# Comparison of policy functions from optimal learning and adaptive control frameworks

Amman, H.M.; Kendrick, D.A.

*Published in:*
Computational Management Science

*DOI:*
10.1007/s10287-014-0215-9

# Comparison of Policy Functions from the Optimal Learning and Adaptive Control Frameworks

Hans M. Amman
Utrecht School of Economics
Utrecht University
Heidelberglaan 8
3584 CS Utrecht, Netherlands
h.m.amman@uu.nl

David A. Kendrick
Department of Economics
University of Texas
Austin, Texas 78712
USA
kendrick@eco.utexas.edu

January 24, 2015

**Abstract**

In this paper we turn our attention to comparing the policy function obtained by Beck and Wieland (2002) to the one obtained with adaptive control methods. It is an integral part of the *optimal learning* method used by Beck and Wieland to obtain a policy function that provides the optimal control as a feedback function of the state of the system. However, computing this function is not necessary when doing Monte Carlo experiments with adaptive control methods. Therefore, we have modified our software in order to obtain the policy function for comparison to the BW results.

***Keywords*** Active learning, dual control, optimal experimentation, stochastic optimization, time-varying parameters, numerical experiments.

***JEL Classification***: C63, E61.

# 1 Preamble

It is a great pleasure for us to be able to contribute this paper to the special issue of Computational Management Science honoring Berc Rustem. Berc was one of the first control engineers to take an active interest in the possible applications of optimal control methods to the stabilization of macroeconometric models. He was an enthusiastic member of a group pulled together at Imperial College by Professor John Westcott with support from a member of Parliament, Jeremy Bray. Jeremy and John had the novel idea that one could develop small optimal control models to help in the stabilization process for the large macro econometric models of the 1970's. Berc was the youngest member of this group and therefore the one who actually wrote much of the computer code that was developed by this pioneering project.

From this auspicious beginning, Berc has gone on the become a major contributor to the use of optimal control methods in economics – especially in the field of robust control where his John Wiley and Princeton University Press books are early and outstanding contributions to a subfield that has gained widespread recognition and use in macroeconomics. Also, Berc has generated enormous externalities of the type that are so valuable to all of us in the academic community. He has not only made many important contributions of his own but through his editorship for many years of the *Journal of Economic Dynamics and Control* and more recently of *Automatica* and *Computational Management Science* he has enable a very large community of scholars to see their work improved while on the path to making it widely available through publication.

# 2 Introduction

This paper is a continuation of work in the Methods Comparison Project[1] to compare various methods of solving *optimal experimentation*, *adaptive (dual) control* or *optimal learning* models, as the subject has been called by various authors. In essence, the methods consider dynamic stochastic models in which the control variables can be used not only to guide the system in desired directions but also to improve the accuracy of estimates of parameters in the models. Thus there is a tradeoff in which experimentation or perturbation of the control variables early in time detracts from reaching current goals but leads to learning or improved parameter estimates and thus improved performance of the system later in time - hence the dual nature of the control.

[1] Currently there are three groups involved in this project (1) Volker Wieland and Gunter Beck, (2) Thomas Cosimano and Michael Gapen and (3) Hans Amman, David Kendrick and Marco Tucci.

The optimal experimentation approach uses perturbation methods, see Cosimano (2008) and Cosimano and Gapen (2005a, 2005b), which are applied in the neighborhood of the augmented linear regulator problems as discussed by Hansen and Sargent (2004). The adaptive or dual control approach, see Kendrick (1981, 2002), Amman (1996) and Tucci (2004), uses methods that draw on earlier work in the engineering literature by Tse and Bar-Shalom (1973). The optimal learning approach uses numerical approximation of the optimal decision rule, see Wieland (2000a, 2000b), in methods that are related to earlier work by Prescott (1972), Taylor (1974) and Kiefer (1989).

In previous work in this project we have compared the mathematical results from the adaptive control approach to those obtained by Beck and Wieland (2002) in Kendrick and Tucci (2006). Also, we have examined the properties of the Beck and Wieland model using the DualPC software in Amman, Kendrick and Tucci (2008) and the problems caused by nonconvexities in this model in Tucci, Kendrick and Amman (2007).

In this paper we turn our attention to comparing the policy function obtained by Beck and Wieland 2002) to the one obtained with adaptive control methods. It is an integral part of the *optimal learning* method used by Beck and Wieland (BW) to obtain a policy function that provides the optimal control as a feedback function of the state of the system. However, computing this function is not necessary when doing Monte Carlo experiments with adaptive control methods. Therefore, we have modified our software in order to obtain the policy function for comparison to the BW results. To facilitate the description of our procedures we provide here a description of the BW model first in their notation and then in the adaptive control notation. This is done in the following two sections.
In doing the comparison we have employed two variants of adaptive control methods – the first based on the DualPC software, Amman and Kendrick (1999b), and the second using a MATLAB program with a parameterized cost-to-go function for adaptive control following the method outlined in the Amman and Ken–drick (1995) paper and the extension of these results in Tucci, Kendrick and Amman (2007). After describing both of these methods we will present in Section 7 of the paper a comparison of the policy function results obtained with (1) these two methods and (2) the Beck and Wieland method.

## 3 The Beck-Wieland Model in Wieland's Notation

The Beck-Wieland model is a one-state, one-control model with a single time varying parameter. The system equation is

$$x_{t+1} = \gamma x_t + \beta_t u_t + \alpha + \varepsilon_t \tag{3.1}$$

where $x_t$ is a state variable, $u_t$ a control variable, $\beta_t$ a stochastic time-varying parameter, $\gamma$ and $\alpha$ are constant coefficients and the identically and independently distributed (iid) random noise term $\varepsilon_t \sim N(0, \sigma_\varepsilon)$.
The initial condition for the system equation is $x_0$, where $x_0$ is the initial state variable. The control variable in equation (3.1) in the Beck-Wieland model has the time subscript $t+1$ rather than $t$; however, we follow here the convention in the engineering literature in which the control variable action in period $t$ affects the system in period $t+1$. Also, the idd noise term $\varepsilon_t$ here has the subscript $t$ rather than the $t+1$ used in BW.
The time-varying parameter equation is

$$\beta_{t+1} = \beta_t + \eta_t \tag{3.2}$$

where $\eta_t$ is an idd additive noise term with $\eta_t \sim N(0, \sigma_\eta)$ and with the initial value $\beta_0$ of the time-varying parameter. Also, the idd noise term here has the subscript $t$ rather than the $t+1$ used in BW.
The criterion function is a quadratic tracking equation where the goal is to find the minimum over the controls $\{u_t\}_{t=0}^{N-1}$ of

$$J = E\left\{\delta^N (x_N - \tilde{x})^2 + \sum_{t=0}^{N-1} \delta^t \left[(x_t - \tilde{x})^2 + \omega (u_t - \tilde{u})^2\right]\right\} \tag{3.3}$$

where $J$ is the criterion value, $E$ the expectations operator, $\delta \in < 0, 1]$ the discount factor, $\tilde{x}$ is the desired state variable, $\omega$ the (relative) weight on deviations of control variables from targets, $\tilde{u}$ the desired control variable.
The criterion function is over a finite horizon here in contrast to Beck and Wieland where it is infinite horizon. Also, the tracking function for the last time period $N$is separated here to indicate that the control variables are optimized only through period $N-1$.
For their numerical experiments Beck and Wieland use the following values $\gamma = 1$, $\alpha = 0$, $\tilde{x} = 0$, $\tilde{u} = 0$, $\omega = 0$, $\delta = 0.95$, $\sigma_\varepsilon = 1$, $\sigma_\eta = 0$, and the initial conditions $x_0 = 0$, $b_0 = -0.5$, $v_0^b = 0.25$.
The symbol $b$ is used to indicate the estimates of the parameter $\beta_t$. Actually, the initial condition $x_0 \in [-3, 3]$ in Figures 1 and 2 in the Beck and Wieland paper and we will use those same ranges in doing the comparisons in this paper.

# 4 The Beck and Wieland Model in Kendrick's Notation

The model in Kendrick (2002) that most closely approximates the Beck and Wieland model, is the one in Chapter 10 since that model includes time-

varying parameters. In addition, some use will be made of the notation in Amman and Kendrick (1999a) because that paper includes the discounting that is used in the BW model but is not present in the Ch. 10 model.
The systems equations in the Ch. 10 model in equation (10.7) are

$$x_{t+1} = A_t(\theta_t)\, x_t + B_t(\theta_t)\, u_t + c_t(\theta_t) + v_t \tag{4.1}$$

where $t \in [0, N-1]$ is the time index, $x_t \in \Re^{(n \times 1)}$ the state vector, $u_t \in \Re^{(m \times 1)}$ the control vector, $v_t \in \Re^{(n \times 1)}$ the idd vector of additive noise terms, $A_t(\theta_t) \in \Re^{(n \times n)}$ the state vector coefficient matrix, $B_t(\theta_t) \in \Re^{(n \times m)}$ the control vector coefficient matrix, $c_t(\theta_t) \in \Re^{(n \times 1)}$ the exogenous coefficient vector, $\theta_t \in \Re^{(s \times 1)}$ vector containing the subset of the coefficients in $A_t(\theta_t)$, $B_t(\theta_t)$ and $c_t(\theta_t)$ that are treated as uncertain.
The matrix $A_t(\theta_t)$ is a function of the subset of the uncertain coefficients in $\theta_t$ which come from that matrix. The same applies to $B_t(\theta_t)$ and $c_t(\theta_t)$.
For the BW model there is a single state variable and a single control variable so these two vectors each have a single element. Also there is single uncertain coefficient so $\theta_t$ is

$$\theta_t = \beta_t \tag{4.2}$$

Comparison of equation (4.1) to the BW model system equation equation(3.1) yields

$$A = \gamma = 1 \qquad B_t = \beta_t \qquad c = \alpha = 0 \qquad v_t = \varepsilon_t$$

Because this paper draws on mathematics from two different sources we will occasionally encounter cases where the same symbol is used for different purposes in the two sources. When this occurs we will rely on the context to communicate the differences, for example in equation (4.1) and in the equation above, $v_t$ is used to indicate the idd additive noise term for the systems equations in the Kendrick framework. In contrast, $v_t$ is used in the Beck and Wieland paper to indicate the variance of the estimate of the $\beta_t$ parameter.
Also, we will be using equation numbers from multiple sources and, here also, we rely on context rather than special fonts to distinguish the sources.
The measurement equation in the Ch. 10 model in equation (10.8) is

$$y_t = H_t x_t + w_t \tag{4.3}$$

where $y_t \in \Re^{(r \times 1)}$ is a measurement vector, $H_t \in \Re^{(r \times n)}$ a measurement coefficient matrix $w_t \in \Re^{(r \times 1)}$ an idd measurement noise vector.
Though Wieland has included measurement errors in one of his papers with Coenen (viz. Coenen and Wieland (2001)) those errors are not included in the BW model; therefore we have

$$H = I \qquad \forall t \ w_t = 0$$

The time-varying parameter equation in the Ch. 10 model, i.e. equation (10.9), is

$$\theta_{t+1} = D_t \theta_t + \eta_t \tag{4.4}$$

where $D_t \in \Re^{(s \times s)}$ the parameter evolution matrix, $\eta_t \in \Re^{(s \times 1)}$ the idd additive noise term of the time-varying parameter. For more general forms of equation (4.4) in the adaptive control context, including the *return to normality* model, see Tucci (2004) page 17.
In the BW model there is a single time-varying parameter, also the coefficient D is one, thus

$$D = I$$

Also, the idd additive noise term in the time-varying parameter equation (4.4) is distributed

$$\eta_t \sim N\left(0, \sigma_\eta\right) \text{ with } \sigma_\eta = 0 \tag{4.5}$$

The initial conditions for the systems equation (3.1) and the parameter evolution equations (4.4) in the Ch. 10 model are

$$x_0 \sim N\left(\hat{x}_{0|0}, \hat{\Sigma}^{xx}_{0|0}\right) \qquad \theta_0 \sim N\left(\hat{\theta}_{0|0}, \hat{\Sigma}^{\theta\theta}_{0|0}\right) \tag{4.6}$$

Since there is no measurement error in the BW model and since the original state is assumed to be zero in the base run we have

$$\hat{x}_{0|0} = 0 \qquad \hat{\Sigma}^{xx}_{0|0} = 0$$

However, in this paper we will solve the BW model for values of the initial condition $\hat{x}_{0|0} \in [-3, 3]$.
Also in the BW model the initial value of the time-varying coefficient is set to -0.5 and the initial variance of that coefficient is set to 0.25 so we have

$$\hat{\theta}_{0|0} = b_0 = -0.5 \qquad \hat{\Sigma}^{\theta\theta}_{0|0} = v^b_0 = 0.25$$

The idd additive noise terms for the systems, measurement and parameter evolution equations in the Ch. 10 model are distributed

$$v_t \sim N\left(0, Q\right) \qquad w_t \sim N\left(0, R\right) \qquad \eta_t \sim N\left(0, \Gamma\right) \tag{4.7}$$

These terms in the BW model are

$$Q = \sigma_\varepsilon = 1.0 \qquad R = 0 \qquad \Gamma = \sigma_\eta = 0$$

Thus there is a variance of one for the idd additive noise term in the systems equations, there is no measurement error and the additive noise term in the parameter evolution equation is set to zero. This last assumption is surprising so we may be misinterpreting the BW paper at this point.
The criterion function in the Ch. 10 model is for a finite horizon model. That criterion with the addition of discounting as in Amman and Kendrick (1999a) may be written as

$$J = E\left\{\delta^N L_N\left(x_N\right) + \sum_{t=0}^{N-1} \delta^t L_t\left(x_t, u_t\right)\right\} \tag{4.8}$$

where $J \in \Re$ is the criterion value, $E$ the expectations operator, $\delta \in< 0, 1]$ the discount factor, $L_N \in \Re$ the criterion function for the terminal period $N$, $x_N \in \Re^{(n\times 1)}$ the state vector for the terminal period $N$, $L_t \in \Re$ the criterion function for period $t$, $x_t \in \Re^{(n\times 1)}$ the state vector for period $t$ and $u_t \in \Re^{(m\times 1)}$ the control vector for period $t$.
The two terms on the right-hand side of equation (4.8) are defined as

$$L_N(x_N) = \frac{1}{2}(x_N - \tilde{x}_N)' W_N (x_N - \tilde{x}_N) \tag{4.9}$$

and

$$\begin{aligned} L_t(x_t, u_t) = \frac{1}{2}\Big[ & (x_t - \tilde{x}_t)' W_t (x_t - \tilde{x}_t) + \\ & (x_t - \tilde{x}_t)' F_t (u_t - \tilde{u}_t) + (u_t - \tilde{u}_t)' \Lambda_t (u_t - \tilde{u}_t)\Big] \end{aligned} \tag{4.10}$$

where $\tilde{x}_N \in \Re^{(n\times 1)}$ the desired state vector for terminal period $N$, $W_N \in \Re^{(n\times n)}$ the symmetric state variable penalty matrix for terminal period $N$, $\tilde{x}_t \in \Re^{(n\times 1)}$ the desired state vector for period $t$, $\tilde{u}_t \in \Re^{(m\times 1)}$ the desired control vector for period $t$, $W_t \in \Re^{(n\times n)}$ the symmetric state variable penalty matrix for period $t$, $F_t \in \Re^{(n\times m)}$ the penalty matrix on state-control variable deviations for period $t$, $\Lambda_t \in \Re^{(m\times m)}$ the symmetric control variable penalty matrix for period $t$.
The comparison of equations (4.9) and (4.10) to the Beck and Wieland model in equation (3.3) above and the use of the parameter values specified in Figure 1 of their article yields

$$W_N = 1 \quad \forall t\ W_t = 1 \quad \forall t\ F_t = 0 \quad \forall t\ \Lambda_t = \omega = 0 \quad \delta = 0.95$$

Thus there is a weight of one on the state variable deviations, no weight on the cross terms, and a weight of zero on the control variable deviations. Also, the discount factor is set at 0.95 so the discount rate is 0.05.

This completes the description of the model in Wieland's notation and in Kendrick's notation. This notation can now be used to discuss the two procedures we have used in making comparisons of the policy functions.

# 5 Modification of the DualPC Software

In summary, from above the adaptive control problem is to find the control variables $(u_0, u_1, \cdots, u_{N-1})$that minimize the criterion function

$$J = E\left\{\delta^N L_N\left(x_N\right) + \sum_{t=0}^{N-1} \delta^t L_t\left(x_t, u_t\right)\right\} \tag{5.1}$$

subject to the systems equations

$$x_{t+1} = A_t\left(\theta_t\right) x_t + B_t\left(\theta_t\right) u_t + c_t\left(\theta_t\right) + v_t \tag{5.2}$$

the measurement equations

$$y_t = H_t x_t + w_t \tag{5.3}$$

and the parameter evolution equations

$$\theta_{t+1} = D_t \theta_t + \eta_t \tag{5.4}$$

from the initial conditions

$$x_0 \sim N\left(\hat{x}_{0|0}, \hat{\Sigma}_{0|0}^{xx}\right) \qquad \theta_0 \sim N\left(\hat{\theta}_{0|0}, \hat{\Sigma}_{0|0}^{\theta\theta}\right) \tag{5.5}$$

This is the problem that we solve with the DualPC software. However, the policy function in Figure 1 of the Beck and Wieland paper is a feedback function of the form

$$u_0 = f\left(x_0\right) \tag{5.6}$$

where $u_0$ is the optimal control vector in period 0, $x_0$ the state vector in period 0 and this function is not automatically calculated by previous versions of the DualPC software.

Therefore to make the comparison it was necessary first to define a discrete grid over the initial state vector, $x_0$. In the BW model there is only one state variable and the grid is defined in their Figure 1 over the range $[-3, 3]$ at roughly 45 points so the spacing between points is about 0.2. We used the same range but with a finer grid of about 240 points with a spacing of 0.025 between points.

Next we created an outside for loop in DualPC over each of these 240 elements so that the problem in equations (5.1) - (5.5) is solved repeatedly. In each pass through the loop we stored the optimal control for period zero, $u_0$, that corresponded to the grid value for $x_0$ in that pass through the loop.

Since the BW model has a single control variable it was necessary to store only one value in each pass through the loop.
Since there is no measurement error in the BW model $x_0$ in that model is not random and none of the other random elements in the model occur before the computation of the zero period optimal control so it was necessary to turn off all of the Monte Carlo capabilities of the DualPC software when doing these calculations.
Table 1 below shows the parameter values that we used for the base run and which correspond to the parameter values that we believe underlie the results in Figure 1 of the Beck and Wieland paper.

Table 5.1:
Parameter Values Used in the Base Run

| Beck & Wieland Notation | Kendrick Notation | Value of Parameter |
|---|---|---|
| $\gamma$ | $A$ | 1.0 |
| $b_0$ | $B$ | -0.5 |
| $\alpha$ | $c$ | 0.0 |
| $\nu_0^b$ | $\Sigma_0^{\theta\theta}$ | 0.50 |
| 1 | $W$ | 1.00 |
| $\omega$ | $\Lambda$ | 0.0001 |
| $\tilde{x}$ | $\tilde{x}$ | 0.0 |
| $\tilde{u}$ | $\tilde{u}$ | 0.0 |
| $\sigma_\varepsilon$ | q | 1.0 |

The results of our calculations are shown below in Figure 5.1. Our policy function for the adaptive control case has the same characteristic S shape as the *optimal* function in Figure 1 of BW though our function is somewhat smoother than the one in BW in part because we used a finer grid for $x_0$. The values for the function in our Figure 5.1 above are close to those in Figure 1 of the BW paper but are not identical.
Also, we have prepared a second plot that is shown below in Figure 5.2 which compares the Dual solution, Cautionary solution and Certainty Equivalence solution for the policy functions. The order and shape of all three of these functions agrees with the results in Figure 1 of the Beck and Wieland paper. We will return to the comparison of our results to the BW results later; however, first it is useful to report on a second set of calculations which we did as a check on the ones we made with the DualPC software.

Figure 5.1:

Policy Function for the BW Model from the DualPC Software

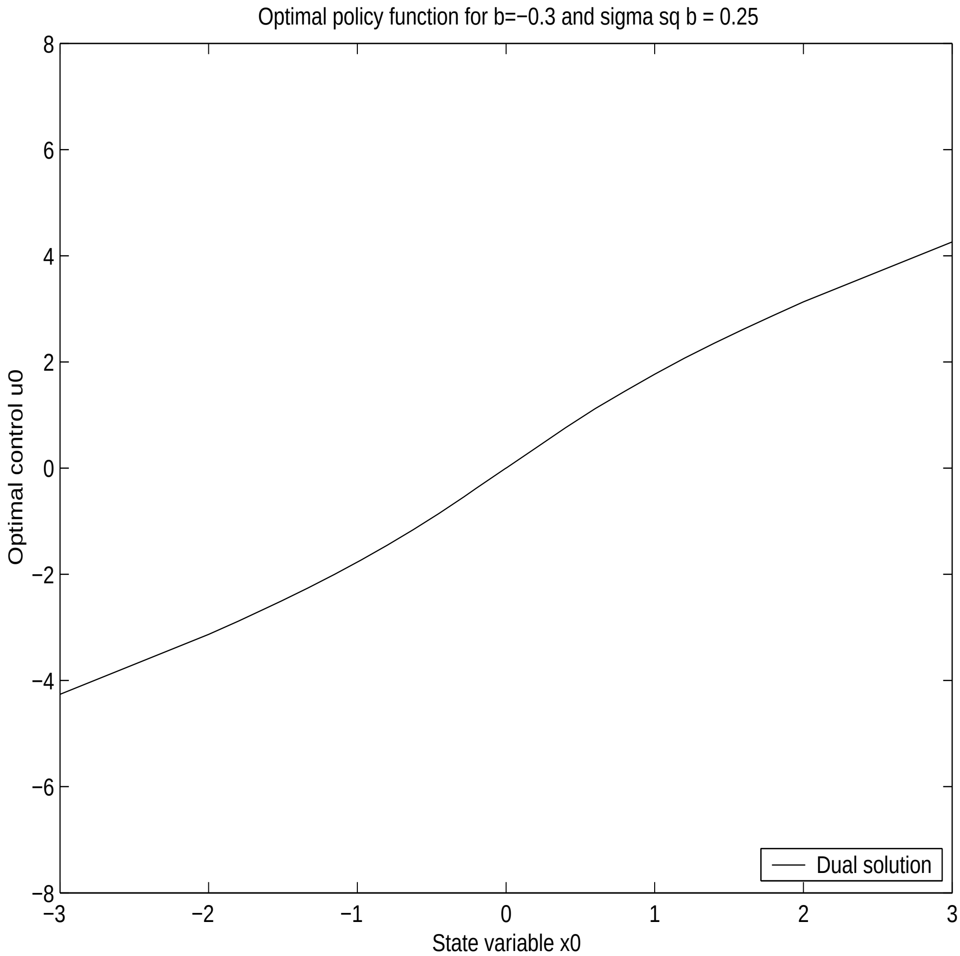
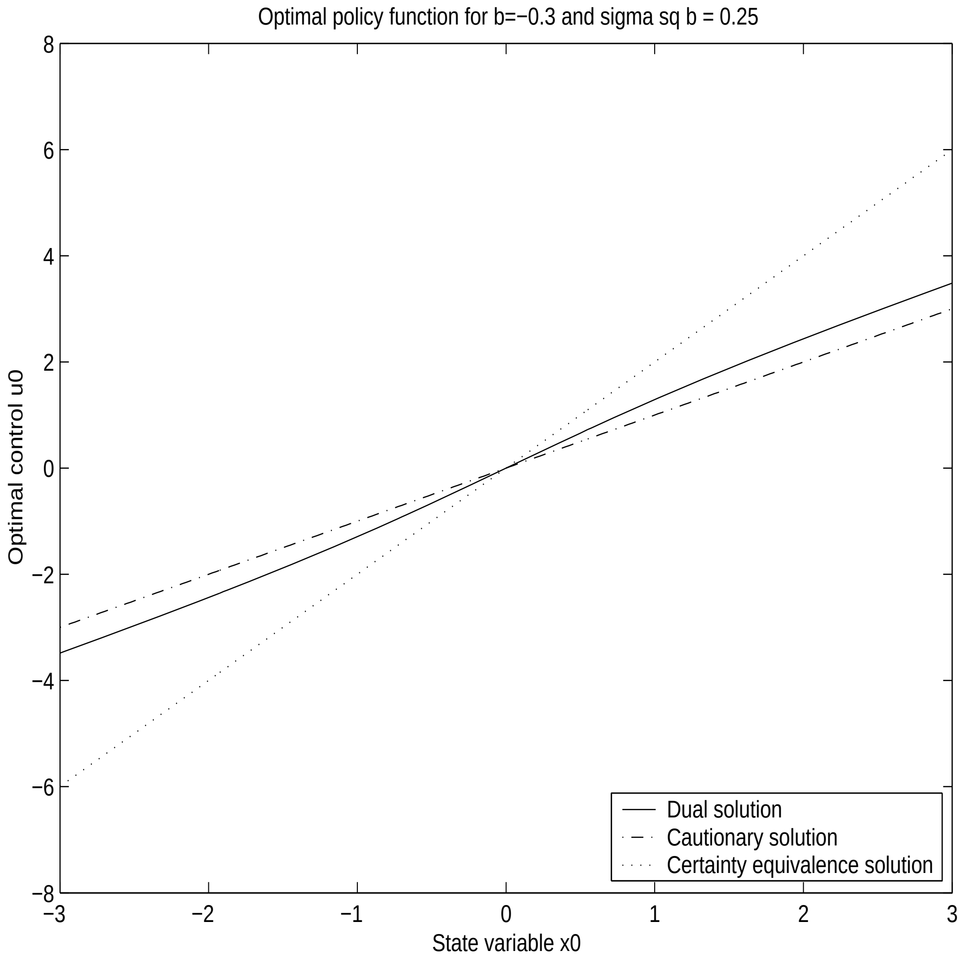


Figure 5.2:

Dual, Cautionary and Certainty Equivalence Policy Functions



# 6 A Parameterized Cost-To-Go Function for Adaptive Control

Since the Beck and Wieland model has a single control variable and a single state variable it is simple enough that it is possible to take an entirely different, and simpler, approach to calculating the policy function. This approach comes from earlier work we did in Amman and Kendrick (1995) when we were analyzing the question of whether or not the cost-to-go function in adaptive control problems was sometimes characterized by non-convexities. In that paper we were able to use a set of parameters from the system equation (4.1) like "$a$" for the state variable matrix, "$b$" for the control variable matrix, "$c$" of the constant vector in the systems equations and "$x_0$" for the initial condition of the system equation. These parameters where then substituted into the cost-to-go function and given a base set of values. This enabled us to obtain the cost-to-go function

$$J_N = f\left(u_0\right) \tag{6.1}$$

where

$$\begin{aligned} J_N &= \text{the cost-to-go with } N \text{ periods to go} \\ u_0 &= \text{the initial period control variable} \end{aligned}$$

so that we could analyze its properties. Indeed when we did this we confirmed that non-convexities would occur in the function at times and it was therefore necessary when solving adaptive control problems to employ global optimization methods which searched over the various local optima.

Recently we have returned to this subject while trying to understand why we had a substantial number of outliers when we did Monte Carlo experiments with adaptive control on the Beck and Wieland model. In this work, which is reported in Tucci, Kendrick and Amman (2007), we found that the outliers were caused by non-convexities which occurred when some combinations of parameter values were generated by the Monte Carlo procedures. A by-product of this work was a spreadsheet that was developed by Marco Tucci to compute the cost-to-go function for various set of parameter values.

When we began the present work on computing the policy function for the Beck and Wieland model we realized that we could build on the spreadsheet and add to it an optimization search over the cost-to-go function so as to obtain the policy function

$$u_0 = f\left(x_0\right) \tag{6.2}$$

for the initial period optimal control. However before progressing far with this we realized that these kinds of calculations could be done in a more

straightforward and easier-to-check manner in a MATLAB program than in an Excel spreadsheet.
Therefore, we developed a MATLAB program that had an outside loop over each of the grid points for $x_0$ as described in the previous section. This was around an inside loop which was in turn across a set of grid points for the control variable $u_0$. This enabled us to search for local optima on each pass through the inside loop while using the outside loop to get the points one-by-one for the policy function (6.2). The details of these calculations are described in Appendix A.
The result of this work was a very efficient way to compute the policy function for the Beck and Wieland model with relatively transparent computations. While we realized that this procedure would be difficult to generalize to models with many state and control variables it worked well for a model with a single state and a single control variable and thus provided a good check on the results obtained from the much more elaborate calculations done with the modified version of the DualPC program.
Indeed we were pleased when we discovered that this approach gave the same values for the policy function of the BW model as the results obtained from the modified DualPC program. This left us free with some confidence to move on to a comparison of these results to those in Figure 1 of the Beck and Wieland paper.

# 7 Comparison of the DualPC and Beck and Wieland Policy Functions

In order to make a comparison between our results and those of Beck and Wieland we went to Volker Wieland's web site at
http://www.volkerwieland.com
and downloaded his Fortran code "A Numerical Dynamic Programming Algorithm for Solving the Optimal Learning Problem". We then ran this program with the parameter values shown in Table 2.
Table 2 Parameter Values Used in Wieland's Program
These values all correspond to those used in the base run as described in Table 1 except for (1) the initial value of the $\boldsymbol{B}$ parameter, $b_0$, which is set at -0.3 here instead of -0.5 and (2) the variance of the estimate of that parameter which is set at 0.25 here instead of 0.50.
We used the same grid for $x_0$ as was used in the Wieland program namely

[-5 -3 -2 -1.8 -1.6 -1.4 -1.2 -1.0 -0.8 -0.6 -0.4 -0.2 -0.0001

0.0001 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 3.0 5.0]

Notice that this grid is not uniform but rather is less widely spaced around zero than at the lower and upper extremes.

| Beck & Wieland Notation | Kendrick Notation | Value of Parameter |
|---|---|---|
| $\gamma$ | $A$ | 1.0 |
| $b_0$ | $B$ | -0.3 |
| $\alpha$ | $c$ | 0.0 |
| $\nu_0^b$ | $\Sigma_0^{\theta\theta}$ | 0.25 |
| 1 | $W$ | 1.00 |
| $\omega$ | $\Lambda$ | 0.0001 |
| $\tilde{x}$ | $\tilde{x}$ | 0.0 |
| $\tilde{u}$ | $\tilde{u}$ | 0.0 |
| $\sigma_\varepsilon$ | q | 1.0 |

We put the parameter values from Table 2 into our MATLAB program for the second of the two methods described above, namely the parameterized cost-to-go function approach. We also imported the results from running Wieland's Fortran code into the MATLAB program and then plotted the results from the Dual and BW methods as shown in Figure 7.1 below.
The two solutions are close for values of $x_0$ below -2 and for values above 2. In the range for $x_0$ between -2 and 2 the BW solution has more aggressive control values than the ones in the Dual solution. We do not yet know why this occurs but are exploring several possible explanations. So, in summary, the Dual and BW methods yield very similar policy functions for the BW model except in the range of values of $x_0$ around zero.

# 8 Conclusions

We have developed two methods for using our adaptive control software to compute the policy function for the Beck and Wieland model. These two methods give identical values for the function so they check against one another. The shape of the policy function they yield is closely similar to the BW policy function and the numerical values of the function are close for values of $x_0$ below -2 and above 2 but differ somewhat for values of $x_0 \in [-2, 2]$. We are investigating why the differences occur in this range.

## References


Amman, H. M.: 1996, Numerical methods for linear-quadratic models, *in* H. M. Amman, D. A. Kendrick and J. Rust (eds), *Handbook of Computational Economics*, Vol. 13 of *Handbook in Economics*, North-Holland Publishers (Elsevier), Amsterdam, the Netherlands, pp. 579–618.

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

Amman, H. M. and Kendrick, D. A.: 1999a, Matrix methods for solving nonlinear dynamic optimization models, *in* R. J. Heijmans, D. S. G. Pollock and A. Satorra (eds), *Innovations in Multivariate Statistical Analysis*, Vol. 30 of *Advanced Studies in Theoretical and Applied Econometrics*, Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 257–276.

Amman, H. M. and Kendrick, D. A.: 1999b, The DualI/DualPC software for optimal control models: User's guide, *Working paper*, Center for Applied Research in Economics, University of Texas, Austin, Texas 78712, USA.

Amman, H. M., Kendrick, D. A. and Tucci, M. P.: 2008, Solving the Beck and Wieland model with optimal experimentation in DUALPC, *Automatica* **44**, 1504–1510.

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

Coenen, G. and Wieland, V.: 2001, Evaluating information variables for monetary policy in a noisy economic environment. European Central Bank, presented at the Seventh International Confernce of the Society for Computational Economics Conference,Yale University, New Haven, USA.

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.

Cosimano, T. F. and Gapen, M. T.: 2005a, Program notes for optimal experimentation and the perturbation method in the neighborhood of the augumented linear regulator problem, *Working paper*, Department of Finance, University of Notre Dame, Notre Dame, Indiana, USA.

Cosimano, T. F. and Gapen, M. T.: 2005b, Recursive methods of dynamic linear economics and optimal experimentation using the perturbation method, *Working paper*, Department of Finance, University of Notre Dame, Notre Dame, Indiana, USA.

Hansen, L. P. and Sargent, T. J.: 2004, Recursive models of dyanmic linear economies. Department of Economics, Univesity of Chicago manuscript.

Kendrick, D. A.: 1981, *Stochastic control for economic models*, first edn, McGraw-Hill Book Company, New York, New York, USA. See also Kendrick (2002).

Kendrick, D. A.: 2002, Stochastic control for economic models. University of Texas.
**URL:** *http://www.utexas.edu/cola/depts/economics/faculty/dak2*

Kendrick, D. A. and Tucci, M. P.: 2006, The Beck and Wieland model in the adaptive control framework, *Working paper*, Department of Economics, University of Texas, Austin, Texas 78712, USA.

Kiefer, N.: 1989, A value function arising in the economics of information, *Journal of Economic Dynamics and Control* **13**, 201–223.

Prescott, E. C.: 1972, The multi-period control problem under uncertainty, *Econometrica* **40**, 1043–1058.

Taylor, J. B.: 1974, Asymptotic properties of multiperiod control rules in the linear regression model, *International Economic Review* **15**, 472–482.

Tse, E.: 1973, Further comments on adaptive stochastic control for a class of linear systems, *IEEE Transactions on Automatic Control* **18**, 324–326.

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.: 2007, The parameter set in an adaptive control Monte Carlo experiment: Some considerations, *Quaderni del Dipartimento di Economia Politica 507*, Università di Siena, Siena, Italy.

Wieland, V.: 2000a, Learning by doing and the value of optimal experimentation, *Journal of Economic Dynamics and Control* **24**, 501–543.

Wieland, V.: 2000b, Monetary policy, parameter uncertainty and optimal learning, *Journal of Monetary Economics* **46**, 199–228.

## Appendix A

Calculations for the *Parameterized Cost-To-Go* Approach

As is discussed in Tucci, Kendrick and Amman (2007) the cost-to-go in adaptive control problems can be divided into three terms, i.e.

$$J_N = J_{D,N} + J_{C,N} + J_{P,N} \tag{A-1}$$

where $J_N$ is the the cost-to-go with $N$ periods remaining, $J_{D,N}$ the deterministic component of the cost-to-go, $J_{C,N}$ the cautionary component of the cost-to-go and $J_{P,N}$ the probing component of the cost-to-go. Furthermore,

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

$$J_{C,2} = \delta_1 u_0^2 + \delta_2 u_0 + \delta_3 \tag{A-3}$$

$$J_{P,2} = \left(\frac{\sigma_b^2 q \varphi_1}{2}\right) \frac{(\varphi_2 u_0 + \varphi_3)^2}{\sigma_b^2 u_0^2 + q} \ . \tag{A-4}$$

and where the three sets of parameters $\psi, \delta$ and $\varphi$ are themselves functions of an underlying set of parameters $\nu$. All four sets of these parameters are in turn functions of the basic parameters of the Beck and Wieland model as defined in Sections 3 and 4. Also, notice that the deterministic and cautionary components are quadratic functions of the control variable in period zero, $u_0$, and the probing component is a function of the ratio of two quadratic functions in $u_0$.

The calculations of these three separate components of the cost-to-go are laid out in considerable detail in the Tucci, Kendrick and Amman (2007) paper. However, for our purposes here it is sufficient to know that many of these four sets of parameters $\psi$, $\delta$, $\varphi$ and $\nu$ are themselves functions of the initial state $x_0$ so we can rewrite the three equations in more general function form as

$$J_{D,N} = f_D\left(x_0, u_0\right) \tag{A-5}$$

$$J_{C,N} = f_C\left(x_0, u_0\right) \tag{A-6}$$

$$J_{P,N} = f_P\left(x_0, u_0\right) \tag{A-7}$$

and thus from equation (A-1) one can think of the total cost-to-go function itself a function of the initial state and the initial control variable, i.e.

$$J_N = f\left(x_0, u_0\right) \tag{A-8}$$

Our MATLAB program builds on the foundation of equation (A-8) to compute the policy function

$$u_0 = f(x_0) \tag{A-9}$$

This is done with a set of two `for` loops. The outside loop is over the grid for $x_0$. As was discussed in Section 5 the range for this grid is $[-3, 3]$ to follow the range used in Figure 5.1 of the Beck and Wieland paper. Also, the grid is set more finely that in the BW paper to provide a smoother function by spacing the grid elements 0.025 apart so as to create 240 points in the $x_0$ grid.

Between the outside and inside `for` loops are the calculations of the sets of parameters $\psi$, $\delta$, $\varphi$ and $\nu$ some of which are functions of $x_0$. Then the inside loop is over a grid for $u_0$. We use a grid search rather than a gradient method to find the optimal control that corresponds to each value of $x_0$ because the cost-to-go function (A-8) may be non-convex.

The range for the $u_0$ grid is $[-10, 10]$ with a spacing between grid points of $10^{-3}$. This is a very fine grid; however the evaluation of the cost-to-function (A-8) and the storage of this value as an element in a vector at each pass through the loop require very little computation.

After the end of the $u_0$ `for` loop we apply the MATLAB function `min` to the vector of values for $J_N$ corresponding to each of the $u_0$ grid values. The minimum value obtained from this operation then gives us the point of the policy function (A-9) corresponding to the current $x_0$ grid point. Therefore by the time we have passed through the outside loop over all the $x_0$ grid points we have completed the construction of the policy function (A-9) and can plot it.

Figure 7.1:

Comparison of the Dual and the BW Policy Functions