambda G^2)[1-\rho(\alpha + b_0 G)^2]^{-1}\right\}(\alpha x_{0,1})^2
\end{aligned} \tag{E-4}$$

when there is no constant term and the desired path for the state and control are zero.

## Appendix F

### The cautionary component

The general formula for the cautionary component of the approximate cost-to-g0, see e.g. Kendrick (1981; 2002, equation (10.50)) or Tucci (2004, equation 2.68), for $t = 0$ and $T = \infty$ looks like

$$J_{C,\infty} = \frac{1}{2}\left(k_1^{xx}\sigma_{1|0}^{xx} + k_1^{\beta\beta}\sigma_{1|0}^{\beta\beta}\right) + k_1^{x\beta}\sigma_{1|0}^{x\beta} + \frac{1}{2}\sum_{j=1}^{\infty}\left(k_{j+1}^{xx} q\right) \tag{F-1}$$

with $k_1^{xx} = \rho k^{xx}$ in the infinite horizon model where $k^{xx}$ is the fixed point solution to the Riccati quantity described in Appendix A and

$$k_1^{\beta x} = \rho k_1^{xx}(\alpha + b_0 G)\left[1-\rho(\alpha + b_0 G)^2\right]^{-1} G x_{0,1} = \bar{k}_1^{\beta x} x_{0,1} \tag{F-2}$$

$$\begin{aligned}
k_1^{\beta\beta} &= \rho k_1^{xx}\left[1+\rho(\alpha + b_0 G)^2\right]\left[1-\rho(\alpha + b_0 G)^2\right]^{-2} G^2 x_{0,1}^2 \\
&\quad -(\rho k_1^{xx})^2\left[1-\rho(\alpha + b_0 G)^2\right]^{-3}\left(\lambda_1 + \rho k_1^{xx} b_0^2\right)^{-1} b_0^2 G^2 x_{0,1}^2 \\
&= \bar{k}_1^{\beta\beta} x_{0,1}^2
\end{aligned} \tag{F-3}$$

derived in Appendix A and Appendix B, where $x_{0,1}^2 \equiv \hat{x}_{1,0}^2$. By using the fact that the projected variances in this case look like $\sigma_{1|0}^{xx} = \sigma_{0|0}^{\beta\beta} u_0^2 + q$, $\sigma_{1|0}^{\beta x} = \sigma_{0|0}^{\beta\beta} u_0$, and $\sigma_{1|0}^{\beta\beta} = \sigma_{0|0}^{\beta\beta}$, after some manipulations the cautionary cost can be rewritten as

$$J_{C,\infty} = \delta_1 u_0^2 + \delta_2 u_0 + \delta_3 \tag{F-4}$$

with

$$\begin{aligned}
\delta_1 &= \frac{1}{2}k_1^{xx}\sigma_{0|0}^{\beta\beta} + \frac{1}{2}\sigma_{0|0}^{\beta\beta}\bar{k}_1^{\beta\beta} b_0^2 + \sigma_{0|0}^{\beta\beta}\bar{k}_1^{\beta x} b_0 \\
\delta_2 &= \sigma_{0|0}^{\beta\beta}\bar{k}_1^{\beta\beta} b_0 \alpha x_0 + \sigma_{0|0}^{\beta\beta}\bar{k}_1^{\beta x}\alpha x_0 \\
&= \sigma_{0|0}^{\beta\beta}\left(\bar{k}_1^{\beta\beta} b_0 + \bar{k}_1^{\beta x}\right)\alpha x_0 \\
\delta_3 &= \frac{1}{2}k_1^{xx} q(1-\rho)^{-1} + \frac{1}{2}\sigma_{0|0}^{\beta\beta}\bar{k}_1^{\beta\beta}\alpha^2 x_0^2
\end{aligned} \tag{F-5}$$

**Appendix G**

**The probing component**

The general formula for the probing component of the approximate cost-to-go, see e.g. Kendrick (1981; 2002, equation (10.51)) or Tucci (2004, equation 2.69), for $t$ = 0 and $T$ = ∞ looks like

$$J_{P,\infty}=\frac{1}{2}\sum_{j=1}^{\infty}\left[p_{j+1}^{x}+u_o\rho^{j}k_1^{xx}b_0+k_{j+1}^{\beta x}b_0\right]^2\left[\rho^{j}\left(\lambda_0+k_1^{xx}b_0^2\right)\right]^{-1}\sigma_{j|j}^{\beta\beta} \tag{G-1}$$

when the unknown parameter β is replaced by its estimate at time 0, i.e. $b_0$, and $k_1^{xx}=\rho k^{xx}$. By comparing the terms of this infinite summation with the definition of submatrix $k^{\beta\beta}$, it is apparent that they have a lot in common. Namely, the $j$–th term multiplying the updated variance corresponds to the 'minus term' in the formula for $k_j^{\beta\beta}$. As shown in Appendix C

$$k_j^{\beta\beta}=\rho\left(\alpha+b_0G\right)^2k_{j-1}^{\beta\beta}=\left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}k_1^{\beta\beta} \tag{G-2}$$

with

$$\begin{aligned}
k_1^{\beta\beta} &= \rho k_1^{xx}\left[1+\rho\left(\alpha+b_0G\right)^2\right]\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-2}G^2x_{0,1}^2\\
&\quad-\left(\rho k_1^{xx}\right)^2\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-3}\left(\lambda_1+\rho k_1^{xx}b_0^2\right)^{-1}b_0G^2x_{0,1}^2\\
&= \bar{k}_{1,1}^{\beta\beta}x_{0,1}^2-\bar{k}_{1,2}^{\beta\beta}x_{0,1}^2\\
&= \bar{k}_1^{\beta\beta}x_{0,1}^2
\end{aligned} \tag{G-3}$$

as given in equation (C-8). Then the probing component can be rewritten as

$$J_{P,\infty}=\frac{1}{2}\sum_{j=1}^{\infty}\left\{\left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}\bar{k}_{1,2}^{\beta\beta}x_{0,1}^2\right\}\sigma_{j|j}^{\beta\beta} \tag{G-4}$$

with $x_{0,1}^2\equiv\hat{x}_{1|0}^2$ as before. By replacing the updated variances in (G-4) with equation (D-9) in Appendix D it yields

$$\begin{aligned}
J_{P,\infty} &= \frac{1}{2}\left[\bar{k}_{1,2}^{\beta\beta}x_{0,1}^2\sigma_{0|0}^{\beta\beta}q\left(u_0^2\sigma_{0|0}^{\beta\beta}+q\right)^{-1}\right]\\
&\quad\times\Sigma_{j=1}^{\infty}\left\{\left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}\left[1+S\Sigma_{\substack{i=2\\ for\, j\geq 2}}^{j}\left(\alpha+b_0G\right)^{2(i-2)}\right]^{-1}\right\}
\end{aligned} \tag{G-5}$$

with $S=G^2\left(\alpha x_0+b_0u_0\right)^2\sigma_{0|0}^{\beta\beta}\left(u_0^2\sigma_{0|0}^{\beta\beta}+q\right)^{-1}$. The infinite sum in (G-5) can alternatively be written as

$$\begin{aligned}
&\sum_{j=1}^{\infty}\left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}\left[1+S\Sigma_{\substack{i=2\\ for\, j\geq 2}}^{j}\left(\alpha+b_0G\right)^{2(i-2)}\right]^{-1}=\\
&1+\rho\left(\alpha+b_0G\right)^2(1+S)^{-1}+\left[\rho\left(\alpha+b_0G\right)^2\right]^2\left[1+S+S\left(\alpha+b_0G\right)^2\right]^{-1}\\
&+\left[\rho\left(\alpha+b_0G\right)^2\right]^3\left[1+S+S\left(\alpha+b_0G\right)^2+S\left(\alpha+b_0G\right)^4\right]^{-1}+\ldots
\end{aligned} \tag{G-6}$$

with

$$\lim_{j\to\infty}\left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}=0$$

when the system is stabilizable, and

$$1<\lim_{j\to\infty}\left[1+S\sum_{\substack{i=2\\ for\, j\geq 2}}^{j}\left(\alpha+b_0G\right)^{2(i-2)}\right]=\left\{1+S\left[1-\left(\alpha+b_0G\right)^2\right]^{-1}\right\}<\infty \tag{G-7}$$

because all quantities are squared quantities or variances. One way to compute this infinite sum is by using the limiting ratio approach. The ratio between any two consecutive terms of equation (G-6) looks like

$$\frac{s_{j+1}}{s_j} = \frac{\left[\rho\left(\alpha+b_0G\right)^2\right]^j\left[1+S\sum_{\substack{i=2\\ for\, j\geq 2}}^{j+1}\left(\alpha+b_0G\right)^{2(i-2)}\right]^{-1}}{\left[\rho\left(\alpha+b_0G\right)^2\right]^{j-1}\left[1+S\sum_{\substack{i=2\\ for\, j\geq 2}}^{j}\left(\alpha+b_0G\right)^{2(i-2)}\right]^{-1}} \tag{G-8}$$

then the limiting ratio is

$$\begin{aligned}\lim_{j\to\infty}\left|\frac{s_{j+1}}{s_j}\right| &= \rho\left(\alpha+b_0G\right)^2\lim_{j\to\infty}\left|\frac{1+S\sum_{\substack{i=2\\ for\, j\geq 2}}^{j}\left(\alpha+b_0G\right)^{2(i-2)}}{1+S\sum_{\substack{i=2\\ for\, j\geq 2}}^{j+1}\left(\alpha+b_0G\right)^{2(i-2)}}\right| \\ &= \rho\left(\alpha+b_0G\right)^2\end{aligned} \tag{G-9}$$

When equation (G-9) is used to compute the infinite sum in (G-6) it yields

$$J_{P,\infty} = \frac{1}{2}\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-1}\sigma_{0|0}^{\beta\beta}q\left(u_0^2\sigma_{0|0}^{\beta\beta}+q\right)^{-1}\bar{k}_{1,2}^{\beta\beta}x_{0,1}^2 \tag{G-10}$$

This means that the probing component can be rearranged as in Amman and Kendrick (1995) and Tucci et al. (2010), namely

$$J_{P,\infty} = \frac{1}{2}\frac{g\left(u_0\right)}{h\left(u_0\right)} \tag{G-11}$$

with

$$h\left(u_0\right) = \left(u_0^2\sigma_{0|0}^{\beta\beta}+q\right)\left(\sigma_{0|0}^{\beta\beta}q\right)^{-1} \tag{G-12}$$

identical to the definition reported in those works and

$$g\left(u_0\right) = \left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-1}\bar{k}_{1,2}^{\beta\beta}x_{0,1}^2 = \phi_1\left(\phi_2u_0+\phi_3\right)^2 \tag{G-13}$$

with

$$\begin{aligned}\phi_1 &= \left[\left(\rho k_1^{xx}\right)^2\left(\lambda_1+\rho k_1^{xx}b_0^2\right)^{-1}\right]\left[1-\rho\left(\alpha+b_0G\right)^2\right]^{-1}b_0^2G^2 \\ \phi_2 &= b_0 \\ \phi_3 &= \alpha x_0\end{aligned} \tag{G-14}$$

**Appendix H**

**Comparing the deterministic component of the approximate cost-to-go in a two-period finitehorizon model with that in an infinite horizon model**

This appendix shows that the parameter definitions in the deterministic component of the approximate cost-to-go associated with the control applied at time 0 reported in Amman and Kendrick (1995) and Tucci et al. (2010) are consistent with those presented in Appendix E. The parameter $\psi_1$ in Tucci et al. (2010, equation 5.3) takes the form

$$\begin{aligned}\psi_1 &= \frac{\lambda_0}{2}+\frac{1}{2}b^2\left\{w_2\left[\alpha\left(1-\frac{b^2w_2}{\lambda_1+b^2w_2}\right)\right]^2+w_1\right. \\ &\qquad \left.+\lambda_1\left(\frac{-1}{\lambda_1+b^2w_2}\right)^2\left(\alpha bw_2\right)^2\right\}\end{aligned} \tag{H-1}$$

when there is no constant term and the desired path for the state and control are zero. Rearranging the terms yields

$$\begin{aligned}\psi_1 &= \frac{\lambda_0}{2}+\frac{1}{2}b^2\left\{w_1+w_2\alpha^2-\alpha^2b^2w_2^2\left[\lambda_1+b^2w_2\right]^{-1}\right\} \\ \psi_1 &= \frac{1}{2}\left(\lambda+b^2k_1^{CE}\right)\end{aligned} \tag{H-2}$$

Similarly, the parameter $\psi_2$ in their equation (24) looks like

$$\begin{aligned}\psi_2 &= w_2b\alpha\left(1-\frac{b^2w_2}{\lambda_1+b^2w_2}\right)\left[b\left(-\frac{1}{\lambda_1+b^2w_2}\right)\alpha^2bw_2x_0+\alpha^2x_0\right] \\ &+ w_1\left(\alpha x_0\right)b+\left(-\frac{\lambda_1}{\lambda_1+b^2w_2}\right)\alpha b^2w_2\left[\left(-\frac{1}{\lambda_1+b^2w_2}\right)\alpha^2bw_2x_0\right]\end{aligned} \tag{H-3}$$

when there is no constant term and the desired path for the state and control are zero and after some minor manipulations it yields

$$\begin{aligned}\psi_2 &= b\left\{w_2\alpha^2 - w_2\alpha^2\left[\lambda_1 + \left(b^2w_2\right)\right]b^2w_2\left(\lambda_1 + b^2w_2\right)^{-2} + w_1\right\}\alpha x_0 \\ \psi_2 &= k_1^{CE} b\alpha x_0\end{aligned} \tag{H-4}$$

Finally, the parameter $\psi_3$ in Tucci et al. (2010, equation 5.3) can be rewritten as

$$\begin{aligned}\psi_3 = \frac{w_2}{2}\left\{\left(-\frac{1}{\lambda_1 + b^2w_2}\right)\alpha^2b^2w_2x_0 + \alpha^2x_0\right\}^2 + \frac{w_1}{2}\left(\alpha x_0\right)^2 \\ +\tfrac{\lambda_1}{2}\left[\left(-\frac{1}{\lambda_1 + b^2w_2}\right)\alpha^2bw_2x_0\right]^2\end{aligned} \tag{H-5}$$

and after explicating the squared terms and simplifying it yields

$$\begin{aligned}\psi_3 = \Bigg\{\left(w_2b^2 + \lambda_1\right)\left(-\frac{1}{\lambda_1 + b^2w_2}\right)^2 b^2w_2^2\alpha^2 + w_2\alpha^2 \\ +2\left(-\frac{1}{\lambda_1 + b^2w_2}\right)b^2w_2^2\alpha^2 + w_1\Bigg\}\frac{1}{2}\left(\alpha x_0\right)^2 = \frac{1}{2}k_1^{CE}\left(\alpha x_0\right)^2\end{aligned} \tag{H-6}$$

It is straightforward that equations (H-2), (H-4) and (H-6) are identical to the equations in (E-4) in Appendix E when the estimate of the unknown parameter β at time 0 is denoted by $b$, instead of $b_0$ as in the present paper, and the finite horizon Riccati quantity is replaced by its 'infinite-horizon' counterpart.

## Appendix I

### Comparing the cautionary component of the approximate cost-to-go in a two-period finitehorizon model with that in an infinite horizon model

This appendix shows that the parameter definitions in the cautionary component of the approximate cost-to-go associated with the control applied at time 0 reported in Amman and Kendrick (1995) and Tucci et al. (2010) are consistent with those presented in Appendix F. In a two-period BWM model with unknown parameter $\beta$, this component looks like

$$J_{C,2} = \frac{1}{2}\left(k_1^{xx}\sigma_{1|0}^{xx} + 2k_1^{x\beta}\sigma_{1|0}^{\beta x} + k_1^{\beta\beta}\sigma_{1|0}^{\beta\beta}\right) + \frac{1}{2}k_2^{xx}q \tag{I-1}$$

with $\sigma_{1|0}^{xx} = \sigma_{0|0}^{\beta\beta}u_0^2 + q$, $\sigma_{1|0}^{\beta x} = \sigma_{0|0}^{\beta\beta}u_0$, $\sigma_{1|0}^{\beta\beta} = \sigma_{0|0}^{\beta\beta}$ and $k_2^{xx} = w_2$. In Tucci et al. (2010, equation 4.1) it takes the form

$$\begin{aligned}J_{C,2} = \frac{\sigma_b^2w_2}{2}\left(\alpha u_0 + u_{0,1}\right)^2 + \frac{\sigma_b^2}{2}\left(-\frac{1}{\lambda_1 + b^2w_2}\right)\left(\alpha bw_2u_0 + bw_2u_{0,1} + w_2x_{0,2}\right)^2 \\ +\tfrac{q}{2}\left[\alpha^2w_2 + w_2 + w_1 + \left(-\frac{1}{\lambda_1 + b^2w_2}\right)\left(\alpha bw_2\right)^2\right] + \frac{\sigma_b^2w_1}{2}u_0^2\end{aligned} \tag{I-2}$$

with $u_{0,1}$ and $x_{0,2}$ the nominal, or $CE$, values of $u_1$ and $x_2$ defined as

$$u_{0,1} = \left(-\frac{1}{\lambda_1 + b^2w_2}\right)\left[\alpha b^2w_2u_0 + \alpha^2bw_2x_0\right] \tag{I-3}$$

$$x_{0,2} = b\left(\alpha - \frac{\alpha b^2w_2}{\lambda_1 + b^2w_2}\right)u_0 + \alpha^2x_0 + \left(-\frac{1}{\lambda_1 + b^2w_2}\right)\alpha^2b^2w_2x_0 \tag{I-4}$$

when there is no constant term and the desired path for the state and control are zero. Then it is convenient to rewrite (I-3) and (I-4) as

$$\begin{aligned}u_{0,1} &= \left(-\frac{\alpha bw_2}{\lambda_1 + b^2w_2}\right)\left(bu_0 + \alpha x_0\right) \\ &= G_1x_{0,1}\end{aligned} \tag{I-5}$$

and

$$\begin{aligned}x_{0,2} &= b\left(\alpha - b\frac{\alpha bw_2}{\lambda_1 + b^2w_2}\right)u_0 + \alpha^2x_0 + b\left(-\frac{\alpha bw_2}{\lambda_1 + b^2w_2}\right)\alpha x_0 \\ &= \left(\alpha + bG_1\right)x_{0,1}\end{aligned} \tag{I-6}$$

respectively, with $G_1$ the usual feedback law in a two-period control problem. Then, equation (I-2) can be rewritten as

$$J_{C,2} = \delta_1 u_0^2 + \delta_2 u_0 + \delta_3 \tag{I-7}$$

with

$$\begin{aligned}
\delta_1 &= \frac{\sigma_b^2}{2}\left[v_1^2\left(w_2 - \frac{4b^2w_2^2}{\lambda_1 + b^2w_2}\right) + w_1\right] \\
&= \frac{\sigma_b^2}{2}\left[(\alpha + bG_1)^2\left(w_2 - \frac{4b^2w_2^2}{\lambda_1 + b^2w_2}\right) + w_1\right] \\
\delta_2 &= \sigma_b^2 w_2 v_1\left\{v_2 - \frac{2bw_2(2bv_2 + v_3)}{\lambda_1 + b^2w_2}\right\} \\
&= \sigma_b^2 w_2(\alpha + bG_1)\left\{G_1\alpha x_0 - \frac{2bw_2(2bG_1\alpha x_0 + \alpha^2 x_0)}{\lambda_1 + b^2w_2}\right\} \\
\delta_3 &= \frac{\sigma_b^2}{2}w_2\left[v_2^2 - \frac{w_2(2bv_2 + v_3)^2}{\lambda_1 + b^2w_2}\right] + \frac{q}{2}\left[\alpha^2 w_2 + w_2 + w_1 - \frac{(\alpha b w_2)^2}{\lambda_1 + b^2w_2}\right] \\
&= \frac{\sigma_b^2}{2}w_2\left[(G_1)^2 - \frac{w_2(\alpha + 2bG_1)^2}{\lambda_1 + b^2w_2}\right](\alpha x_0)^2 + \frac{q}{2}\left(w_2 + k_1^{xx}\right)
\end{aligned} \tag{I-8}$$

because the quantities defined in Tucci et al. (2010, equation 4.4) look like

$$\begin{aligned}
v_1 &= \alpha\left(1 - \frac{b^2w_2}{\lambda_1 + b^2w_2}\right) \\
&= \alpha + bG_1
\end{aligned} \tag{I-9}$$

$$\begin{aligned}
v_2 &= \left(-\frac{1}{\lambda_1 + b^2w_2}\right)\alpha^2 b w_2 x_0 \\
&= G_1 \alpha x_0
\end{aligned} \tag{I-10}$$

$$v_3 = \alpha^2 x_0 \tag{I-11}$$

in this simpler setup. Equations (I-8) are identical to equations (F-5) in Appendix F when the estimate of the unknown parameter β at time 0 is denoted by $b$, instead of $b_0$ as in the rest of the present paper, because

$$\begin{aligned}
k_1^{\beta x} &= 2w_2(\alpha + bG_1)G_1 x_{0,1} \\
&= \tilde{k}_1^{\beta x} x_{0,1}
\end{aligned} \tag{I-12}$$

$$\begin{aligned}
k_1^{\beta\beta} &= w_2 G_1^2 x_{0,1}^2 + w_2^2(\alpha + 2bG_1)^2\left[-\left(\lambda_1 + b^2w_2\right)\right]^{-1} x_{0,1}^2 \\
&= \tilde{k}_1^{\beta\beta} x_{0,1}^2
\end{aligned} \tag{I-13}$$

in the two-period horizon, and $\delta_1$ in equation (I-8) can be rearranged as

$$\begin{aligned}
\delta_1 = \frac{\sigma_b^2}{2}\Big\{& w_2\alpha^2 + G_1\alpha b w_2 + w_1 + 2w_2\left[\alpha + (\alpha + 2bG_1)\right]G_1 b \\
& + w_2 G_1^2 b^2 + \left(-\frac{1}{\lambda_1 + b^2w_2}\right) w_2^2(\alpha + 2bG_1)^2 b^2\Big\}
\end{aligned} \tag{I-14}$$

with the first three terms in braces corresponding to $k_1^{xx}$, the fourth term to $\tilde{k}_1^{\beta x} b$ and the last two to $\tilde{k}_1^{\beta\beta} b^2$.

## Appendix J

### Comparing the probing component of the approximate cost-to-go in a two-period finite horizon model with that in an infinite horizon model

This appendix shows that the parameter definitions in the probing component of the approximate cost-to-go associated with the control applied at time 0 reported in Amman and Kendrick (1995) and Tucci et al. (2010) are consistent with those presented in Appendix G. In Tucci et al. (2010), the function $h(u_0)$ in this component is identical to equation (G-12) in Appendix G and their $g(u_0)$, labeled equation (8), takes the form

$$g(u_0) = \left(\frac{w_2^2}{\lambda_1 + b^2 w_2}\right)(bu_{0,1} + x_{0,2})^2 \tag{J-1}$$

with $u_{0,1}$ and $x_{0,2}$ the nominal, or *CE*, values of $u_1$ and $x_2$ defined as

$$u_{0,1} = \left(-\frac{1}{\lambda_1 + b^2 w_2}\right)\left[\alpha b^2 w_2 u_0 + \alpha^2 b w_2 x_0\right] \tag{J-2}$$

$$x_{0,2} = b\left(\alpha - \frac{\alpha b^2 w_2}{\lambda_1 + b^2 w_2}\right) u_0 + \alpha^2 x_0 + \left(-\frac{1}{\lambda_1 + b^2 w_2}\right)\alpha^2 b^2 w_2 x_0 \tag{J-3}$$

when there is no constant term and the desired path for the state and control are zero. Then it is straightforward to rewrite (J-2) and (J-3) as

$$\begin{aligned} u_{0,1} &= \left(-\frac{\alpha b w_2}{\lambda_1 + b^2 w_2}\right)(bu_0 + \alpha x_0) \\ &= G_1 x_{0,1} \end{aligned} \tag{J-4}$$

and

$$\begin{aligned} x_{0,2} &= b\left(\alpha - b\frac{\alpha b w_2}{\lambda_1 + b^2 w_2}\right) u_0 + \alpha^2 x_0 + b\left(-\frac{\alpha b w_2}{\lambda_1 + b^2 w_2}\right)\alpha x_0 \\ &= (\alpha + bG_1)\, x_{0,1} \end{aligned} \tag{J-5}$$

respectively, with $G_1$ the usual optimal control law in a two-period control problem. Using equations (J-4) and (J-5) in (J-1) and rearranging it yields

$$\begin{aligned} g(u_0) &= \left(\frac{w_2^2}{\lambda_1 + b^2 w_2}\right)(\alpha + 2bG_1)^2 x_{0,1}^2 \\ &= \phi_1 (\phi_2 u_0 + \phi_3)^2 \end{aligned} \tag{J-6}$$

where the old definitions simplify to, in this simpler setup,

$$\begin{aligned} \phi_1 &= \left(\frac{w_2^2}{\lambda_1 + b^2 w_2}\right) \\ \phi_2 &= \alpha b\left(1 - \frac{2b^2 w_2}{\lambda_1 + b^2 w_2}\right) = (\alpha + 2bG_1)\, b \\ \phi_3 &= 2b\left(-\frac{1}{\lambda_1 + b^2 w_2}\right)\alpha^2 b w_2 x_0 + \alpha^2 x_0 \\ &= (\alpha + 2bG_1)\,\alpha x_0 \end{aligned} \tag{J-7}$$

Equations (J-7) are identical to equations (G-14) in Appendix G when the estimate of the unknown parameter β at time 0 is denoted by $b$, instead of $b_0$ as in the present paper, the finite horizon Riccati quantity $w_2$ is replaced by its *infinite-horizon* counterpart $\rho k_1^{xx}$ and the infinite path for the nominal state and control are taken into account. By doing so, the usual optimal control law in a two-period control problem is replaced by the infinite sum of the time-invariant feedback matrix.