



UvA-DARE (Digital Academic Repository)

Learning about learning in dynamic economic models

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

Publication date

2014

Document Version

Author accepted manuscript

Published in

Handbook of computational economics. - Vol. 3

[Link to publication](#)

Citation for published version (APA):

Kendrick, D. A., Amman, H. M., & Tucci, M. P. (2014). Learning about learning in dynamic economic models. In K. Schmedders, & K. L. Judd (Eds.), *Handbook of computational economics. - Vol. 3* (pp. 1-35). (Handbooks in Economics). Elsevier, North-Holland.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Learning About Learning in Dynamic Economic Models ¹

David A. Kendrick, Hans M. Amman and Marco P. Tucci

Department of Economics, University of Texas, Austin, Texas 78712, USA.

*Utrecht School of Economics, Utrecht University, Heidelberglaan 8, 3584 CS
Utrecht, the Netherlands.*

*Dipartimento di Economia Politica, Università di Siena, Piazza S. Francesco, 7,
53100 Siena, Italy*

Abstract

This chapter of the *Handbook of Computational Economics* is mostly about research on *active* learning and is confined to discussion of learning in dynamic models in which the systems equations are linear, the criterion function is quadratic and the additive noise terms are Gaussian. Though there is much work on learning in more general systems, it is useful here to focus on models with these specifications since more general systems can be approximated in this way and since much of the early work on learning has been done with these quadratic-linear-gaussian systems.

We begin with what has been learned about learning in dynamic economic models in the last few decades. Then we progress to a discussion of what we hope to learn in the future from a new project that is just getting underway. However before doing either of these it is useful to provide a short description of the mathematical framework that will be used in the chapter.

Key words: Active learning, dual control, optimal experimentation, stochastic optimization, time-varying parameters, forward looking variables, numerical experiments.

JEL Classification: C63, E61.

Email addresses: kendrick@austin.utexas.edu (David A. Kendrick),
h.m.amman@uu.nl (Hans M. Amman), tucci@unisi.it (Marco P. Tucci).

¹ Corresponding author David A. Kendrick. We are indebted to a referee for many helpful comments and suggestions.

1 Introduction

It is common sense that one can learn about an economic system as time passes. One can observe the inputs and outputs of the system and make inferences about the relationship between the variables that characterize the system. An example is a macroeconomic model with inputs like government expenditure, taxes and the money supply and outputs like gross domestic product, inflation and unemployment. Another example is a microeconomic model with inputs like the price level and resources such as capital, labor and energy and outputs like production, sales and profit.

In the control theory framework one can model the inputs as control variables and the outputs as state variables and characterize the relationship between the two with the functional form of the system equations and the parameters of those equations. In this context previously obtained data may be used to calculate the means and variances of the estimates of the parameters and of the additive noise terms in the system equations. Learning then occurs as more observations are obtained with the passage of time and these observations are used to modify the means and to reduce the variances of the estimates of the parameters and of the additive noise terms. This is called *passive* learning because no deliberate attempt is made to increase the learning done in each period.

In contrast *active* learning occurs when the control variables are chosen in each period in such a way as to perturb the system and thus increase the speed of learning.² However, this is done only at some cost in moving the economic system away from the paths that would otherwise be followed

This chapter is mostly about research on *active* learning and is confined to discussion of learning in dynamic models in which the systems equations are linear, the criterion function is quadratic and the additive noise terms are Gaussian. Though there is much work on learning in more general systems, it is useful here to focus on models with these specifications since more general systems can be approximated in this way and since much of the early work on learning has been done with these quadratic-linear-gaussian systems.

We begin with what has been learned about learning in dynamic economic models in the last few decades. Then we progress to a discussion of what we hope to learn in the future from a new project that is just getting underway. However, before doing either of these it is useful to provide a short description of the mathematical framework that will be used in the chapter.

² A more recent term for active learning is *optimal experimentation*, see for instance Beck and Wieland (2002)

2 The Framework

The framework consists of two parts: an optimization part and a learning part. The optimization part of the framework consists of an objective that has to be minimized, *the criterion value*, and the constraints that bind this objective. In economics the constraints define the dynamics of the system (state). We start with the constraints.

As mentioned above, optimal control models, like those used in Kendrick (1980, 1981, 2002), have linear *system equations* of the form

$$x_{t+1} = A_t(\theta_t)x_t + B_t(\theta_t)u_t + c_t(\theta_t) + v_t \quad (2.1)$$

where $t \in \{0, \dots, N-1\}$ is the (discrete) time index, $x_t \in \mathbb{R}^{(n \times 1)}$ a state vector, $u_t \in \mathbb{R}^{(m \times 1)}$ a control vector, $v_t \in \mathbb{R}^{(n \times 1)}$ an independently and identically distributed (i.i.d.) additive noise term with $v_t \sim N(0, \Sigma_t^{vv})$, $A_t \in \mathbb{R}^{(n \times n)}$ a transition matrix, $B_t \in \mathbb{R}^{(n \times m)}$ a control coefficient matrix, and $c_t \in \mathbb{R}^{(n \times 1)}$ an exogenous variables vector. The vector $\theta_t \in \mathbb{R}^{(s \times 1)}$ is a vector containing the subset of the coefficients in A_t , B_t and c_t that are treated as uncertain. The matrix A_t is a function of the subset of the uncertain coefficients in θ_t which come from that matrix. The same applies to B_t and c_t .

This class of models permit consideration of situations that are common in economics where the state variables x_t may not all be directly observed or may not be observed without noise. The equations for this specification are called the *measurement equations* and may be written as

$$y_t = H_t x_t + \zeta_t \quad (2.2)$$

where $y_t \in \mathbb{R}^{(r \times 1)}$ is a measurement vector, $H_t \in \mathbb{R}^{(r \times n)}$ a measurement coefficient matrix, $\zeta_t \in \mathbb{R}^{(r \times 1)}$ a i.i.d. measurement noise vector with $\zeta_t \sim N(0, \Sigma_t^{\zeta\zeta})$.

The parameter estimates $\hat{\theta}_{t|t}$ of the true θ_t change over time³; while in most specifications of this class of models the true values of the parameters are assumed to remain constant. However, in some specifications the true values of the parameters are themselves assumed to be time varying. In these cases one can use parameter *evolution equations* of the form

$$\theta_{t+1} = D_t \theta_t + \eta_t \quad (2.3)$$

³ Following Bayesian literature, we make a distinction between prior estimates $\hat{\theta}_{t|t-1}$ and posterior estimates $\hat{\theta}_{t|t}$ of the true parameter vector θ_t .

where $D_t \in \mathbb{R}^{(s \times s)}$ the parameter evolution matrix, $\eta_t \in \mathbb{R}^{(s \times 1)}$ an iid additive noise term $\eta_t \sim N(0, \Sigma_t^{\eta\eta})$. When a more general form of equation (2.3) is needed, the law of motion of the time varying parameters can be written as

$$\theta_{t+1} = D_t \theta_t + (I - D_t) \bar{\theta} + \eta_t \quad (2.4)$$

where I is the identity matrix and $\bar{\theta}$ is the unconditional mean of the stochastic parameter. This is the form used by Tucci (1997, 1998, 2004) to model a wide variety of time varying parameter specifications. For example, when $\bar{\theta} = 0$, equation (2.4) reduces to equation (2.3). Also, when D_t and $\Sigma_t^{\eta\eta}$ are zero, θ_t becomes the usual time-invariant case. In contrast, if D_t is equal to zero, but $\Sigma_t^{\eta\eta}$ is nonzero, then equation (2.4) describes a random parameter, i.e. a parameter varying randomly about the fixed mean $\bar{\theta}$. If, on the other hand, $\bar{\theta}$ is equal to zero, equation (2.4) models a vector-autoregressive process of order one with mean zero. Also, random walk parameters may be modeled by setting D_t equal to one. Finally, equation (2.4) can be used to represent a lack of knowledge about the parameters. For example, when the true parameter associated with the control variable is constant, but unknown, then setting D_t equal to zero and $\Sigma_t^{\eta\eta}$ not equal to zero allows one to interpret θ_t as the time-varying estimate of the unknown constant parameter $\bar{\theta}$ based on observations through period t . When this is the case one can interpret $\Sigma_t^{\eta\eta}$ as the covariance of the estimates based on all information available at time t .

The initial conditions for the systems equation and the parameter evolution equations model are

$$\begin{aligned} x_0 &\sim N(\bar{x}_0, \Sigma_0^{xx}) \\ \theta_0 &\sim N(\bar{\theta}_0, \Sigma_0^{\theta\theta}) \end{aligned} \quad (2.5)$$

where E is the expectations operator. The expected states at $t = 0$, $\bar{x}_0 = E[x_0]$, $\bar{\theta}_0 = E[\theta_0]$, and their covariances Σ_0^{xx} , $\Sigma_0^{\theta\theta}$ are assumed to be known.⁴

The criterion function for this class of finite horizon model may be written with discounting terms as in Amman and Kendrick (1999b) as

$$J = E\{C_N\} = E\left\{ \delta^N L_N(x_N) + \sum_{t=0}^{N-1} \delta^t L_t(x_t, u_t) \right\} \quad (2.6)$$

⁴ For simulation purposes we usual start with (good) estimates of the variables $\hat{x}_{0|0}$, $\hat{\theta}_{0|0}$, $\hat{\Sigma}_{0|0}^{xx}$ and $\hat{\Sigma}_{0|0}^{\theta\theta}$ and update their prior and posterior estimates when new information comes available. See equations (2.12) and (2.14).

$J \in \mathbb{R}^+$ is the (scalar) value of the criterion, a discount factor $\delta \in \mathbb{R}^+$ usual defined on the interval $]0, 1]$, $L_t \in \mathbb{R}^+$ the criterion function and $N \in \mathbb{N}$ the terminal period.⁵ The two terms on the right-hand side of equation (2.6) are defined as

$$L_N(x_N) = \frac{1}{2}(x_N - \tilde{x}_N)'W_N(x_N - \tilde{x}_N) + w'_N(x_N - \tilde{x}_N) \quad (2.7)$$

and

$$\begin{aligned} L_t(x_t, u_t) &= \frac{1}{2}(x_t - \tilde{x}_t)'W_t(x_t - \tilde{x}_t) + (x_t - \tilde{x}_t)'F_t(u_t - \tilde{u}_t) \\ &\quad + \frac{1}{2}(u_t - \tilde{u}_t)'\Lambda_t(u_t - \tilde{u}_t) + w'_t(x_t - \tilde{x}_t) + \lambda'_t(u_t - \tilde{u}_t) \end{aligned} \quad (2.8)$$

where $\tilde{x}_t \in \mathbb{R}^{(n \times 1)}$ the desired state vector, $\tilde{u}_t \in \mathbb{R}^{(m \times 1)}$ the desired control vector. $W_t \in \mathbb{R}^{(n \times n)}$, $F_t \in \mathbb{R}^{(n \times m)}$, $\Lambda_t \in \mathbb{R}^{(m \times m)}$ are penalty matrices on the deviation of states and controls from their desired paths. W_t is assumed to semi-positive definite and Λ_t to be positive definite. $w_t \in \mathbb{R}^{(n \times 1)}$ and $\lambda_t \in \mathbb{R}^{(m \times 1)}$ are penalty vectors on the deviation of states and controls from their desired paths.

In summary, the stochastic optimal control model is specified to find the set of control variables $\{u_t\}_{t=0}^{N-1}$ that will minimize the criterion function (2.6) subject to (2.7)-(2.8), the system equations (2.1), the measurement equation (2.2) and the parameter evolution equation (2.3) or (2.4). Given the initial conditions (2.5).

This brings us to the second part of the framework. After the optimal control vector u_t^* is chosen in each time period the outputs of the system are observed and the means and variances of the parameters and of the state vectors are updated.

In this optimization procedure the original state vector x_t is augmented with the parameter vector θ_t to create a new state

⁵ A referee has noted that as $N \rightarrow \infty$ there should exist a unique time-independent function $J(x, u^*)$ (the value function or the expected cost-to-go) where u^* is the optimal decision rule (active learning). See Kiefer and Nyarko (1989) for a proof of the existence of this function.

$$z_t = \begin{bmatrix} x_t \\ \theta_t \end{bmatrix} \quad (2.9)$$

thus the corresponding covariance matrix for the augmented state is written as

$$\Sigma_t^{zz} = \begin{bmatrix} \Sigma_t^{xx} & \Sigma_t^{x\theta} \\ \Sigma_t^{\theta x} & \Sigma_t^{\theta\theta} \end{bmatrix} \quad (2.10)$$

where $\Sigma_t^{\theta\theta}$ is the covariance for the parameter estimates as previously defined. In general, Σ_t^{zz} is not known and has to be estimated. Hence, we have

$$\hat{\Sigma}_{t|t}^{zz} = \begin{bmatrix} \hat{\Sigma}_{t|t}^{xx} & \hat{\Sigma}_{t|t}^{x\theta} \\ \hat{\Sigma}_{t|t}^{\theta x} & \hat{\Sigma}_{t|t}^{\theta\theta} \end{bmatrix} \quad (2.11)$$

Also, for all four elements of the covariance matrix $\hat{\Sigma}_{t|t}^{zz}$ a distinction will be made between the current period covariance $\hat{\Sigma}_{t|t}^{zz}$, the projected covariance next period before new measurements are made $\hat{\Sigma}_{t+1|t}^{zz}$ and the projected variance next period after new measurements are made $\hat{\Sigma}_{t+1|t+1}^{zz}$.⁶

The estimation process used in Kendrick (1981, 2002) is the Kalman filter. The mathematical expression for the updating of the means of the uncertain parameters is⁷

$$\hat{\theta}_{t+1|t+1} = \hat{\theta}_{t+1|t} + \hat{\Sigma}_{t+1|t}^{\theta x} H'_{t+1} S_{t+1}^{-1} [y_{t+1} - H_{t+1} \hat{x}_{t+1|t}] \quad (2.12)$$

where

$$S_{t+1} = H_{t+1} \hat{\Sigma}_{t+1|t}^{xx} H'_{t+1} + \Sigma_{t+1}^{\zeta\zeta} \quad (2.13)$$

From these equations one can see that the change in the parameter estimate from the projection $\hat{\theta}_{t+1|t}$ to the post-measurement updated values $\hat{\theta}_{t+1|t+1}$ de-

⁶ Like in footnote (3), $\hat{\Sigma}_{t+1|t}^{zz}$ is called the prior estimate and $\hat{\Sigma}_{t+1|t+1}^{zz}$ the posterior estimate of the covariance matrix Σ_{t+1}^{zz} .

⁷ Equations (2.12), (2.13) and (2.14) are based on Kendrick (1981, 2002), equations (10.61), (10.68) and (10.69).

depends on the difference between the actual measurement y_{t+1} and the expected measurement $H_{t+1}\hat{x}_{t+1|t}$ via the measurement equation (2.2). From equation (2.12) one can see that the magnitude of this adjustment depends positively on the estimated covariance of the original state vector x_t and the parameter vector θ_t in $\hat{\Sigma}_{t+1|t}^{\theta x}$. Also, equation (2.13) shows that the magnitude of the adjustment depends inversely on the estimated degree of uncertainty about the original states $\hat{\Sigma}_{t+1|t}^{xx}$ and the variance of the i.i.d. additive noise term in the measurement equations, which is $\Sigma_{t+1}^{\zeta\zeta}$.

In parallel to the updating equation for the parameter means is the updating equation for the parameter covariances which is shown below

$$\hat{\Sigma}_{t+1|t+1}^{\theta\theta} = \hat{\Sigma}_{t+1|t}^{\theta\theta} - \hat{\Sigma}_{t+1|t}^{\theta x} H'_{t+1} S_{t+1}^{-1} H_{t+1} \hat{\Sigma}_{t+1|t}^{x\theta} \quad (2.14)$$

Note that the degree of learning in period t is represented by the decrease in the covariance from the prior projection $\hat{\Sigma}_{t+1|t}^{\theta\theta}$ to the posterior measurement update $\hat{\Sigma}_{t+1|t+1}^{\theta\theta}$. Furthermore, in equation (2.14), this is positively related to the covariance of the original state vector x_t and the parameter vector θ_t as represented by the estimated matrix $\hat{\Sigma}_{t+1|t}^{\theta x}$. Also, from (2.13) and (2.14) the learning is inversely related to the degree of uncertainty about the original states $\hat{\Sigma}_{t+1|t}^{xx}$ and the additive noise term in the measurement equations, that is $\Sigma_{t+1}^{\zeta\zeta}$.

This completes the discussion of the mathematical framework so, with this framework in mind, we turn to a discussion first of what has already been learned about learning in this type of dynamic economic models.

3 What We Have Learned

Research on passive and active learning stochastic control of economic models has been underway for more than thirty years dating back to the early work of Prescott (1972), MacRae (1972) and Taylor (1974) and including work by Norman (1976), Kendrick (1976, 1980, 1981), Easley and Kiefer (1988), Kiefer (1989), Kiefer and Nyarko (1989), Tucci (1989), Mizrach (1991), Aghion et al. (1991), Amman (1996), Amman and Kendrick (1995, 1997), Wieland (2000a, 2000b), Beck and Wieland (2002), Tucci (2004), Cosimano and Gapen (2005b, 2005a), Cosimano (2008), Tesfaselassie, Schaling and Eijffinger (2007) and others. What have we learned about learning from this research?

3.1 Active Perturbation

The most basic expectation about active learning comes in comparing it to passive learning. For example, compare Expected Optimal Feedback (*EOF*) with Dual Control (*DC*).⁸ In *EOF* the choice of control takes into account the uncertainty in the parameter estimates and may therefore exhibit a tendency toward caution as compared to methods that do not consider the parameter uncertainty. However, in *EOF* no consideration is given to the impact of the current choice of control on future decreases in the uncertainty of the parameter estimates. In contrast, in the *DC* method the choice of control in each period depends on projections of the future covariance matrices of the parameters $\hat{\Sigma}_{t|t}^{\theta\theta}$ for all future time periods.

Thus one expects that the controls variables will be more active in the early time periods in the active control (*DC*) solutions than in the passive control (*EOF*) solutions. Our experience is that this expectation proves to be true. In fact, it appears that the perturbations in the control variables with the *DC* method are used not to learn all parameters equally but rather to learn more quickly those parameter that are most crucial to the effective control of the system. Neither we, nor so far as we know, others have yet quantified these kinds of results; however, our causal observations of the models we have used suggests that this class of results holds.

3.2 Rapid Decrease in Parameter Variances in the First Few Periods

Also, we have observed that in most active learning solutions there is a very sharp decrease in the elements of the covariance matrix of the parameters, $\hat{\Sigma}_{t|t}^{\theta\theta}$, in the first few time periods. Of course, the speed of this learning is related to the cost of the perturbations in driving the state variables further away from their desired paths than would otherwise occur. If this cost is high, the learning will progress more slowly. In addition, when the cost of control is high, i.e. in situations where the penalty weights in the Λ matrix in the criterion function are relatively large, then the control will not be very active and the learning will be slower, i.e. the elements in the $\hat{\Sigma}_{t|t}^{\theta\theta}$ matrix will decline more slowly over time.

Conversely, if there is large initial uncertainty in the parameter estimates, $\hat{\Sigma}_{0|0}^{\theta\theta}$, then it is difficult to effectively control the system and there is reason to provide larger perturbations in order to quickly reduce this uncertainty.

⁸ For a classification of the various methods that are used in stochastic control of economic models and the nomenclature that is associated with these methods see Kendrick and Amman (2006).

Exactly how this tradeoff will play out in macroeconomic and microeconomic systems is a matter that deserves substantial attention in future research.

3.3 *Non-convexities*

Also, we have learned that the $\Sigma_t^{\theta\theta}$ matrices play a major role in determining whether or not there are nonconvexities in the cost-to-go function. As we have related elsewhere, Kendrick (2005), in the early work on dual control we did not expect that the cost-to-go function would exhibit nonconvexities and we did not make any allowance for this. However, we accidentally discovered that local optima existed in the cost-to-go functions, Kendrick (1978), Norman, Norman and Palash (1979), and this was confirmed by theoretical research by Mizrach (1991) and by numerical research by Amman and Kendrick (1995). Also, the presence of the nonconvexities has been confirmed by Wieland (2000a) using a different solution method, namely value function iteration.

Therefore as a guard against the non-convexity problem we have incorporated grid search methods into our DualPC software systems.⁹ However the unvarnished grid search system may be inefficient. This occurs because in our experience the nonconvexities are most likely to occur in the first few periods when the $\Sigma_t^{\theta\theta}$ matrix elements are relatively large. We have come to this realization in part by the way that the cost-to-go is divided into three components in the Tse and Bar-Shalom (1973) framework that we have used. These three components are labeled deterministic, caution and probing. In our experience the deterministic term is always convex and the cautionary term is almost always convex so these two terms do not generally cause a problem. However, the third component, namely the probing cost term can have a concave shape when the elements of $\hat{\Sigma}_{t|t}^{\theta\theta}$ are large. In this situation the sum of the deterministic, caution and probing terms may be a non-convex function with local optima and grid search methods or some global search optimization must be used.

However, in the cases where the $\hat{\Sigma}_{t|t}^{\theta\theta}$ declines rapidly over time the probing term diminishes in size relative to the other two terms and change to a less pronounced concave shape. Thus the cost-to-go function will be non-convex in the first few periods but will become convex in later time periods. In this situation one would like to use a global optimization procedure that checks in each time period to see whether or not nonconvexities are present. In those time periods where nonconvexities are present, time-consuming and thorough grid search methods must be use.

In contrast, in time periods where the cost-to-go function is apparently convex,

⁹ See Amman and Kendrick (1997, 1999c).

then efficient gradient methods can be used.¹⁰ We are still in the early stages of gaining experience with these approaches; however, it appears that they may be extremely useful in permitting the rapid solution of active learning control problems with substantial numbers of Monte Carlo runs in situations where there are nonconvexities in the cost-to-go function.

Sometimes mathematical results derived for one purpose have the fortunate side effect that they can later be used for another purpose. Such is the case of the analytical results that were originally derived to track down the sources of nonconvexities in small models. These results in Mizrach (1991) and Amman and Kendrick (1994b, 1995) allow one to fully characterize the three components of the cost-to-go function for the simplest one-state, one-control, one unknown parameter, quadratic linear adaptive control problem with a time horizon of two periods. Therefore, in Tucci et al. (2010) we have used these results as a starting point to compare the *average* or *representative* cost-to-go with different parameter sets and thus to analyze the effects of these different parameter sets on individual runs of a Monte Carlo experiment.¹¹ The representative cost-to-go helps to sort out the basic characteristics of the different parameter sets. The Monte-Carlo results are useful in reconciling the theoretical results in Mizrach (1991) and Amman and Kendrick (1994a, 1995) with the computational findings in Tucci (1998, 2004) and to shed some light on the outlier problem' discussed in Amman et al. (2008).

3.4 Rankings

One of the key questions in the field of learning in dynamic economic models is whether there is enough difference between the solutions provided by different methods to justify coping with the complexity of the more sophisticated methods. We have addressed this question primarily by comparing three methods: Optimal Feedback (*OF*), Expected Optimal Feedback (*EOF*) and Dual Control (*DC*). We have discussed these methods in detail in Amman, Kendrick and Tucci (2008); however, the basic difference can be outlined in a few sentences.

The *OF* method is the most simple since it ignores the uncertainty in the parameter matrix and considers only the additive noise uncertainty v_t in the system equations. Thus the solution method is very fast. The *EOF* method

¹⁰ Marco Tucci has devised just such an optimization procedure, viz Tucci (1998, 2004).

¹¹ The term *parameter set* is used here to include both the parameters and their covariance and the values used for the penalty matrices, desired paths for the states and controls and the initial states.

considers the uncertainty in both the additive noise term in the system equations and in the parameters in the θ_t vector which are elements from the A , B and c matrices from the system equations. The DC is more complex than the OF method, but is still computationally very efficient. However, this method does not consider the potential effect of change in the control today on the future uncertainty of the parameter estimates as represented in the $\Sigma_{t|t}^{\theta\theta}$ matrices – matters that are considered in the DC method. The DC method is substantially more complex mathematically than the first two and is also much less efficient computationally – especially so in cases where non-convexities arise.

When we first started learning about learning we expected that there would be a clear ranking between these three methods as measured by the criterion function values from the three solution methods over substantial numbers of Monte Carlo runs. We thought that DC would be clearly better than EOF which in turn would be clearly better than OF . Our experience so far has turned out to be more complex.

In each case when we have done the comparisons of the methods we have done it in two ways. One way is to count the percentage on the Monte Carlo runs in which each method had the lowest average criterion value across Monte Carlo runs when compared to the other two methods. The other way is to compute the average criterion value for each of the three methods over the Monte Carlo runs. A recent example of our work in this realm can be found in Amman, Kendrick and Tucci (2008) where we compare the three methods over solutions to the models from Beck and Wieland (2002). These results are typical of those that we have obtained with other models as well. Basically we find that the simple OF method does well relative to the more complex EOF and DC methods when the “number of runs” comparison is used. In contrast, the DC method does better than the EOF method which does better than OF method when one compares *average values*.

We believe that the difference between these two types of results can be explained in rather simple terms. It appears that the less sophisticated OF method may work better in most cases; however in situations where it does not work well it does very poorly. For example, in a Monte Carlo run where the initial values of the parameter estimates from the random drawing are relative close to the true values, then the OF methods does well relative to the EOF and DC methods. However, when the initial values of the parameter estimates are further from the true values then there is a premium on treating those values with caution, as is done in the EOF method, or of actively perturbing the system in order to increase the rate of learning and thereby obtaining parameter estimates closer to the true values in later periods as is done in the DC method.

Since in our judgment the average value method is the most useful approach

in most situations where learning occurs we believe that the *DC* method is better than the *EOF* and *OF* method. However, this leaves open the question of how much better. If we compute the variance of the criterion values we usually find that the difference between the *DC* method and the *EOF* method is statistically significant; however one may want to address this question in a larger realm than purely one of statistical significance.

Also, our experience is that the ranking between these three methods can be very model dependent.¹² With some models one may get the expected ranking of *DC* better than *EOF* better than *OF*; however, with other models this ranking may be different. We do not yet have enough experience with different models to say much about which properties of models affect the rankings; however we will address this question more in the following section of this chapter.

3.5 *Time-Varying Parameters*

There is one other more subtle question that we have begun to consider in this area. This comes from comparison of the methods across two versions of the same model, one without and one with time-varying parameters. (Here we mean time varying values of the true parameters and not of the parameter estimates since in all learning situations there are time varying values of the parameter estimates.) The Beck and Wieland (2002) model offers a good laboratory for this kind of experiment since it has both constant parameter and time-varying parameter versions.¹³ Using the “average value” approach it would seem that the difference between the *DC* method and the *EOF* method might diminish when one moves from the constant parameter to the time-varying parameter version of the model. This occurs because the gain from learning diminishes when the true values of the parameters themselves change over time. Our results in Amman, Kendrick and Tucci (2008) are of this variety.

However, it can also be argued that in situations where the true values of the parameters are time varying it will be even more important to perturb the system to try to continue to track those changing true values. Thus the jury may be out on this more subtle question for some time to come.

¹² See the discussion in Section 8 of Tucci et al. (2010).

¹³ Tucci et al. (2007) derive the appropriate formulae for the determination of *EOF* control in the presence of parameters following a *Return to Normality model*.

3.6 Model Sizes

Our experience with learning about learning is so far limited to small models with only a few equations and unknown parameters. However, as the speed of computers continues to increase it may well be possible to do large numbers of Monte Carlo runs using even the more sophisticated *DC* methods and even here in cases where there may be some nonconvexities in the cost-to-go function.¹⁴ The crucial issue will probably not be so much the number of uncertain parameters but rather the number of control variables since this governs the number of dimensions in which one must search in cases where there are nonconvexities.

4 What We Hope to Learn

The field of learning in dynamic economic models got off to a strong start twenty or thirty years ago but then went through a period where there was less interest among economists in the subject. That has changed in recent years with an increase in contributions from two groups, one at the Goethe University Frankfurt led by Volker Wieland and one at Notre Dame University led by Thomas Cosimano. These two groups and our own each use different methods for solving learning models so we have launched a small and informal project to compare the three methods and thus to learn their comparative advantage.

The first session of the project included presentation of the three methods. Wieland presented a dynamic programming method with numerical approximation of optimal decision rules, Wieland (2000a, 2000b), that is based on previous work on authors such as Prescott (1972), Taylor (1974) and Kiefer (1989). Cosimano discussed his perturbation method applied in the neighborhood of the augmented linear regulator problem (see Cosimano (2008) and Cosimano and Gapen (2005b, 2005a) which draws on work by Hansen and Sargent (2001, 2004, 2008). We talked about the adaptive control method (see Kendrick (1981, 2002) and Amman and Kendrick (1999a, 1999d, 2003) that draws on earlier work in the engineering literature by Tse (1973), Tse and Athans (1972), Tse and Bar-Shalom (1973) and Tse et al. (1973).

After the session we decided to use all three methods to solve the same model and thus to begin a new phase of learning about learning in dynamic economic models. The first model we selected to solve was the Beck and Wieland (2002)

¹⁴A recent paper on computational speed in this arena is Mathur and Morozov (2009).

model with both constant parameter and time-varying parameter versions. In this work Gunter Beck has joined with Volker Wieland, Michael Gapen is working with Thomas Cosimano and Marco Tucci has become a member of our group.

The focus of this Methods Comparison project will be to study the comparative advantage of the three methods in a number of dimensions including accuracy and efficiency and to use the three methods to further explore a number of questions about learning. These questions are what characteristics of models cause different rankings among *OF*, *EOF* and *DC* methods, whether or not non-convexities will turn out to be a common or uncommon problem in economic models, whether measurement errors are large enough that they should be routinely included in learning models and whether the use of constant or time-varying parameters in the models changes the answers to any of the questions above.

4.1 Accuracy

It would appear that the dynamic programming method used by Wieland and Beck will produce more accurate solutions than either of the other two approaches since they both use approximation methods. However, it remains to be seen how large this loss of accuracy is for each method relative to the dynamic programming methods.

4.2 Efficiency

It would also appear that the approximation methods used by (1) Cosimano and Gapen and (2) by our group will be more efficient and thus be able to solve larger models and do larger numbers of Monte Carlo runs on those models. However, it remains to be seen whether or not that difference is substantial and what dimensions of the models are most important in determining this relative efficiency. It may be that it is not the number of state variables or of control variables that is the crucial measure of size here but rather the number of parameters that are treated as unknown. For example, a relatively large model might have only a few parameters that need to be learned and thus the least efficient of the three methods will still be very useful for that model.

4.3 Rankings

In the process of applying the three methods to a number of different economic models we hope to begin to sort out the question of what characteristics of different economic models results in differences in the rankings between the *OF*, *EOF*, *DC* and related methods. From the discussion above it appears likely that the size of the initial parameter uncertainty in $\hat{\Sigma}_{0|0}^{\theta\theta}$ will be a crucial element; however, it also seems likely that the nature of the B matrix and the degree of controllability of the model may also be very important. Likewise it may be that the nature and size of the measurement errors may play a large role in determining the rankings of different methods on various types of models.

4.4 Non-convexities

As discussed above, it appears that the size of the elements in the initial estimate of the parameter covariance matrix $\hat{\Sigma}_{0|0}^{\theta\theta}$ play a crucial role in determining whether or not there are non-convexities in the cost-to-go function. If the elements in this matrix are relatively small in macroeconomic, or finance or microeconomic models then there will probably be little or no problem with local optima in those models. On the other hand if these elements are relatively large in models from some applications areas then it will be necessary to redouble our efforts in include efficient global search methods in our computer codes.

For models with a single control, grid search methods may prove to be an effective way to obtain the global optimal solution, even though these methods may not be the most efficient methods. However in models with more than one control - and even in some models with a single control - it may be wise to use more sophisticated global search methods. Our experience suggests that nonconvexities are much more common and subtle than the theoretical results suggests. For example Amman and Kendrick (1995, page 465) found that when the MacRae (1972) parameter set is used the cost-to-go function becomes non-convex when the variance of the estimated parameter is set to 2. However Tucci (1998, 2004) found nonconvexities in 28 percent of the cases of a Monte Carlo experiment with the same parameter set and a variance equal to 0.5 for the unknown parameter. Thus it may be wise to employ sophisticated global optimization methods with most models until one has gained some confidence that nonconvexities are not present.

4.5 *Measurement Errors*

Measurement errors have not been commonly considered in economic models, yet it seems likely that they may serve as an important limitation of learning in economic models. The good news here is that stochastic control methods commonly are equipped to handle measurement errors and we have included them in the models we have experimented with for some years. Also, there is recently new attention to this area of economics as in the work of Coenen et al. (2005).

4.6 *Time varying parameters*

The methods used by all three groups include, to various degrees, the ability to handle time-varying parameter specifications. In particular, Tucci (1989, 2004) has done considerable work in this area and has laid the ground work for exploration of the effects of time-varying parameters on the rankings of *OF*, *EOF* and *DC* methods and of the relative efficiency of different approaches to solving economic models with learning.

4.7 *Monte Carlo Runs*

While much useful knowledge can be obtained from analysis of *representative* runs of stochastic control models, when comparison of solution methods are being considered, Monte Carlo experiments have been the method of choice. For each Monte Carlo run random values of the system noise v_t and the initial parameter estimate $\hat{\theta}_{0|0}$ are generated using the means and covariances described above. This corresponds to running the model repeatedly with different additive noise terms and with different initial values of the estimate of the parameter θ_0 . Figure 4.1 shows (for a model with a single uncertain parameter) the probability density function of the initial parameter estimate for selected values of its covariance. It highlights the effect of this quantity on the actual values of the uncertain parameter used in a Monte Carlo experiment.

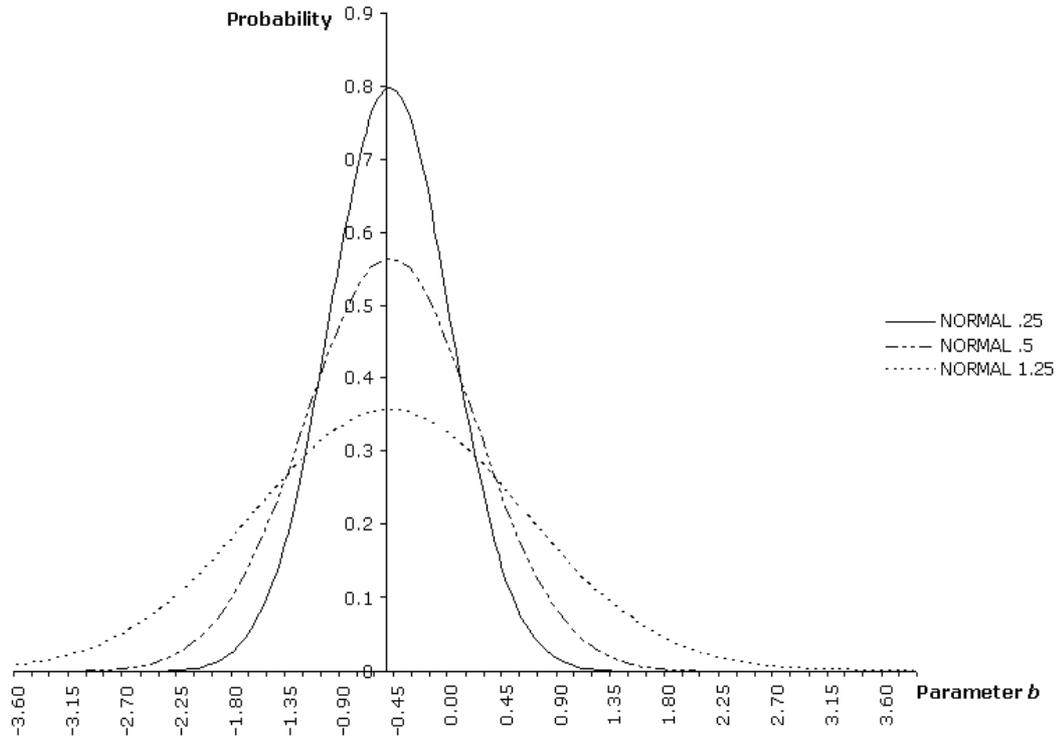
¹⁵

Therefore it is important to analyze the shape of the approximate cost-to-go function when the initial parameter estimate takes on values in different intervals. Recall that even though the probability that the parameter will take on a specific value is zero, the probability that it will fall in a certain interval is not zero. In studying the effect of changes in the parameter (when

¹⁵ For further discussion of these issues see Tucci et al. (2010).

Figure 4.1.

Probability density function of the initial parameter estimate $\hat{\theta}_{0|0}$
 when $\hat{\Sigma}_{0|0}^{\theta\theta} = 0.25$, $\hat{\Sigma}_{0|0}^{\theta\theta} = 0.50$ and $\hat{\Sigma}_{0|0}^{\theta\theta} = 1.25$.



that parameter is the one multiplied by the control variable) Amman and Kendrick (1995, page 470) observed that for the MacRae (1972) model the cost-to-go function will be convex for values either substantially above or below zero but will be non-convex for values close to zero. However, they did not investigate the relationship between this result and the outcomes of the Monte Carlo experiments. In Tucci et al. (2010) we have moved in this direction studying more closely this model. However more general models should also be considered

5 Algorithms and Codes

One of the most important aspects of research in this area is that the mathematics of the algorithms for learning in dynamic economic models is *complex*. It is not difficult but it is complex - both in the sense of there begin a lot of it and in the sense of requiring close attention to a substantial amount of detail.

The mathematics has cost-to-go functions, Riccati equations, Kalman filters and other components that are themselves not difficult; however, when they are combined in an active learning framework the whole requires prolonged and careful work to master.

Also, this is reflected in the computer codes. They too are not inherently difficult to understand but they require the development of many substantive subroutines that must be integrated carefully.

The result of this situation is that until recently there have not been many alternative algorithms and codes developed in this area. This meant that it has been difficult to cross check both mathematical and numerical results and therefore to increase the probability that both the math and the codes are error free. However, this is now changing. Within our own group Tucci (2002) has developed a variant of the adaptive control code which includes a variety of ways to specify time-varying parameters. In the process he has modified the mathematics in Kendrick (1981, 2002) and thus provided an independent derivation of the many mathematical results. Also, Beck and Wieland (2002) and Cosimano (2008) and Cosimano and Gapen (2005b, 2005a) have provided detailed results and derivations of those results that permit cross checking of aspects of the mathematics. A small beginning of cross checking of mathematical results is in Kendrick (2006) which compares results obtained by Beck and Wieland (2002) with those obtained by our group.

Also, we now have a variety of computer codes available that permit cross checking of numerical results. The original Dual code by Kendrick has been extensively revised and updated by Hans Amman to create DualPC which is an efficient code in Fortran. Kendrick has developed a Windows version of Dual in the C language with an easy to use interface (DualI) to permit low cost entry to the field. Amman has developed code in MATLAB that is useful for solving models with rational expectations. Tucci has created a version of Dual with a variety of time-varying parameter specifications (DualTVP). We have found it most useful to run these codes against one another in a continuing process of checking numerical results.

Also Wieland has developed code in Fortran to implement his algorithm and has made it available at his web site (<http://www.volkerwieland.com>) and Cosimano and Gapen have developed code in MATLAB to implement their algorithm. We anticipate in the future that all three groups will be able to solve a number of different macroeconomic and financial models and thus facilitate numerical checking across codes. As an example of this, in the following section we report on work we have done recently on the Beck and Wieland (2002) model.

Recently, a research group headed by Reinhard Neck at Klagenfurt University in Austria has extended their OPTCON code for models with nonlinear system equation to include a Kalman filter for learning parameters.¹⁶ They have informally reported that they have been able to replicate results from two test models in the Dual software for both OF and EOF solutions.

6 A Showcase on Active Learning

In this section we will present a simple model that fits into the framework we have presented in Section 2 and we focus on the issue of *rankings* as discussed in Section 3.4. Thus we compare optimal feedback (*OF*), expected optimal feedback (*EOF*) and dual control (*DC*) methods. We begin with a brief presentation of the BW model followed by the results for the constant parameter version of the model and then progress to the results for the version with *time-varying* parameters.

6.1 Outline of the Beck and Wieland Model

Following Beck and Wieland (2002) the decision maker is faced with a linear stochastic optimization problem of the form

$$J = \underset{[u_t]_{t=0}^{T-1}}{\text{Min}} E \left[\delta^T (x_T - \hat{x}_T)^2 + \sum_{t=0}^{T-1} \delta^t \{ (x_t - \hat{x}_t)^2 + \lambda (u_t - \hat{u}_t)^2 \} \right] \quad (6.1)$$

subject to the equations

$$x_{t+1} = \alpha x_t + \theta_t u_t + c + v_t \quad (6.2)$$

$$\theta_{t+1} = \theta_t + \eta_t \quad (6.3)$$

In fact the model goes back to an earlier strand of literature in the early Sev-

¹⁶ See Neck (1984), Matulka and Neck (1992) and Blueschke-Nikolaeva, Blueschke and Neck (2010).

encies, see MacRae (1975). The model contains one uncertain parameter θ_t , with an initial estimate of its value $\hat{\theta}_{0|0} = b_0$, and an initial estimate of its variance $\hat{\Sigma}_{0|0}^{\theta\theta} = \hat{\sigma}_{\theta_0} = v_0^b$. The parameters α and c are constant, $v_t \sim N(0, \sigma_v)$ and $\eta_t \sim N(0, \sigma_\eta)$. Beck and Wieland assume in their paper that $T \rightarrow \infty$. In contrast, we will assume that the planning horizon is finite, hence $T < \infty$. Furthermore, we have adopted the *timing* convention from Kendrick (1981, 2002) where the control, u_t , has a lagged response on the state x_t .

For the simulations in the next paragraphs we will use the following numerical values for experiments: $\alpha = 1$, $b_0 = -0.50$, $v_0^b = 1.25$, $c = 0$, $\sigma_v = 1$, $\sigma_\eta = 0$ for the constant parameter case and $\sigma_\eta = 0.04$ for the time varying parameter case (see the next two sections). Furthermore, $\lambda = 10^{-6}$, $T = 10$, $\forall t \hat{x}_t = 0$, $\forall t \hat{u}_t = 0$, $\delta = 1$, $x_0 = 1$.

With this set of parameters, the above model can be solved in *DualPC*, see Amman et al. (2008), allowing us to simulate the various situations in the next sections.

6.2 Constant Parameters

We make a distinction, at each time step, between the *true* parameters, β_t , and the *estimates* of those parameters, b_t . In this version of the model the true values of the parameters are constant but the estimates change over time. In contrast, in the time-varying parameter version of the model in Section 6.3, both the true parameters and the estimates change over time.

The parameters are the same as those used with the versions of the model solved with *DualI* except for the discount factor, δ , which is set at 0.95 in the *DualI* versions and at 1.00 in the *DualPC* versions. The reason for this is that the *DualPC* software does not yet support discounting.

We used the *DualPC* software to run 10,000 Monte Carlo in which we compared the criterion values obtained with three different methods *OF*, *EOF* and *DC*. As indicated above, the first two methods are described in the complementary paper, Kendrick and Tucci (2006). For each Monte Carlo run random values of the system noise ϵ_t and the initial parameter estimate \hat{b}_0 are generated using the means and covariances described above. This corresponds to running the model repeatedly with different additive noise terms and with different initial values of the estimate of the parameter b_t .

The *DC* (adaptive control) methods used here is the one described in Chapters 9 and 10 of Kendrick (2002). In addition, the *DualPC* software includes a grid search method that is designed to deal with possible non-convexities

in the cost-to-go function. This is a two level grid search that begins in the neighborhood of the OF solution. The best grid point obtained in the first search then provides the starting point for the second level search which is done in finer detail over a lesser range than the first grid search.

When we applied the OF , EOF and DC methods to the BW model we found that in a substantial number of the runs the criterion value for one or more of the methods was unusually large. Or, to say this in another way, the distribution of criterion values had a long right tail.

This outlier problem may be caused by the initial parameter estimates for the uncertain parameter, which are themselves outliers in either the right or left tails of that distribution. Assuming that the uncertain parameter has mean $-.5$ and variance 1.25 implies that the initial value used in the Monte Carlo runs is in the interval $(-1.62, .62)$ in approximately 68% of the cases, and in the intervals $(-2.74, 1.74)$ and $(-3.86, 2.86)$ in approximately 95% and 99%, respectively, of the cases. Alternatively put, the initial estimate of the unknown parameter is outside the *narrower* interval $(-1.62, .62)$ approximately 32% of times and this is obviously reflected in the associated value of the criterion function. To see how the various methods perform in the different situations we decided to run the comparison three times. In the first test we kept the runs in which the criterion value for all of the three methods was less than or equal to 100. In the second test we set this boundary at 200 and in the third run we set the boundary at 500. Thus in the three test we include a larger and larger number of the runs that we are uneasy about. Therefore we are inclined to give more credence to the test with the lower cut off values.

Our calculations were performed on a 64 bit AMD Dual Opteron Linux machine. Each simulation run of 10,000 Monte Carlo runs took about an hour of processing time on one processor. In comparing the results we looked first at the percentage of runs in which each method proved to have the lowest criterion value. These results are shown in Table 6.1.

Table 6.1

Percentage of Runs in Which Each Method
had the Lowest Criterion

J	OF	EOF	DC
$J < 100$	44.2	18.1	37.7
$J < 200$	43.0	18.4	38.6
$J < 500$	41.7	19.0	39.2

The first line of Table 6.1 shows that for the $J < 100$ case the percentage of Monte Carlo runs for which each method obtained the lowest criterion value was *OF* 44 percent, *EOF* 18 percent and *DC* 38 percent. These results proved to be relatively constant across the three rows of Table 6.1 which indicated that the outlier problem does not seem to affect the relative performance of the three methods.

The second way we compared the results was by examining the average criterion value for each method with their standard errors.¹⁷ These results are shown in Table 6.2.

Table 6.2

Average Criterion Value for Each Method

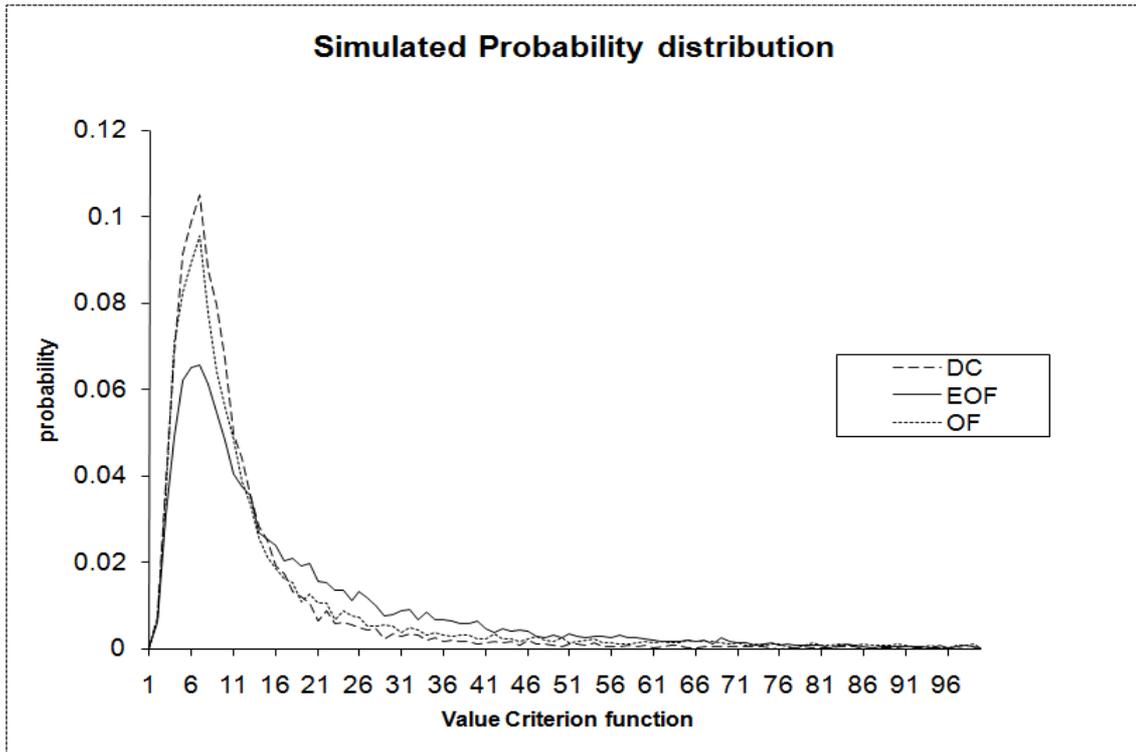
J	<i>OF</i>	<i>EOF</i>	<i>DC</i>
$J < 100$	14.247 (0.158)	17.527 (0.163)	10.979 (0.103)
$J < 200$	18.021 (0.267)	18.627 (0.188)	11.732 (0.133)
$J < 500$	25.970 (0.564)	18.803 (0.188)	12.142 (0.155)

The first row in Table 6.2 shows that the *OF* and *EOF* methods do not do as well as the *DC* method in the $J < 100$ case. This is the result to which we currently assign the most credence because it describes a situation where the estimated parameter is not "too far" from the true unknown value. The simulated probability density function for *OF*, *EOF* and *DC*, $J < 100$, is plotted in Figure 6.1.

Looking down the columns in Table 6.2 we see, not unexpectedly, that the average criterion values increase as more of the outliers are included. However, it is worth noting that the criterion values in the *EOF* column do not increase as rapidly as do those in the *OF* column. This is consistent with our results in Amman and Kendrick (1999e), that when the *OF* solutions are bad they may be really bad, i.e. when you have an outlier estimate for the

¹⁷ The standard errors are printed between parentheses and defined as $\frac{s}{\sqrt{n}}$ where s is the standard deviation in the sample and n the number of monte carlo runs. With the help of the standard errors it is possible to compute the confidence intervals for the various methods. For instance, the 95% percent confidence for the mean of the *DC* solution is $10.979 \pm 2 \times 0.103$, cf. Glasserman (2004, page 541).

Figure 6.1.



parameter value and treat it as though you trust that it is correct one can get a seriously bad solution. In these cases *EOF* is better, because it is cautious, and *DC* is better yet, because it devotes a part of the control energy to experiments to help learn the parameter value. Next we turn to the version of the BW model with time-varying parameters.

6.3 Time-Varying Parameters Version

In this version of the BW model the true value of the parameter is time varying and follows a first order Markov process. The major change from the first version of the model is that the variance of the additive noise term in the parameter evolution equation, σ_η is not zero, as in the previous version, but is 0.04. Also, recall that for both versions of the model solved with the *DualPC* software, the discount factor, δ , is not 0.95 but rather 1.00.

Just as with the constant parameter version of the model, in comparing the results we looked first at the percentage of the 10,000 Monte Carlo runs in which each method proved to have the lowest criterion value. These results are shown in Table 6.3.

Table 6.3

Percentage of Runs in Which Each Method had the Lowest Criterion

J	OF	EOF	DC
$J < 100$	30.9	28.0	41.1
$J < 200$	29.4	28.9	41.7
$J < 500$	27.9	30.6	41.5

The first line of Table 6.3 shows that for the $J < 100$ case the OF method and the EOF method each had the lowest criterion value in roughly 30 percent of the Monte Carlo runs and that the DC method had the lowest criterion value in roughly 40 percent of the runs. So the DC method proves to be the best of the three when the comparison is done in this way.

Then a comparison of the second and third rows in Table 6.3 to each other and to the first row shows that the percentage of the Monte Carlo runs in which each method had the lowest criterion value was not affected much by the number of the outliers that were included. Again the outliers do not seem to affect the relative performance of OF , EOF and DC .

The second way we compared the results was by examining the average criterion value. These results with their standard errors are shown in Table 6.4.¹⁸

Table 6.4

Average Criterion Value for Each Method

J	OF	EOF	DC
$J < 100$	18.612	17.289	12.529
	(0.197)	(0.167)	(0.126)
$J < 200$	26.511	19.362	15.433
	(0.357)	(0.207)	(0.212)
$J < 500$	43.733	20.254	20.493
	(0.788)	(0.219)	(0.423)

The first row in Table 6.4 shows that the OF method does not do as well as the EOF method which in turn does not do as well as the DC methods in the $J < 100$ case. This is the result to which we currently assign the most credence.

¹⁸The standard errors are between parentheses. See also footnote 17.

Looking down the columns in Table 6.4. shows, not unexpectedly, that the average criterion values increase as more of the outliers are included. However, it is worth noting that the criterion values in the *EOF* column do not increase as rapidly as do those in either of the other columns.

Overall the most important results from these Monte Carlo experiments to provide ranking among methods using the Beck and Wieland model is that the Dual Control (*DC*) method is better than the Expected Optimal Feedback (*EOF*) and the Optimal Feedback (*OF*) method in the constant parameter case. Also, in the time-varying parameter case the *DC* method is better than the *EOF* method which in turn is better than the *OF* method.

7 Learning with Forward Looking Variables

In the mid 1970s the use of optimal control techniques for deriving an optimal macroeconomic policy, e.g. Pindyck (1973) and Chow (1975), came under scrutiny. The critique by Lucas (1976) argued that it is difficult to determine optimal macroeconomic policies because the announcement of these policies results in changes in behavior by economic agents and thus changes in the parameters on which the optimal policy was based.

In this view one of the major drawbacks was that control methods could not deal with *forward looking variables* or rational expectations (RE). Subsequently a number of generic methods to solve models with forward looking variables were developed. For instance, Fair and Taylor (1983) used an iterative method for solving RE models and, in the tradition of Theil (1964), Fisher, Holly and Hughes Hallett (1986) used a method based on stacking the model variables. Blanchard and Kahn (1980) and Anderson and Moore (1985) both presented methods based on the saddle point property.

A computationally attractive and widely applicable method was introduced by Sims (2001) and applied in a control framework by Amman and Kendrick (2003). Sims solves forward looking model by using a *generalized inverse* approach based on the *QZ decomposition*. In the following subsection we will show how Sims' approach can be used with forward looking models that include learning.

7.1 Extending the framework

The state equation (2.1) can be augmented to accommodate for forward looking variables like this

$$x_{t+1} = A_t(\theta_t)x_t + B_t(\theta_t)u_t + C_t(\theta_t)z_t + \sum_{j=1}^k P_{j,t}(\theta_t)E_t x_{t+j} + v_t \quad (7.1)$$

where additionally to Section 2 the matrix $P_{j,t}(\theta_t) \in \mathbb{R}^{(n \times n)}$ is a parameter matrix, $C_t(\theta_t) \in \mathbb{R}^{(n \times s)}$ is a parameter matrix for the exogenous variables, $z_t \in \mathbb{R}^s$ is the vector of exogenous variables, $E_t x_{t+j} \in \mathbb{R}^n$ is the expected state for time $t+j$ at time t , k being the maximum lead in the expectations formation.

In order to compute the admissible set of instruments we have to eliminate the rational expectations from the model. In order to apply Sims' method we first put (7.1) in the form

$$\Gamma_{0,t}\pi_{t+1} = \Gamma_{1,t}\pi_t + \Gamma_{2,t}u_t + \Gamma_{3,t}z_t + \Gamma_{4,t}v_t \quad (7.2)$$

where

$$\Gamma_{0,t} = \begin{bmatrix} I - P_{1,t} & -P_{2,t} & \dots & -P_{k-1,t} & -P_{k,t} \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & & \ddots & 0 & 0 \\ 0 & \dots & & I & 0 \end{bmatrix}$$

$$\Gamma_{1,t} = \begin{bmatrix} A_t & 0 & 0 & \dots & 0 \\ 0 & I & 0 & \dots & 0 \\ 0 & 0 & I & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & I \end{bmatrix} \quad \Gamma_{2,t} = \begin{bmatrix} B_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Gamma_{3,t} = \begin{bmatrix} C_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Gamma_{4,t} = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and the augmented state vector

$$\pi_t = \begin{bmatrix} x_t \\ Ex_{t+1} \\ Ex_{t+2} \\ \vdots \\ Ex_{t+k-1} \end{bmatrix} \quad (7.3)$$

Taking the generalized eigenvalues of (7.2) allows us to decompose the system matrices $\Gamma_{0,t}$ and $\Gamma_{1,t}$ in the following manner, *viz.* Coleman and Van Loan (1988) or Moler and Stewart (1973),

$$\begin{aligned} \Lambda_t &= Q_t \Gamma_{0,t} Z_t \\ \Omega_t &= Q_t \Gamma_{1,t} Z_t \end{aligned}$$

with $Z_t' Z_t = I$ and $Q_t' Q_t = I$. The matrices Λ_t and Ω_t are upper triangular matrices and the generalized eigenvalues are $\forall i \omega_{i,i}/\lambda_{i,i}$. If we use the transformation $w_t = Z_t' \pi_t$ and $w_{t+1} = Z_t' \pi_{t+1}$ we can write (7.2)

$$\Lambda_t w_{t+1} = \Omega_t w_t + Q_t \Gamma_{2,t} u_t + Q_t \Gamma_{3,t} z_t + Q_t \Gamma_4 v_t \quad (7.4)$$

It is possible to reorder the matrices Z_t and Q_t in such a fashion that the diagonal elements of the matrices Λ_t and Ω_t contain the generalized eigenvalues in ascending order. In that case we can write (7.4) as follows

$$\begin{aligned} & \begin{bmatrix} \Lambda_{11,t} & \Lambda_{12,t} \\ 0 & \Lambda_{22,t} \end{bmatrix} \begin{bmatrix} w_{1,t+1} \\ w_{2,t+1} \end{bmatrix} = \\ & \begin{bmatrix} \Omega_{11,t} & \Omega_{12,t} \\ 0 & \Omega_{22,t} \end{bmatrix} \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} + \begin{bmatrix} Q_{1,t} \\ Q_{2,t} \end{bmatrix} \Gamma_{2,t} u_t + \begin{bmatrix} Q_{1,t} \\ Q_{2,t} \end{bmatrix} \Gamma_{3,t} z_t + \begin{bmatrix} Q_{1,t} \\ Q_{2,t} \end{bmatrix} \Gamma_4 v_t \quad (7.5) \end{aligned}$$

where the unstable eigenvalues are in lower right corner, that is the matrices $\Lambda_{22,t}$ and $\Omega_{22,t}$. By forward propagation and taking expectations, it is possible to derive $w_{2,t}$ as a function of future instruments and exogenous variables

$$\gamma_t = w_{2,t} = - \sum_{j=0}^{\infty} \tilde{M}_{t,j} \Omega_{22,t+j}^{-1} Q_{2,t+j} (\Gamma_{2,t+j} u_{t+j} + \Gamma_{3,t+j} z_{t+j}) \quad (7.6)$$

The matrix $\tilde{M}_{t,j}$ is defined as

$$\tilde{M}_{t,j} = \prod_{i=0}^{j-1} M_{t+i} \quad \text{for } j > 0$$

and

$$\tilde{M}_{t,j} = I \quad \text{for } j = 0$$

with

$$M_t = \Omega_{22,t}^{-1} \Lambda_{22,t}$$

Given the fact that $\Omega_{22,t}$ contains the eigenvalues outside the unit circle, we have applied the following condition in deriving (7.6)

$$\lim_{j \rightarrow \infty} \tilde{M}_{t,j} = 0$$

In contrast to Sims, M_t is not time invariant since we explicitly want to allow for time dependent matrices in the model. Reinserting (7.6) into (7.5) gives us

$$\tilde{\Lambda}_t w_{t+1} = \tilde{\Omega}_t w_t + \tilde{\Gamma}_{2,t} u_t + \tilde{\Gamma}_{3,t} z_t + \tilde{\Gamma}_{4,t} v_t + \tilde{\gamma}_t \quad (7.7)$$

with

$$\begin{aligned} \tilde{\Lambda} &= \begin{bmatrix} \Lambda_{11,t} & \Lambda_{12,t} \\ 0 & I \end{bmatrix} & \tilde{\Omega} &= \begin{bmatrix} \Omega_{11,t} & \Omega_{12,t} \\ 0 & 0 \end{bmatrix} & \tilde{\Gamma}_{2,t} &= \begin{bmatrix} Q_{1,t} \\ 0 \end{bmatrix} \Gamma_{2,t} \\ \tilde{\Gamma}_{3,t} &= \begin{bmatrix} Q_{1,t} \\ 0 \end{bmatrix} \Gamma_{3,t} & \tilde{\Gamma}_{4,t} &= \begin{bmatrix} Q_{1,t} \\ 0 \end{bmatrix} \Gamma_4 & \tilde{\gamma}_t &= \begin{bmatrix} 0 \\ \gamma_t \end{bmatrix} \end{aligned}$$

Knowing that $\pi_t = Z_t w_t$ and $\pi_{t+1} = Z_t w_{t+1}$ we can write (7.7) as

$$\pi_{t+1} = \tilde{A}_t(\theta_t)\pi_t + \tilde{B}_t(\theta_t)u_t + \tilde{C}_t(\theta_t)\tilde{z}_t + v_t \quad (7.8)$$

with

$$\tilde{A}_t = Z_t\tilde{\Lambda}_t^{-1}\tilde{\Omega}_tZ_t' \quad \tilde{B}_t = Z_t\tilde{\Lambda}_t^{-1}\tilde{\Gamma}_{2,t} \quad \tilde{C}_t = \begin{bmatrix} Z_t\tilde{\Lambda}_t^{-1}\tilde{\Gamma}_{3,t} & Z_t\tilde{\Lambda}_t^{-1} \end{bmatrix} \quad (7.9)$$

and

$$\tilde{\Lambda}_t^{-1} = \begin{bmatrix} \Lambda_{11,t}^{-1} & -\Lambda_{11,t}^{-1}\Lambda_{12,t} \\ 0 & I \end{bmatrix} \quad \tilde{z}_t = \begin{bmatrix} z_t \\ \tilde{\gamma}_t \end{bmatrix} \quad \tilde{v}_t = Z_t\tilde{\Lambda}_t^{-1}\tilde{\Gamma}_{4,t}v_t \quad (7.10)$$

Now that we have the augmented state equation (7.1) in the form of equation (7.8) we can compute the admissible controls as a function of the expected future states

$$u_t = u(x_t, E_t x_{t+1}, \dots, E_t x_T) \quad (7.11)$$

Once we have applied these controls, we can go back to the augmented system and compute the expected future state $\{E_t x_{t+1}, E_t x_{t+2} \dots\}$ and estimate the parameters $\hat{\theta}_{t|t}$ when new information on the state x_{t+1} comes available. Hence, like with the controls $\{u_t, u_{t+1} \dots\}$ the same holds for the estimation of the parameter that is a function of expected future states.

$$\hat{\theta}_{t|t} = \theta(x_t, E_t x_{t+1}, \dots, E_t x_T) \quad (7.12)$$

More detail about the corresponding algorithm is presented in Amman and Kendrick (1999a).

7.2 An example

In an earlier paper, Amman and Kendrick (2003), we introduced a model in which one can address the question of how best to do macroeconomic policy formulation in an environment in which agents may change their behavior over time, as well as in response to economic policy pronouncements. To mimic the Lucas critique in our example we will assume that the random vector θ_t is influenced by the real money supply in the following way

$$\theta_{t+1} = \theta_t + \left(\frac{\Delta M_t^R}{M_t^R} \right) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \theta_t + \eta_t \quad (7.13)$$

which means that an increase in the real money supply has an increasing effect on inflation and a decreasing effect on output.¹⁹ The policy maker will be unaware of this relationship and will try to estimate θ_t based on the random walk assumption.

Consider first the unrealistic case in which the policy makers knows exactly the component of the drift in the parameters which is due to policy changes. We call this scenario "Known Parameters". This is accomplished in the model by correcting the estimation of the parameters each time period when the updating occurs as in equations (2.12)-(2.13), that is

$$\hat{\theta}_{t|t+1} = \hat{\theta}_{t|t} + \hat{\Sigma}_{t+1|t}^{\theta x} (\hat{\Sigma}_{t+1|t}^{xx})^{-1} (x_{t+1} - \hat{x}_{t+1|t}) + \left(\frac{\Delta M_t^R}{M_t^R} \right) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \theta_t \quad (7.14)$$

so the policy maker is aware of the parameter shifts described in equation (7.13). Also, the parameters in this scenario are stochastic and time varying due to η_t . Note that equation (7.14) is slightly simpler than equation (2.12) due to the fact there is no measurement error and therefore $\forall t H_t = I$ and $\Sigma_t^{\zeta\zeta} = 0$. While, in reality, it would be impossible for policy makers to know exactly the parameter shifts like those modeled in equation (7.13) the "Known Parameter" case provides us with an infeasible but useful standard with which to compare other policy selection methods.

In the face of the fact that parameter drift may be occurring in the economy, policy may be determined in a deterministic manner treating all parameters as known and constant - i.e. a "Deterministic" scenario. In this case the true values of the parameters drift but the policy maker does not know what these values are and does not modify his or her policy in any way because of the stochastic elements of the model. One can then compare these two cases to

¹⁹ A referee has observed that we use a Lucas style supply curve (inverted as an equation determining inflation at time t , conditional on $t - 1$ expectation of time t inflation) in the Amman and Kendrick (2003) article. The referee suggested that the extension to the New-Keynesian Phillips curve (current inflation is then a function of current time t expectation of $t + 1$ inflation) would be interesting and could in principle be done with the methods we use (that is with the Sims (2001) method for solving RE models that we have integrated into the DC algorithm).

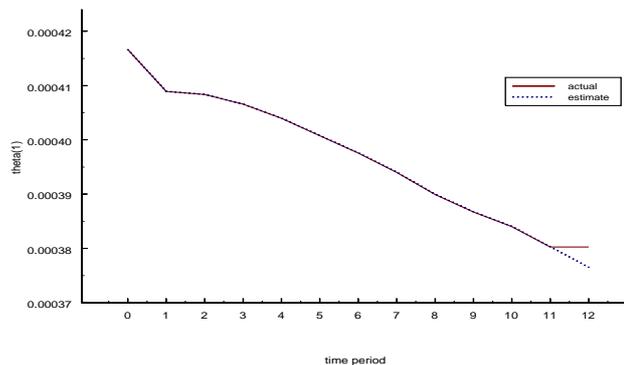
asks whether or not knowing the parameter drift makes a difference. In order to do this we performed 1000 Monte Carlo runs with the model above and obtained the results shown in Table 7.1.

Table 7.1
Comparison of the "Known Parameters" and "Deterministic" Scenarios

<i>Scenario</i>	$E\{J_T\}$	J_T^{min}	J_T^{max}	σ_{J_T}
Known Parameters	1.638	1.280	1.808	0.089
Deterministic	1.675	1.386	3.591	0.193

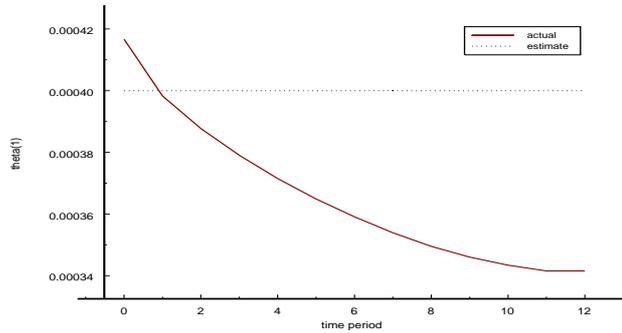
In both of these cases the true parameters are drifting over time and also changing in response to shifts in policy values; however, in the "Known Parameters" case the policy maker knows the effect of policy changes on the drift and in the "Deterministic" case he or she uses parameter values which are not correct, but treats those values as though they are (i) correct and (ii) that there is no uncertainty attached to the parameter estimates. Not surprisingly, the average criterion values is better at 1.638 for the known parameter case than at 1.675 for the deterministic case. Also from the max and min comparisons one can see that there is considerably more variability in the deterministic case than in the known parameter case and this is confirmed in the last column which shows the standard deviation of the criterion values across the 1000 Monte Carlo runs.

Figure 7.1. Known Parameters ($\theta(1)$)



As an example of the nature of these results, compare Figures 7.1 and 7.2 which show the values of parameter $\theta(1)$ in the two cases. The actual value of the parameter drifts down in both cases but changes by different amounts because the use of the policy variables is not the same in the two cases. In

Figure 7.2. Deterministic ($\theta(1)$)



contrast the estimated value tracks the actual very closely in the "Known Parameter" case but does not track and indeed does not change in the deterministic case.

So for the example at hand the scenario of "Known Parameters" is better than the scenario of ignoring the fact of the unknown parameters in the "Deterministic" case. Thus the Lucas critique is valid in this case. Ignoring the fact that the parameters are drifting and changing in response to policy pronouncements results in higher loss functions for the performance of the economy.

Would it be possible to do better with some other policy formulation method? One alternative approach would be to use game theory between the policy makers and the agents. Another, and the one which is examined here, is for the policy makers to (1) treat seriously the fact that they are using parameter estimates rather than the true parameter values when making policy and (2) update parameter estimates as the true values of the parameters drift over time and shift in response to policy announcements. We call this scenario the *Learning* approach. One can then asks whether the learning approach is indeed better than the Deterministic case. This comparison is provided in Table 7.2.

Table 7.2
Comparison of the "Deterministic" and "Learning" Scenarios

<i>Scenario</i>	$E\{J_T\}$	J_T^{min}	J_T^{max}	σ_{J_T}
Deterministic	1.675	1.386	3.591	0.193
Learning	1.646	1.292	1.817	0.090

The average value of the criterion function over the 1000 Monte Carlo runs is

better at 1.646 for the "Learning" case than at 1.675 for the "Deterministic" case. Also, the variability is lower for the learning case than the deterministic case as is shown in the remaining three columns of Table 7.2. Thus, for the case at hand, it is better to take account of the uncertainty in the parameters when determining policy in an environment where parameters are changing. Figures 7.3 and 7.4 show the true and estimated parameter estimates for the deterministic and learning scenarios - this time for parameter $\theta(2)$.

Figure 7.3. Deterministic ($\theta(2)$)

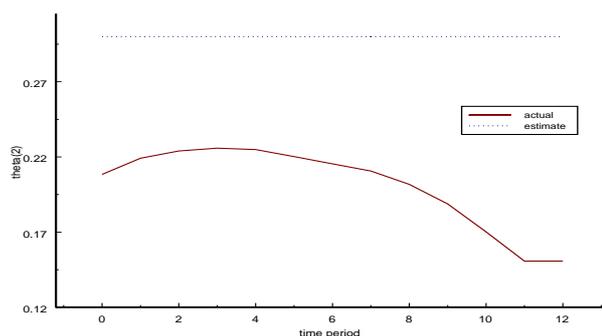
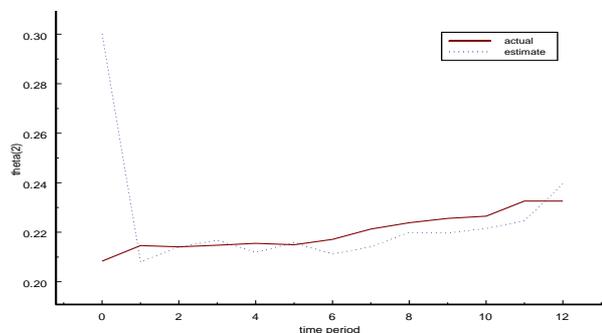


Figure 7.4. Learning ($\theta(2)$)



As before, in the deterministic case there is no change in the parameter estimate in Figure 7.3. In contrast Figure 7.4 shows that there is a very rapid learning of the $\theta(2)$ parameter in the learning scenario and that this parameter estimate then tracks closely the drift in the actual parameter over time.

Finally, one can ask whether the Lucas critique still carries heavy weight when policy makers treat seriously the uncertainty of parameter estimates in determining policy levels. Table 7.3 provides an indication of how this question

may be answered by providing a comparison of the "Known Parameters" and the "Learning" scenarios.

Table 7.3

Comparison of the "Known Parameters" and "Learning" Scenarios

<i>Scenario</i>	$E\{J_T\}$	J_T^{min}	J_T^{max}	σ_{J_T}
1: Learning	1.646	1.292	1.817	0.090
2: Known Parameters	1.638	1.280	1.808	0.089

There is little difference between these two cases in either the means in the first column or in the indicators of variance in the remaining three columns. Thus for this commonly used small macroeconomic model, the effects of the Lucas critique are substantially mitigated if one uses policy determination methods which treat the uncertainty in parameter estimates seriously and track the drifting parameters over time by using Kalman filter estimators. Moreover, it is not possible for policy makers to have as much knowledge about parameter drift as the "Known Parameter" method implies; therefore, the difference between the "Learning" method and a feasible method similar to the "Known Parameter" method would be even smaller.

A comparison of Figures 7.5 and 7.6 show that the "Known Parameters" scenario provides a better tracking of $\theta(1)$ than does the "Learning" scenario. However, a comparison of Figures 7.7 and 7.8 shows that the "Learning" method does almost as well at tracking $\theta(2)$ as does the "Known Parameters" method.

Figure 7.5. Known Parameters ($\theta(1)$)

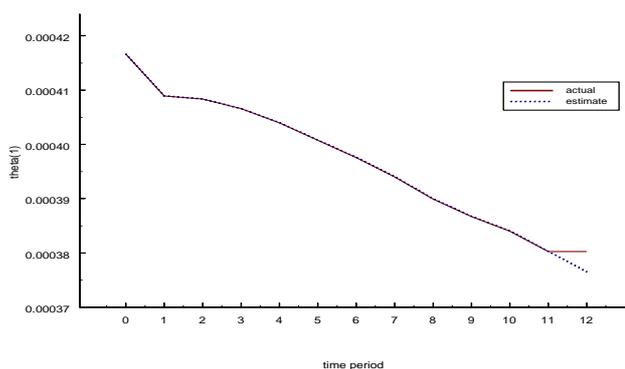


Figure 7.6. Learning ($\theta(1)$)

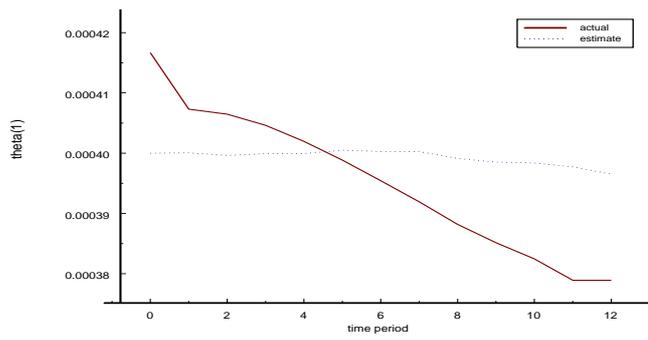


Figure 7.7. Known Parameters ($\theta(2)$)

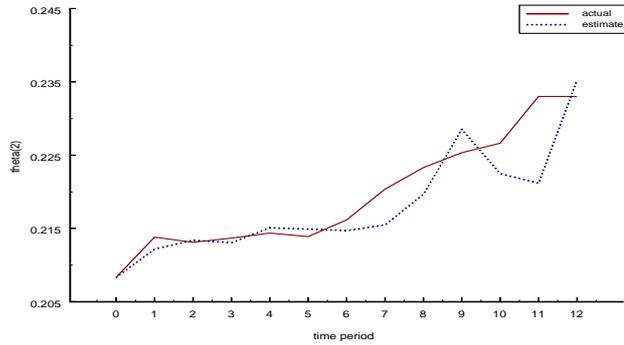
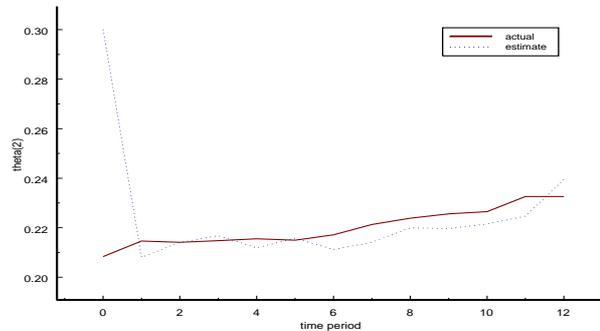


Figure 7.8. Learning ($\theta(2)$)



In summary, forward variables and the effects of policy choices on the parameters modeling agent behavior can be included in stochastic control settings with learning. Furthermore, these methods can be used to mitigate the effects of the Lucas critique.

8 Other Applications of Active Learning

A referee has suggested that we add a section to our original paper that would provide a list of areas of applications with references to a number of articles that study learning and control in both microeconomic and macroeconomic applications. Apparently many of these applications use bandit-type frameworks where unknown parameters are known to take on only one of two possible values. However, in the referees opinion some of these might be improved with the methods presented in this article. The areas and the articles are as follows:

- Monopolistic profit maximization with unknown demand: Kiefer (1989) or Trefler (1993).
- Experimental consumption, drug use: Crawford and Shum (2005) or Grossman, Kihlstrom and Mirman (1977).
- Investment and growth under uncertainty: El-Gamal and Sundaram (1993) or Bertocci and Spagat (1998)
- Inflation targeting and monetary policy: Wieland (2006), Ellison (2006) or Svensson and Williams (2007).

9 Summary

Thus looking back one can see that much has been learned about learning in dynamic economic models. It has been confirmed that in many cases the expected ranking between Optimal Feedback, Expected Optimal Feedback and Dual Control methods holds, but we have also learned that this is not always the case. We have been surprised to find that non-convexities can occur in the cost-to-go function but have been able to track down in the mathematics the reasons that this occurs and have been able to confirm numerically in the computer codes that these effects are significant. We have developed algorithms and computer codes for global search methods that are specifically tailored to deal with the way that non-convexities appear in economic models in early time periods but disappear in the later time periods of the same model.

We have also learned how to incorporate forward variables and the effects of policy choices on the parameters that model agent behavior into stochastic control models with learning and time-varying parameters. And, in addition, we have found that models in this class can be used to show how one can mitigate the effects of the Lucas critique.

However, we have not yet been able to learn the characteristics of various economic models which change the ranking among the methods. We do not yet know whether or not non-convexities will occur in most economic models or in only a few models. We do not yet know whether measurement errors will be large enough in many economic settings to substantially alter the results when they are included or excluded.

There are now a number of research groups actively working in this field and this has opened the possibility for fruitful checking of both mathematical and numerical results across algorithms and codes as well as the development and use of a variety of models from different fields of economics in order to learn more about learning.

References

- Aghion, P., Bolton, P., Harris, C. and Jullien, B.: 1991, Optimal learning by experimentation, *Review of Economic Studies* **58**, 621–654.
- 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.: 1994a, Active learning: Monte carlo results, *Journal of Economic Dynamics and Control* **18**, 119–124.
- Amman, H. M. and Kendrick, D. A.: 1994b, Nonconvexities in stochastic control problems: An analysis, in W. W. Cooper and A. B. Whinston (eds), *New Directions in Computational Economics*, Vol. 4 of *Advances in Computational Economics*, Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 57–94.
- 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.: 1997, Active learning: A correction, *Journal of Economic Dynamics and Control* **21**, 1613–1614.
- Amman, H. M. and Kendrick, D. A.: 1999a, Linear-quadratic optimization for models with rational expectations, *Macroeconomic Dynamics* **3**, 534–543.
- Amman, H. M. and Kendrick, D. A.: 1999b, 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.: 1999c, 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. and Kendrick, D. A.: 1999d, Should macroeconomic policy makers consider parameter covariances, *Computational Economics* **14**, 263–267.
- Amman, H. M. and Kendrick, D. A.: 1999e, Should macroeconomic policy makers consider parameter covariances, *Computational Economics* **14**, 263–267.
- Amman, H. M. and Kendrick, D. A.: 2003, Mitigation of the Lucas critique with stochastic control methods, *Journal of Economic Dynamics and Control* **27**, 2035–2057.
- 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.
- Anderson, G. and Moore, G.: 1985, A linear algebraic procedure for solving linear perfect foresight models, *Economic Letters* **17**, 247–252.
- Beck, G. and Wieland, V.: 2002, Learning and control in a changing economic environment, *Journal of Economic Dynamics and Control* **26**, 1359–1377.

- Bertocchi, G. and Spagat, M.: 1998, Growth under uncertainty with experimentation, *Journal of Economic Dynamics and Control* **23**, 209–231.
- Blanchard, O. J. and Kahn, C. M.: 1980, The solution of linear difference models under rational expectations, *Econometrica* **48**, 1305–1311.
- Blueschke-Nikolaeva, V., Blueschke, D. and Neck, R.: 2010, Optimal control of nonlinear dynamic econometric models: An algorithm and an application, *Technical report*, Klagenfurt University, Klagenfurt, Austria.
- Chow, G. C.: 1975, A solution to optimal control of linear systems with unknown parameters, *Review of Economics and Statistics* **57**(3), 338–345.
URL: <http://www.jstor.org/stable/1923918?origin=JSTOR-pdf>
- Coenen, G., Levin, A. and Wieland, V.: 2005, Data uncertainty and the role of money as an information variable for monetary policy, *European Economic Review* **49**, 975–1006.
- Coleman, T. F. and van Loan, C.: 1988, *Handbook for matrix computations*, SIAM, Philadelphia, Pennsylvania, 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 augmented 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.
- Crawford, G. and Shum, M.: 2005, Uncertainty and learning in pharmaceutical demand, *Econometrica* **73**, 1137–1173.
- Easley, D. and Kiefer, N. M.: 1988, Controlling a stochastic process with unknown parameters, *Econometrica* **56**, 1045–1064.
- El-Gamal, M. and Sundaram, R.: 1993, Bayesian economists ... bayesian agents, *Journal of Economic Dynamics and Control* **17**, 355–383.
- Ellison, M.: 2006, The learning cost of interest rate reversals, *Journal of Monetary Economics* **53**, 1895–1907.
- Fair, R. C. and Taylor, J. B.: 1983, Solution and maximum likelihood estimation of dynamic rational expectations models, *Econometrica* **51**, 1169–1185.
- Fisher, P. G., Holly, S. and Hughes-Hallett, A. J.: 1986, Efficient solution techniques for dynamic non-linear rational expectations models, *Journal of Economic Dynamics and Control* **10**, 139–145.
- Glasserman, P.: 2004, *Monte Carlo methods in financial engineering*, Springer, New York, New York.
- Grossman, S., Kihlstrom, R. and Mirman, L.: 1977, A bayesian approach to the production of information and learning by doing, *Review of Economic Studies* **44**, 533–547.
- Hansen, L. P. and Sargent, T. J.: 2001, Elements of robust control and fil-

- tering for macroeconomics. Draft downloaded from Sargents web site at www.stanford.edu/~sargent, version of 23 March 2001.
- Hansen, L. P. and Sargent, T. J.: 2004, Recursive models of dynamic linear economies. Department of Economics, University of Chicago manuscript.
- Hansen, L. P. and Sargent, T. J.: 2008, *Robustness*, Princeton University Press, Princeton, NJ.
- Kendrick, D. A.: 1976, Applications of control theory to macroeconomics, *Annals of Economic and Social Measurement* **5**, 171–190.
- Kendrick, D. A.: 1978, Non-convexities from probing in adaptive control problems, *Economic Letters* **1**, 347–351.
- Kendrick, D. A.: 1980, Control theory with application to economics, in K. J. Arrow and M. D. Intriligator (eds), *Handbook of Mathematical Economics*, Vol. 1 of *Handbook in Economics*, North-Holland Publishers (Elsevier), Amsterdam, the Netherlands, chapter 4, pp. 111–158.
- 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.: 1997, The DualI/DualPC software for optimal control models, in H. M. Amman, B. Rustem and A. B. Whinston (eds), *Computational Approaches to Economic Problems*, Vol. 6 of *Advances in Computational Economics*, Kluwer Academic Publishers, Dordrecht, the Netherlands, p. 363.
- 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.: 2005, Stochastic control for economic models: Past, present and paths ahead, *Journal of Economic Dynamics and Control* **29**, 3–30.
- Kendrick, D. A. and Amman, H. M.: 2006, A classification system for economic stochastic control models, *Computational Economics* **27**, 453–481.
- 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.
- Kendrick, D. A., Tucci, M. P. and Amman, H. M.: 2006, DualI: software for solving stochastic control problems, *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.
- Kiefer, N. and Nyarko, Y.: 1989, Optimal control of an unknown linear process with learning, *International Economic Review* **30**, 571–586.
- Lucas, R. E.: 1976, Econometric policy evaluation: A critique, in K. Brunner and A. Meltzer (eds), *The Phillips curve and the labor markets*, pp. 19–46. Supplementary series to the *Journal of Monetary Economics*.
- MacRae, E. C.: 1972, Linear decision with experimentation, *Annals of Economic and Social Measurement* **1**, 437–448.
- MacRae, E. C.: 1975, An adaptive learning role for multiperiod decision prob-

- lems, *Econometrica* **43**, 893–906.
- Mathur, S. and Morozov, S.: 2009, Massively parallel computation using graphics processors with application to optimal experimentation in dynamic control, *Technical report*, University of Munich Personal RePec Archive. URL: <http://mpira.ub.uni-muenchen.de/16721>.
- Matulka, J. and Neck, R.: 1992, An algorithm for the optimal control of nonlinear stochastic models, *Annals of Operations Research* **37**, 375–401.
- Mizrach, B.: 1991, Non-convexities in an stochastic control problem with learning, *Journal of Economic Dynamics and Control* **15**, 515–538.
- Moler, C. B. and Stewart, G. W.: 1973, An algorithm for generalized matrix eigenvalue problems, *SIAM Journal on Numerical Analysis* **10**, 241–256.
- Neck, R.: 1984, Stochastic control theory and operational research, *European Journal of Operations Research* **17**, 283–301.
- Norman, A. L.: 1976, First order dual control, *Annals of Economic and Social Measurement* **5**, 311–322.
- Norman, A. L., Norman, M. R. and Palash, C. J.: 1979, Multiple Relative Maxima in Optimal Macroeconomic Policy: An Illustration, *Southern Economic Journal* **46**, 274–729.
- Pindyck, R. S.: 1973, *Optimal planning for economic stabilization*, North-Holland, Amsterdam, the Netherlands.
- Prescott, E. C.: 1972, The multi-period control problem under uncertainty, *Econometrica* **40**, 1043–1058.
- Sims, C. A.: 2001, Solving linear rational expectations models, *Computational Economics* **20**, 1–20.
- Svensson, L. and Williams, N.: 2007, Bayesian and adaptive optimal policy under model uncertainty, *NBER Working Paper 13414*, NBER.
- Taylor, J. B.: 1974, Asymptotic properties of multiperiod control rules in the linear regression model, *International Economic Review* **15**, 472–482.
- Tesfaselassie, M. F., Schaling, E. and Eijffinger, S.: 2007, Learning about the term structure and optimal rules for inflation targeting, *Working paper*, Tilburg University, Tilburg, the Netherlands.
- Theil, H.: 1964, *Optimal Decision Rules for Government and Industry*, North-Holland, Amsterdam, the Netherlands.
- Trefler, D.: 1993, The ignorant monopolist: Optimal learning with endogenous information, *International Economic Review* **34**, 565–581.
- Tse, E.: 1973, Further comments on adaptive stochastic control for a class of linear systems, *IEEE Transactions on Automatic Control* **18**, 324–326.
- Tse, E. and Athans, M.: 1972, Adaptive stochastic control for a class of linear systems, *IEEE Transactions on Automatic Control* **17**, 38–52.
- Tse, E. and Bar-Shalom, Y.: 1973, An actively adaptive control for linear systems, *IEEE Transactions on Automatic Control* **18**, 109–117.
- Tse, E., Bar-Shalom, Y. and Meier, L.: 1973, Wide sense adaptive dual control for nonlinear stochastic systems, *IEEE Transactions on Automatic Control* **18**, 98–108.
- Tucci, M. P.: 1989, *Time-varying Parameters in Adaptive Control*, PhD thesis,

- Department of Economics, University of Texas, Austin, Texas 78712, USA.
- Tucci, M. P.: 1997, Adaptive control in the presence of time-varying parameters, *Journal of Economic Dynamics and Control* **22**, 39–47.
- Tucci, M. P.: 1998, The nonconvexities problem in adaptive control models: A simple computational solution, *Computational Economics* **12**, 203–222.
- Tucci, M. P.: 2002, A note on global optimization in adaptive control, econometrics and macroeconomics, *Journal of Economic Dynamics and Control* **26**, 1739–1764.
- 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, Expected optimal feedback with time-varying parameters, *Quaderni del Dipartimento di Economia Politica 497*, Università di Siena, Siena, Italy.
- 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.
- 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.
- Wieland, V.: 2006, Monetary policy and uncertainty about the natural unemployment rate: Brainard-style conservatism versus experimental activism, *Advances in Macroeconomics* **6**, 1–34.