UNIVERSITÀ DI SIENA
1240

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Hans M. Amman
Marco P. Tucci

The DUAL Approach in an Infinite Horizon Model

n. 766 – Dicembre 2017

Abstract. In this paper we deliver the solution for the DUAL approach Kendrick (1981; 2002) with an infinite horizon. The results of this solutions form the basis for the paper Amman and Tucci (2017).

Keywords: Optimal experimentation, value function, approximation method, adaptive control, active learning, time-varying parameters, numerical experiments.

JEL Classification: C63, E61, E62.

Authors:

Hans M. Amman (amman@uva.nl), Faculty of Economics and Business, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, the Netherlands.

Marco P. Tucci (tucci@unisi.it), Dipartimento di Economia Politica Università di Siena, Piazza S. Francesco, 7, 53100 Siena, Italy.

The DUAL Approach in an Infinite Horizon Model

1. Introduction

Most of the literature dealing with the DUAL control method based on Tse and Bar-Shalom (1973) and Kendrick (1981; 2002) seminal works deals with finite horizon problems. In these pages the DUAL solution to the BMW infinite horizon model is reported. In Section 2 the problem is stated and the augmented system treating the stochastic parameter as an additional state variable is defined. Then the one-period ahead projection of the mean and variance of the augmented state is obtained (Section 3). In Section 4 the formula for the nominal path for the state and control in the infinite horizon are presented and the time-invariant feedback rule defined. The appropriate Riccati quantities for the augmented system are derived (Section 5). Section 6 contains the formulae for the updated covariances of the augmented system. Finally the new approximate cost-to-go is presented for the special case where the desired path for the state and control are set equal to 0 and the linear system has no constant (Section 7). A numerical example, based on BW dataset, comparing the DUAL infinite solution optimal control with the two-period finite horizon solution discussed in Tucci et al. (2010) is presented in Section 8. The major conclusions are summarized in Section 9. For the reader's sake, most of the technical derivations are confined to a number of short appendices.

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

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

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

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

with x_t and u_t the state and control variables, respectively, and the tilde indicating the desired path of the specified variable. Also α, β and γ are the parameters of the system equation and ε_{t+1} is an error term identically and independently distributed (i.i.d.) normal with mean zero and variance q. Finally, the initial state x_0 and the penalty weights w's and λ'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.[1]

Following Tse and Bar-Shalom (1973) methods for solving active learning

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

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.[2] This approximate cost-to-go is decomposed into three terms and, for the present problem, written as

$$J_N = J_{D,N} + J_{C,N} + J_{P,N} \tag{2-3}$$

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.[3]

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 $t = 0, 1, \ldots, \infty$, where $t = 0$ denotes the current period, which minimizes the linear functional

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

with the desired path for the state and control set equal to 0, x_t subject to the system equation (2-2) 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 (2-2) and (2-4) 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, ..., ∞ minimizing

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

with W_t^* having w_t on the top left corner and zeros elsewhere. subject to

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

[3]See Kendrick (1981; 2002, pages 97-98) for an introduction to this decomposition.

the discrete-time system equations, with no measurement equation,

$$z_{t+1} = A^z z_t + \beta_t^z u_t + \gamma^z + \varepsilon_t^z \tag{2-6}$$

with the arrays defined as

$$z_t = \begin{bmatrix} x_t \\ \beta_t \end{bmatrix}, A^z = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}, \beta_t^z = \begin{bmatrix} \beta_t \\ 0 \end{bmatrix}, \gamma^z = \begin{bmatrix} \gamma \\ 0 \end{bmatrix}, \varepsilon_t^z = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix} \tag{2-7}$$

Problems (2-2) and (2-4) and (2-5)-(2-7) 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 β) stochastic control problem" as noted in Kendrick (1981; 2002, page 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,** after control at time zero is applied, is

$$\hat{x}_{1|0} = \alpha x_0 + b_0 u_0^\tau + \gamma \tag{3-1}$$

$$b_{1|0} = b_0 \tag{3-2}$$

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_0^τ being the search control at iteration τ, with the Certainty Equivalence (CE) solution being the first search control, i.e. $u_0^1 \equiv u_0^{CE}$. The projected mean of the parameter is equal to its current estimate because the unknown parameter is assumed constant.

For the BMW problem with no measurement error, the projected variances look like[4]

[4]See, e.g., Kendrick (1981; 2002, Chapter 10, page 102) or Tucci (2004, chapter 2, pages 21-22) for details.

$$\begin{aligned}
\sigma^{xx}_{1|0} &= (u^{\tau}_0)^2\,\sigma^{\beta\beta}_{0|0} + q \\
\sigma^{\beta x}_{1|0} &= \sigma^{\beta\beta}_{0|0} u^{\tau}_0 \\
\sigma^{\beta\beta}_{1|0} &= \sigma^{\beta\beta}_{0|0} \equiv \sigma^2_b
\end{aligned} \tag{3-3}$$

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,

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

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.

$$G = -\left(\lambda + \rho k^{CE} b_0^2\right)^{-1} \alpha\rho k^{CE} b_0 \tag{4-1}$$

$$g = -\left(\lambda + \rho k^{CE} b_0^2\right)^{-1} b_0 \left(\rho k^{CE}\gamma + \rho p^{CE}\right) \tag{4-2}$$

with $k^{CE}_{j+1} = \rho k^{CE}_j$ *and* $p^{CE}_{j+1} = \rho p^{CE}_j\ \forall j$, where k^{CE} and p^{CE} are the fixed point solutions to the usual Riccati recursions[5]

$$k^{CE} = w + \alpha^2 \rho k^{CE} - \left(\alpha\rho k^{CE} b_0\right)^2 \left(\lambda + \rho k^{CE} b_0^2\right)^{-1} \tag{4-3}$$

and

[5] 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).

$$p^{CE} = \alpha \left(\rho k^{CE}\gamma + \rho p^{CE}\right) - \alpha\rho k^{CE} b_0^2 \left(\lambda + \rho k^{CE} b_0^2\right)^{-1} \left(\rho k^{CE}\gamma + \rho p^{CE}\right) \tag{4-4}$$

respectively. Then g can be rewritten as

$$g = G\alpha^{-1}\gamma \left(1 + \rho p^*\right) \tag{4-5}$$

with $p^* = \left[1 - \rho\left(\alpha + b_0 G\right)\right]^{-1} \left(\alpha + b_0 G\right)$. 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 γ. Namely

$$\begin{aligned} u_{0,j} &= G_{0,j} x_{0,j} + g_{0,j} \\ u_{0,j} &= G_{0,j} \hat{x}_{1|0} + g_{0,j} \end{aligned} \tag{4-6}$$

with

$$G_{0,j} = G\left(\alpha + b_0 G\right)^{j-1} \tag{4-7}$$

$$g_{0,j} = G\alpha^{-1}\gamma\left(\alpha + b_0 G + b_0 G\rho p^*\right) \sum_{i=1}^{j-1} \left(\alpha + b_0 G\right)^{i-1} + g \tag{4-8}$$

$$\text{for } j = 2, 3, \ldots$$

and the nominal control at time j can be rewritten as

$$x_{0,j} = \left(\alpha + b_0 G\right)^{j-1} \hat{x}_{1|0} + \alpha^{-1}\gamma\left(\alpha + b_0 G + b_0 G\rho p^*\right) \sum_{i=1}^{j-1} \left(\alpha + b_0 G\right)^{i-1} \tag{4-9}$$

In the special case where $\gamma = 0$, the nominal state and control are simply

$$u_{0,j} = G_{0,j} x_{0,j} = G_{0,j} \hat{x}_{1|0} \tag{4-10}$$

and

$$x_{0,j} = \left(\alpha + b_0 G\right)^{j-1} \hat{x}_{1|0} \tag{4-11}$$

5. Riccati Equations for the Arrays of the Augmented System

The $\mathbf{K}$ and $\mathbf{p}$ Riccati arrays of the augmented system are partitioned as

$$K_j = \begin{bmatrix} k_j^{xx} & k_j^{x\beta} \\ k_j^{\beta x} & k_j^{\beta\beta} \end{bmatrix}, \qquad p_j = \begin{bmatrix} p_j^{x} \\ p_j^{\beta} \end{bmatrix} \tag{5-1}$$

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. $\alpha + b_0 G$ is stable, and $\gamma = 0$ the quantities $k^{x\beta} = k^{\beta x}$and $k^{\beta\beta}$ reduce to

$$k_j^{\beta x} = \left[\rho\left(\alpha{+}b_0 G\right)\right]^{j-1} k_1^{\beta x} \tag{5-2}$$

with

$$\begin{aligned} k_1^{\beta x} &= 2\rho k_1^{xx}\left(\alpha + b_0 G\right)\left[1 - \rho\left(\alpha + b_0 G\right)^2\right]^{-1} G x_{0,1} \\ &= \tilde{k}_1^{\beta x} x_{0,1} \end{aligned} \tag{5-3}$$

as shown in Appendix B and Appendix F, and

$$\begin{aligned} k_j^{\beta\beta} &= \rho\left(\alpha + b_0 G\right)^2 k_{j-1}^{\beta\beta} = \left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} k_1^{\beta\beta} \\ &= \left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} \tilde{k}_1^{\beta\beta} x_{0,1}^2 \end{aligned} \tag{5-4}$$

with

$$\begin{aligned} \tilde{k}_1^{\beta\beta} = \rho k_1^{xx}&\left[1 + 3\rho\left(\alpha + b_0 G\right)^2\right]\left[1 - \rho\left(\alpha + b_0 G\right)^2\right]^{-2} G^2 \\ &-\left\{\rho k_1^{xx}\left(\alpha + 2b_0 G\left[1 - \rho\left(\alpha + b_0 G\right)^2\right]^{-1}\right)\right\}^2 \\ &\qquad \left(\lambda_1 + \rho k_1^{xx} b_0^2\right)^{-1}\left[1 - \delta\left(\alpha + b_0 G\right)^2\right]^{-1} \end{aligned} \tag{5-5}$$

as shown in Appendix C and Appendix F.[6] The elements of the $\mathbf{p}$ Riccati

[6] This compares with $k_1^{\beta x} = 2w_2\left(\alpha + bG_1\right) G_1 x_{0,1}$ and $k_1^{\beta\beta} = w_2 G_1^2 x_{0,1}^2 + w_2^2\left(\alpha + 2bG_1\right)^2\left[-\left(\lambda_1 + b^2 w_2\right)\right]^{-1} x_{0,1}^2$ in the two-period finite horizon

vector are defined as

$$p_j^x = k_j^{CE} x_{oj} + p_j^{CE} \tag{5-6}$$

and

$$\begin{aligned} p^\beta &= u_o p_{j+1}^x + p_{j+1}^\beta - \left[p_{j+1}^x + u_o k_{j+1}^{xx} b_0 + k_{j+1}^{\beta x} b_0 \right] \\ &\times \left(\lambda + k_{j+1}^{xx} b_0^2 \right)^{-1} \left(\lambda u_o + p_{j+1}^x b_0 \right) \end{aligned} \tag{5-7}$$

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 BMW problem the updating equations for the covariances of the augmented system look like[7]

$$\Sigma_{j|j} = \begin{bmatrix} O & O \\ -\sigma_{j|j-1}^{\beta x} \left(\sigma_{j|j-1}^{xx} \right)^{-1} & 1 \end{bmatrix} \Sigma_{j|j-1} \tag{6-1}$$

then the elements of the updated covariance matrix are defined as

$$\sigma_{j|j}^{xx} = 0, \sigma_{j|j}^{x\beta} \equiv \sigma_{j|j}^{\beta x} = 0, \sigma_{j|j}^{\beta\beta} = \sigma_{j|j-1}^{\beta\beta} - \sigma_{j|j-1}^{\beta x} \left(\sigma_{j|j-1}^{xx} \right)^{-1} \sigma_{j|j-1}^{x\beta} \tag{6-2}$$

where the projected covariances take the form in (3-3) when j and *j-1* replace 1 and 0, respectively. Combining (6-2) and (3-3), it yields, for $j = 1$,

$$\sigma_{1|1}^{\beta\beta} = \sigma_{1|0}^{\beta\beta} - \sigma_{1|0}^{\beta x} \left(\sigma_{1|0}^{xx} \right)^{-1} \sigma_{1|0}^{x\beta} = \sigma_b^2 q \left(u_0^2 \sigma_b^2 + q \right)^{-1} \tag{6-3}$$

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

---

model.

[7] See, e.g., Kendrick (1981; 2002, Chapter 10, page 103) or Tucci (2004, chapter 2, pages 27-28) for details..

$$
\begin{aligned}
\sigma_{j|j}^{\beta\beta} &= \sigma_b^2 q \left(\sigma_b^2 \sum_{i=0}^{j-1} u_{0,i}^2 + q \right)^{-1} \\
&= \sigma_b^2 q \left(\sigma_b^2 u_0^2 + q \right)^{-1} \left[1 + S \sum_{\substack{l=2 \\ for j \geq 2}}^{j} (\alpha + b_0 G)^{2(l-2)} \right]^{-1} \qquad \text{(6-4)}
\end{aligned}
$$

with

$$
S = G^2 \left(\alpha x_0 + b_0 u_0 \right)^2 \sigma_{0|0}^{\beta\beta} \left(u_0^2 \sigma_{0|0}^{\beta\beta} + q \right)^{-1} \qquad \text{(6-5)}
$$

and

$$
u_{0,0} \equiv u_0 \qquad \text{(6-6)}
$$

7. The Approximate Cost-to-Go

As in Kendrick (1981; 2002, Chapter 10) the approximate cost-to-go associated with the 'search' control u_t^τ 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 10.36 in Kendrick (1981; 2002),

$$
J_{D,\infty} = \frac{1}{2} \lambda_0 u_0^2 + \frac{1}{2} \hat{x}'_{1|0} K_{0,1}^{CE} \hat{x}_{1|0} + p'^{CE}_{0,1} \hat{x}_{1|0} \qquad \text{(7-1)}
$$

For the model at hand, equation (7-1) can be rewritten as

$$
J_{D,\infty} = \psi_1 u_0^2 + \psi_2 u_0 + \psi_3 \qquad \text{(7-2)}
$$

with

$$
\begin{aligned}
\psi_1 &= \frac{1}{2} \left(\lambda + b_0^2 k_{0,1}^{CE} \right) \\
\psi_2 &= \left[\left(\alpha x_0 + \gamma \right) k_{0,1}^{CE} + p_{0,1}^{CE} \ \right] b_0 \\
\psi_3 &= \frac{1}{2} \left(\alpha x_0 + \gamma \right)^2 k_{0,1}^{CE} + p_{0,1}^{CE} \left(\alpha x_0 + \gamma \right) \qquad \text{(7-3)}
\end{aligned}
$$

where $k_{0,1}^{CE} \equiv k_1^{xx} = \rho k^{xx}$. The parameters in equation (7-3) simplify to

$$
\begin{aligned}
\psi_1 &= \frac{1}{2}\left(\lambda + b_0^2 \rho k^{xx}\right) \\
\psi_2 &= \rho k^{xx} b_0 \alpha x_0 \\
\psi_3 &= \frac{1}{2}\rho k^{xx}\left(\alpha x_0\right)^2
\end{aligned} \tag{7-4}
$$

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

$$
J_{C,\infty} = \frac{1}{2}\left[k_1^{xx}\left(\sigma_b^2 u_0^2 + q\right) + k_1^{\beta\beta}\sigma_b^2\right] + k_1^{x\beta}\sigma_b^2 u_0 + \frac{1}{2}\sum_{j=1}^{\infty}\left(\rho^j k_1^{xx} q\right) \tag{7-5}
$$

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

$$
J_{C,\infty} = \delta_1 u_0^2 + \delta_2 u_0 + \delta_3 \tag{7-6}
$$

with

$$
\begin{aligned}
\delta_1 &= \frac{1}{2}\sigma_b^2\left(k_1^{xx} + \tilde{k}_1^{\beta\beta} b_0^2 + 2\tilde{k}_1^{\beta x} b_0\right) \\
\delta_2 &= \sigma_b^2\left(\tilde{k}_1^{\beta\beta} b_0 + \tilde{k}_1^{\beta x}\right)\alpha x_0 \\
\delta_3 &= \frac{1}{2}k_1^{xx} q\left(1-\rho\right)^{-1} + \frac{1}{2}\sigma_b^2\tilde{k}_1^{\beta\beta}\alpha^2 x_0^2
\end{aligned} \tag{7-7}
$$

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

$$
J_{P,\infty} = \frac{1}{2}\sum_{j=1}^{\infty}\left[p_{j+1}^{x} + u_o \rho^j k_1^{xx} b_0 + k_{j+1}^{\beta x} b_0\right]^2\left[\rho^j\left(\lambda_0 + k_1^{xx} b_0^2\right)\right]^{-1}\sigma_{j|j}^{\beta\beta} \tag{7-8}
$$

Similarly to Amman and Kendrick (1995) and Tucci et al. (2010), equation (7-8) can be rewritten as

$$
J_{P,\infty} = \frac{1}{2}\frac{g\left(u_0\right)}{h\left(u_0\right)} \tag{7-9}
$$

with

$$h(u_0) = \left(u_0^2 \sigma_b^2 + q\right)\left(\sigma_b^2 q\right)^{-1} \tag{7-10}$$

and

$$g(u_0) = \phi_1 (\phi_2 u_0 + \phi_3)^2 \tag{7-11}$$

with

$$\begin{aligned} \phi_1 &= \left[\rho \left(k_1^{xx}\right)^2 \left(\lambda + k_1^{xx} b_0^2\right)^{-1}\right]\left[1 - \rho(\alpha + b_0 G)^2\right]^{-2} \\ \phi_2 &= \left\{\alpha + 2b_0 G \left[1 - \rho(\alpha + b_0 G)^2\right]^{-1}\right\} b_0 \\ \phi_3 &= \left\{\alpha + 2b_0 G \left[1 - \rho(\alpha + b_0 G)^2\right]^{-1}\right\} \alpha x_0 \end{aligned} \tag{7-12}$$

as shown in Appendix G. At this point by substituting (7-3), (7-6) and (7-9) into (7-1) yields

$$\begin{aligned} J_\infty &= (\psi_1 + \delta_1) u_0^2 + (\psi_2 + \delta_2) u_0 + (\psi_3 + \delta_3) + \\ &\quad \left(\frac{\sigma_b^2 q}{2}\right) \frac{\phi_1 (\phi_2 u_0 + \phi_3)^2}{\left(\sigma_b^2 u_0^2 + q\right)} \end{aligned} \tag{7-13}$$

with the parameters defined as in (7-4), (7-7) and (7-12). 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, page 1367) which translates to

$$\alpha = 1, b_0 = -0.5, \gamma = 0, q = 1, \sigma_{0|0}^{\beta\beta} = \sigma_b^2 = 0.25, w = 1, \lambda = 0, \rho = 0.95 \tag{8-1}$$

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

$$\begin{aligned} k^{CE} &= 1 + \rho k^{CE} - 0.25 \left(\rho k^{CE}\right)^2 \left(0.25 \rho k^{CE}\right)^{-1} \\ &= 1 + \rho k^{CE} - \rho k^{CE} = 1 \end{aligned} \tag{8-2}$$

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

$$G = -\left(0.25 \rho k^{CE}\right)^{-1} \rho k^{CE} \left(-0.5\right) = 2 \tag{8-3}$$

It follows that the relevant terms for the computation of the approximate cost-to-go described in the previous section 7 specialize to

$$(\alpha + b_0 G) = 1 + 2\left(-0.5\right) = 0 \tag{8-4}$$

$$\begin{aligned} \rho k_1^{xx} &= \rho(\rho k^{xx}) = \rho^2 k^{xx} = (0.95)^2 \\ \tilde{k}_1^{\beta x} &= 2(0.95)^2(0)[1-(0.95)(0)^2]^{-1}2 = 0 \\ \tilde{k}_1^{\beta\beta} &= (0.95)^2 2^2 - \left\{(0.95)^2[1-2(1)^{-1}]\right\}^2 [0.25(0.95)^2]^{-1}(1)^{-1} = 0 \end{aligned} \tag{8-5}$$

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

$$\begin{aligned} \psi_1 &= \frac{1}{2}(0.25)0.95 = 0.119 \\ \psi_2 &= 0.95\left(-0.5\right)x_0 = -0.475 x_0 \\ \psi_3 &= \frac{1}{2}(0.95)x_0^2 = 0.475 x_0^2 \end{aligned} \tag{8-6}$$

$$\begin{aligned} \delta_1 &= \frac{1}{2}0.25(0.95) = 0.119 \\ \delta_2 &= 0 \\ \delta_3 &= \frac{1}{2}(0.95)(1)(0.05)^{-1} = 9.5 \end{aligned} \tag{8-7}$$

and

$$
\begin{aligned}
\phi_1 &= 0.95\,(0.95)^2\,(0.25 * 0.95)^{-1} = 0.95^2 * 4 \\
\phi_2 &= [1 + 4\,(-0.5)]\,(-0.5) = 0.5 \\
\phi_3 &= [1 + 4\,(-0.5)]\,x_0 = (-1)\,x_0 \qquad \text{(8-8)}
\end{aligned}
$$

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 δ_1 in the cautionary component and ϕ_1 in the probing one. It is noteworthy that the coefficient δ_2 in the cautionary component and ϕ_2 and ϕ_3 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 δ_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 ρ.

9. Conclusion

In these pages the DUAL solution to the BMW infinite horizon model has been presented. The appropriate Riccati quantities for the augmented system have been derived and the time-invariant feedback rule defined. When the desired path for the state and control are set equal to 0 and the linear system has no constant term, the new approximate cost-to-go looks identical to that associated with the finite horizon solution discussed in Amman and Kendrick (1995) and Tucci et al. (2010). Namely, the deterministic and cautionary component are quadratic functions of the time-0 control, and the probing component is the ratio of two quadratic functions in the time-0 control. Moreover the new definitions are perfectly consistent with those associated to the two-period finite horizon model.

Appendices

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[8]

$$\begin{aligned} u_{0,1} &= G_1 x_{0,1} + g_1 \\ &= \left(-\frac{1}{\lambda_1 + \beta^2 w_2} \right) \left[\alpha \beta w_2 x_{0,1} + \beta w_2 \left(\gamma - \tilde{x}_2 \right) - \lambda_1 \tilde{u}_1 \right] \end{aligned} \quad \text{(A-1)}$$

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

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

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 x_2 is

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

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

[8] See, e.g., Tucci et al. (2010).

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

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 γ. Namely,

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

with

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

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

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

$$x_{0,j} = \left[\prod_{i=1}^{j-1}(\alpha+\beta G_i)\right] x_{0,1} + \alpha^{-1}\gamma \sum_{i=1}^{j-1}\left[\prod_{l=i}^{j-1}(\alpha+\beta G_l)\right]$$

$$+\alpha^{-1}\sum_{i=1}^{j-1}\left[\prod_{l=i+1}^{j-1}(\alpha+\beta G_l)\right]\beta G_i k_{i+1}^{-1} p_{i+1} \quad \text{(A-9)}$$

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

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

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

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

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

and

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

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

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

with

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

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

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

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

with

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

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

where

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

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

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

$$x_{0,j} = (\alpha + \beta G)^{j-1} x_{0,1} \tag{A-22}$$

$$\text{for} \quad j = 2, 3, \ldots$$

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

In the BMW model, when the unknown parameter β 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

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

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,

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

then it follows that

$$\begin{aligned} p_2^x &= k_2^{xx}x_{0,2} + p_2^{CE} \\ &= \rho k_1^{xx}x_{0,2} + \rho p_1^{CE} \\ &= \rho k_1^{xx}\left(\alpha + b_0 G\right)x_{0,1} + c_2^p \end{aligned} \tag{B-4}$$

where

$$c_2^p = \rho k_1^{xx}\left(\alpha + b_0 G\right)\alpha^{-1}\gamma\left(1 + \rho p^*\right) \tag{B-5}$$

Therefore

$$\begin{aligned} p_2^x G &= \left[\rho k_1^{xx}\left(\alpha + b_0 G\right)x_{0,1} + \rho k_1^{xx}\left(\alpha + b_0 G\right)\alpha^{-1}\gamma\left(1 + \rho p^*\right)\right]G \\ &= \rho k_1^{xx}\left(\alpha + b_0 G\right)\left(G x_{0,1} + g\right) \end{aligned} \tag{B-6}$$

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

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

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

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

By using equation (A1.14) in Appendix A for the nominal control, it follows that $k_1^{\beta x}$ 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 (A1.16a) in Appendix A, into (A2.11a) 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

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

with

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

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

because

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

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

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

The double summation on the right hand side is equal to

$$\textstyle\sum_{j=2}^{\infty} \left[\sum_{i=1}^{j-1} (\alpha + b_0 G)^{i-1}\right] \rho^{j-1} (\alpha + b_0 G)^{j-1} \sum_{j=2}^{\infty} \rho^{j-1} (\alpha + bG)^{j-1}$$

$$+ (\alpha + bG) \sum_{j=3}^{\infty} \rho^{j-1} (\alpha + bG)^{j-1}$$

$$+ (\alpha + bG)^2 \sum_{j=4}^{\infty} \rho^{j-1} (\alpha + bG)^{j-1}$$

$$+ (\alpha + G\beta)^3 \sum_{j=5}^{\infty} \rho^{j-1} (\alpha + bG)^{j-1} + ...$$

$$=$$

$$\rho (\alpha b G) \left[1 + \rho (\alpha + bG)^2 + \rho^2 (\alpha + bG)^4 + ...\right]$$

$$\times \sum_{j=1}^{\infty} \rho^{j-1} (\alpha + bG)^{j-1}$$

$$=$$

$$\rho (\alpha + b_0 G) \left[1 - \rho (\alpha + b_0 G)^2\right]^{-1} [1 - \rho (\alpha + b_0 G)]^{-1} \tag{B-19}$$

when the system is stable and ρ <1, then

$$\begin{aligned} c_1^{\beta x} &= 2\rho k_1^{xx} (\alpha + b_0 G) g + 2\rho k_1^{xx} \big(\alpha + b_0 G\big) \\ &\times \left\{ g\big(\alpha + b_0 G\big) \rho [1 - \rho (\alpha + b_0 G)]^{-1} + G\alpha^{-1} \gamma (\alpha + b_0 G + b_0 G \rho p^*) \right. \\ &\times \left. (\alpha + b_0 G) \rho \left[1 - (\alpha + b_0 G)^2 \rho\right]^{-1} [1 - (\alpha + b_0 G) \rho]^{-1} \right\} \\ &= 2\rho k_1^{xx} (\alpha + b_0 G) g + 2\rho k_1^{xx} (\alpha + b_0 G)^2 \rho [1 - \rho (\alpha + b_0 G)]^{-1} \\ &\quad \times \left\{ g + (g_{0,2} - g) \left[1 - \rho (\alpha + b_0 G)^2\right]^{-1} \right\} \end{aligned} \tag{B-20}$$

Therefore when the system is stable and ρ <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 (A1.14) 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,

$$k_j^{\beta x} = k_j^{\beta x}(x_{0,j}) + c_j^{\beta x} = [\rho(\alpha + b_0 G)]^{j-1} k_1^{\beta x}(x_{0,1})$$
$$+\rho^{j-1} c_1^{\beta x} + 2\sum_{i=2}^{j} \rho^i k_1^{xx}(\alpha + b_0 G)[1 - \rho(\alpha + b_0 G)]^{-1}(g_{0,2} - g) \qquad \text{(B-27)}$$

equation (B-27) simplifies to

$$\begin{aligned} k_j^{\beta x} &= [\rho(\alpha + b_0 G)]^{j-1} \left\{ 2\rho k_1^{xx}(\alpha + b_0 G)\left[1 - \rho(\alpha + b_0 G)^2\right]^{-1} G x_{0,1} \right\} \\ &= \tilde{k}_1^{\beta x} x_{0,1} \qquad \text{(B-28)} \end{aligned}$$

when the constant term γ is zero.

Appendix C. Deriving submatrix $k^{\beta\beta}$ of the augmented system in the infinite horizon model

In the BMW model, when the unknown parameter β is replaced by its estimate at time 0 b_0, the general formula for $k^{\beta\beta}$, see e.g. Kendrick (1981; 2002, equation 10.42) or Tucci (2004, equation 2.57), specializes to

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

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

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

with

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

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

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

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

$$\begin{aligned} \sum_{j=1}^{\infty} \rho^j k_1^{xx} u_{0,j}^2 &= \sum_{j=1}^{\infty} \rho^j k_1^{xx} \left(G_{0,j} x_{0,1} \right)^2 \\ &= \rho k_1^{xx} \left[1 - \rho \left(\alpha + b_0 G \right)^2 \right]^{-1} G^2 x_{0,1}^2 \end{aligned} \tag{C-5}$$

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

$$\begin{aligned} 2 \sum_{j=1}^{\infty} k_{j+1}^{\beta x} x_{0,j+1} G_{0,j} x_{0,1} &= \\ 4 \rho k_1^{xx} \left[1 - \rho \left(\alpha + b_0 G \right)^2 \right]^{-2} \rho \left(\alpha + b_0 G \right)^2 G^2 x_{0,1}^2 \end{aligned} \tag{C-6}$$

and the squared portion is

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

Then equation (C-4) specializes to

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

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

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

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 Appendix D combining (3-3) and (6-2) in the text, it follows that the updated variance of the stochastic parameter β in the BMW 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

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

with

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

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

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

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

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

then using the definition of the nominal control and rearranging yields

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

By repeating this procedure it can be shown that in general

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

Appendix E. The deterministic component

The deterministic component of the approximate cost-to-go can be written as in Kendrick (1981; 2002, equation 10.35), i.e.

$$J_{D,T-t} = \frac{1}{2}\lambda_t u_t^2 + \frac{1}{2}k_{t+1}^{CE}\hat{x}_{t+1|t}^2 + p_{t+1}^{CE}\hat{x}_{t+1|t} \tag{E-1}$$

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

$$J_{D,\infty} = \frac{1}{2}\lambda_0 u_0^2 + \frac{1}{2}\rho k^{CE}\left(\alpha x_0 + b_0 u_0 + \gamma\right)^2 + \rho p^{CE}\left(\alpha x_0 + b_0 u_0 + \gamma\right) \tag{E-2}$$

where k^{CE} and p^{CE} are the fixed point solutions to the usual Riccati equations, ρ is the discount factor and the unknown parameter β 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

$$\begin{aligned} \psi_1 &= \frac{1}{2}\left(\lambda + \rho k^{CE} b_0^2\right) \\ \psi_2 &= \rho k^{CE} b_0 \alpha x_0 \\ \psi_3 &= \frac{1}{2}\rho k^{CE}\left(\alpha x_0\right)^2 \end{aligned} \tag{E-4}$$

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

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

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

derived in Appendix B and Appendix C, 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

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

Appendix G. The probing component

The general formula for the probing component of the approximate cost-to-g0, see e.g. Kendrick (1981; 2002, equation 10.51) or Tucci (2004, equation 2.69), for $t = 0$ and $T = \infty$ 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

$$k_j^{\beta\beta} = \rho\left(\alpha + b_0 G\right)^2 k_{j-1}^{\beta\beta} = \left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} k_1^{\beta\beta} \tag{G-2}$$

with

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

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

$$J_{P,\infty} = \frac{1}{2}\sum_{j=1}^{\infty}\left\{\left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} \tilde{k}_{1,2}^{\beta\beta} x_{0,1}^2\right\} \sigma_{j|j}^{\beta\beta} \tag{G-4}$$

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

$$J_{P,\infty} = \frac{1}{2}\left[\tilde{k}_{1,2}^{\beta\beta} x_{0,1}^2 \sigma_{0|0}^{\beta\beta} q \left(u_0^2 \sigma_{0|0}^{\beta\beta} + q\right)^{-1}\right] \times \sum_{j=1}^{\infty} \left\{ \left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} \left[1 + S \sum_{\substack{i=2 \\ for j \geq 2}}^{j} \left(\alpha + b_0 G\right)^{2(i-2)}\right]^{-1} \right\} \tag{G-5}$$

with $S = G^2 \left(\alpha x_0 + b_0 u_0\right)^2 \sigma_{0|0}^{\beta\beta} \left(u_0^2 \sigma_{0|0}^{\beta\beta} + q\right)^{-1}$. The infinite sum in (G-5) can alternatively be written as

$$\sum_{j=1}^{\infty} \left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} \left[1 + S \sum_{\substack{i=2 \\ for j \geq 2}}^{j} \left(\alpha + b_0 G\right)^{2(i-2)}\right]^{-1} = 1 + \rho\left(\alpha + b_0 G\right)^2 \left(1 + S\right)^{-1} + \left[\rho\left(\alpha + b_0 G\right)^2\right]^2 \left[1 + S + S\left(\alpha + b_0 G\right)^2\right]^{-1} + \left[\rho\left(\alpha + b_0 G\right)^2\right]^3 \left[1 + S + S\left(\alpha + b_0 G\right)^2 + S\left(\alpha + b_0 G\right)^4\right]^{-1} + \ldots \tag{G-6}$$

with

$$\lim_{j \to \infty} \left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} = 0$$

when the system is stabilizable, and

$$1 < \lim_{j \to \infty} \left[1 + S \sum_{\substack{i=2 \\ for j \geq 2}}^{j} \left(\alpha + b_0 G\right)^{2(i-2)}\right] = \left\{1 + S\left[1 - \left(\alpha + b_0 G\right)^2\right]^{-1}\right\} < \infty \tag{G-7}$$

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

$$\frac{s_{j+1}}{s_j} = \frac{\left[\rho\left(\alpha + b_0 G\right)^2\right]^{j} \left[1 + S \sum_{\substack{i=2 \\ for j \geq 2}}^{j+1} \left(\alpha + b_0 G\right)^{2(i-2)}\right]^{-1}}{\left[\rho\left(\alpha + b_0 G\right)^2\right]^{j-1} \left[1 + S \sum_{\substack{i=2 \\ for j \geq 2}}^{j} \left(\alpha + b_0 G\right)^{2(i-2)}\right]^{-1}} \tag{G-8}$$

then the limiting ratio is

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

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

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

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

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

with

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

Appendix H. Comparing the deterministic 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 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 ψ_1 in Tucci et al. (2010, equation 5.3) takes the form

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

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

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

Similarly, the parameter ψ_2 in their equation (5-3) looks like

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

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

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

Finally, the parameter ψ_3 in Tucci et al. (2010, equation 5.3) can be rewritten as

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

and after explicating the squared terms and simplifying it yields

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

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 finite horizon 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 BMW model with unknown parameter β, 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

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

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

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

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

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

$$\begin{aligned}u_{0,1} &= \left(-\frac{\alpha b w_2}{\lambda_1 + b^2 w_2}\right)\left(b u_0 + \alpha x_0\right) \\ &= G_1 x_{0,1}\end{aligned} \tag{I-5}$$

and

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

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

$$
J_{C,2} = \delta_1 u_0^2 + \delta_2 u_0 + \delta_3 \qquad \text{(I-7)}
$$

with

$$
\begin{aligned}
\delta_1 &= \frac{\sigma_b^2}{2}\left[\nu_1^2\left(w_2 - \frac{4b^2 w_2^2}{\lambda_1 + b^2 w_2}\right) + w_1\right] \\
&= \frac{\sigma_b^2}{2}\left[\left(\alpha + bG_1\right)^2\left(w_2 - \frac{4b^2 w_2^2}{\lambda_1 + b^2 w_2}\right) + w_1\right] \\
\delta_2 &= \sigma_b^2 w_2 \nu_1 \left\{\nu_2 - \frac{2bw_2\left(2b\nu_2 + \nu_3\right)}{\lambda_1 + b^2 w_2}\right\} \\
&= \sigma_b^2 w_2 \left(\alpha + bG_1\right)\left\{G_1 \alpha x_0 - \frac{2bw_2\left(2bG_1\alpha x_0 + \alpha^2 x_0\right)}{\lambda_1 + b^2 w_2}\right\} \\
\delta_3 &= \frac{\sigma_b^2}{2} w_2 \left[\nu_2^2 - \frac{w_2\left(2b\nu_2 + \nu_3\right)^2}{\lambda_1 + b^2 w_2}\right] + \frac{q}{2}\left[\alpha^2 w_2 + w_2 + w_1 - \frac{\left(\alpha b w_2\right)^2}{\lambda_1 + b^2 w_2}\right] \\
&= \frac{\sigma_b^2}{2} w_2 \left[\left(G_1\right)^2 - \frac{w_2\left(\alpha + 2bG_1\right)^2}{\lambda_1 + b^2 w_2}\right]\left(\alpha x_0\right)^2 + \frac{q}{2}\left(w_2 + k_1^{xx}\right) \qquad \text{(I-8)}
\end{aligned}
$$

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

$$
\begin{aligned}
\nu_1 &= \alpha\left(1 - \frac{b^2 w_2}{\lambda_1 + b^2 w_2}\right) \\
&= \alpha + bG_1 \qquad \text{(I-9)} \\
\nu_2 &= \left(-\frac{1}{\lambda_1 + b^2 w_2}\right)\alpha^2 b w_2 x_0 \\
&= G_1 \alpha x_0 \qquad \text{(I-10)} \\
\nu_3 &= \alpha^2 x_0 \qquad \text{(I-11)}
\end{aligned}
$$

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

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

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

in the two-period horizon, and δ_1 in equation (I-8) can be rearranged as

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

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 J 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 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 (3-1), takes the form

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

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

$$\begin{aligned} u_{0,1} &= \left(-\frac{\alpha b w_2}{\lambda_1 + b^2 w_2} \right) (b u_0 + \alpha x_0) \\ &= G_1 x_{0,1} \end{aligned} \tag{J-4}$$

and

$$\begin{aligned} x_{0,2} &= b \left(\alpha - b \frac{\alpha b w_2}{\lambda_1 + b^2 w_2} \right) u_0 + \alpha^2 x_0 + b \left(-\frac{\alpha b w_2}{\lambda_1 + b^2 w_2} \right) \alpha x_0 \\ &= (\alpha + b G_1) x_{0,1} \end{aligned} \tag{J-5}$$

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 ρk_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.

References

Amman, H. M. and Kendrick, D. A.: 1995, Nonconvexities in stochastic control models, *International Economic Review* **36**, 455–475.

Amman, H. M. and Neudecker, H.: 1997, Numerical solution methods of the algebraic matrix riccati equation, *Journal of Economic Dynamics and Control* **21**, 363–370.

Amman, H. and Tucci, M.: 2017, How active is active learning: value function method vs an approximation, *Working paper*, Dipartimento di Economia Politica e Statistica, Università di Siena.

Beck, G. and Wieland, V.: 2002, Learning and control in a changing economic environment, *Journal of Economic Dynamics and Control* **26**, 1359–1377.

Hansen, L. P. and Sargent, T. J.: 2007, *Robustness*, Princeton University Press, Princeton, NJ.

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

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

Koning de, W. L.: 1982, Infinite horizon optimal control of linear discrete time systems with stochastic parameters, *Automatica* **18**, 443–453.

Tse, E. and Bar-Shalom, Y.: 1973, An actively adaptive control for linear systems with random parameters, *IEEE Transactions on Automatic Control* **18**, 109–117.

Tucci, M. P.: 2004, *The Rational Expectation Hypothesis, Time-varying Parameters and Adaptive Control*, Springer, Dordrecht, the Netherlands.

Tucci, M. P., Kendrick, D. A. and Amman, H. M.: 2010, The parameter set in an adaptive control Monte Carlo experiment: Some considerations, *Journal of Economic Dynamics and Control* **34**, 1531–1549.