NBER WORKING PAPER SERIES

LINEAR-QUADRATIC APPROXIMATION OF OPTIMAL POLICY PROBLEMS

Pierpaolo Benigno
Michael Woodford

Working Paper 12672
http://www.nber.org/papers/w12672

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2006

An earlier version of this paper was presented as a Plenary Lecture at the 10th Annual Conference on Computing in Economics and Finance, Amsterdam, July 2004. We would like to thank Filippo Altissimo, Vasco Curdia, Wouter Den Haan, Ken Judd, Jinill Kim, Andy Levin, Paul Levine, Diego Rodriguez Palenzuela, and Joseph Pearlman for comments, and the National Science Foundation for research support through a grant to the NBER. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Linear-Quadratic Approximation of Optimal Policy Problems
Pierpaolo Benigno and Michael Woodford
NBER Working Paper No. 12672
November 2006, Revised September 2007
JEL No. C61,C63

ABSTRACT

We consider a general class of nonlinear optimal policy problems involving forward-looking constraints (such as the Euler equations that are typically present as structural equations in DSGE models), and show that it is possible, under regularity conditions that are straightforward to check, to derive a problem with linear constraints and a quadratic objective that approximates the exact problem. The LQ approximate problem is computationally simple to solve, even in the case of moderately large state spaces and flexibly parameterized disturbance processes, and its solution represents a local linear approximation to the optimal policy for the exact model in the case that stochastic disturbances are small enough. We derive the second-order conditions that must be satisfied in order for the LQ problem to have a solution, and show that these are stronger, in general, than those required for LQ problems without forward-looking constraints. We also show how the same linear approximations to the model structural equations and quadratic approximation to the exact welfare measure can be used to correctly rank alternative simple policy rules, again in the case of small enough shocks.

Pierpaolo Benigno
Dipartimento di Scienze Economiche e Aziendali
Luiss Guido Carli
Via Tommasini, 1
00162 Rome - Italy
and NBER
pbenigno@luiss.it

Michael Woodford
Department of Economics
Columbia University
420 W. 118th Street
New York, NY 10027
and NBER
michael.woodford@columbia.edu

Linear-quadratic (LQ) optimal-control problems have been the subject of an extensive literature.[1] General characterizations of their solutions and useful numerical algorithms to compute them are now available, allowing models with fairly large state spaces, complicated dynamic linkages, and a range of alternative informational assumptions to be handled.[2] And the extension of the classic results of the engineering control literature to the case of forward-looking systems of the kind that naturally arise in economic policy problems when one allows for rational expectations on the part of the private sector has proven to be fairly straightforward.[3]

An important question, however, is whether optimal policy problems of economic interest should take this convenient form. It is easy enough to apply LQ methodology if one specifies an *ad hoc* quadratic loss function on the basis of informal consideration of the kinds of instability in the economy that one would like to reduce, and posits linear structural relations that capture certain features of economic time series without requiring these relations to have explicit choice-theoretic foundations, as in early applications to problems of monetary policy.[4] But it is highly unlikely that the analysis of optimal policy in a DSGE model will involve either an exactly quadratic utility function or exactly linear constraints.

We shall nonetheless argue that LQ problems can usefully be employed as approximations to exact optimal policy problems in a fairly broad range of cases. Since an LQ problem necessarily leads to an optimal decision rule that is linear, the most that one could hope to obtain with any generality would be for the solution to the LQ problem to represent a *local linear approximation* to the actual optimal policy — that is, a first-order Taylor approximation to the true, nonlinear optimal policy rule. In this paper we present conditions under which this will be the case, and show how to derive an LQ approximate problem corresponding to any member of a general class of optimal policy problems.

The conditions under which the solution to an LQ approximate problem will yield a correct local linear approximation to optimal policy are in fact more restrictive than might be expected, and than some of the literature on numerical methods for

[1]Important references include Bertsekas (1976), Chow (1975), Hansen and Sargent (2004), Kendrick (1981), Kwakernaak and Sivan (1972), and Sargent (1987). See Kendrick (2005) for an overview of the use of LQ methods in economics.

[2]For numerical algorithms see, among others, Amman and Kendrick (1999), Diaz-Gimenez (1999), Gerali and Lippi (2005), Hansen and Sargent (2004), and Söderlind (1999).

[3]See, *e.g.*, Backus and Driffill (1986) for a useful review.

[4]Notable examples include Kalchbrenner and Tinsley (1975) and Leroy and Waud (1977).

the analysis of DSGE models has suggested.[5] In particular, it does *not* suffice that the objective and constraints of the exact problem be continuously differentiable a sufficient number of times, that the solution to the LQ approximate problem imply a stationary evolution of the endogenous variables, and that the exogenous disturbances be small enough (though each of these conditions is obviously *necessary,* except in highly special cases). An approach that simply computes a second-order Taylor-series approximation to the utility function and a first-order Taylor-series approximation to the model structural relations in order to define an approximate LQ problem — what we shall call "naive LQ approximation" — may yield a linear policy rule with coefficients very different from those of a correct linear approximation to the optimal policy in the case of small enough disturbances, as the example of optimal dynamic tax policy considered in Benigno and Woodford (2006a) shows.[6]

Nonetheless, it is quite generally possible to construct an alternative quadratic objective function — one that also represents a correct local second-order approximation to expected utility under any feasible policy, but that does not imply the same linear characterization of optimal policy when used as the objective for an LQ problem — which *will* result in a correct local LQ approximation. The approach that we use is essentially the one introduced by Fleming (1971), and used by Magill (1977) to derive a local LQ approximation to a continuous-time multi-sector optimal growth model. Here we extend the work of Fleming and Magill by showing how a similar method can be used in the context of discrete-time dynamic optimization problems of the kind that typically arise in the literatures on optimal monetary and fiscal policy, and showing how the method can be extended to the case where some of the structural relations are forward-looking, as is almost inevitably the case in optimal policy problems.[7]

In section 1, we first explain the problem with naive LQ approximation in the context of a simple static optimization problem, and introduce the general idea of our alternative approach. We offer additional comparisons there of the approach that we propose to other possible approaches to the local characterization of optimal pol-

---

[5]This is stressed by Judd (1998, pp. 507-508), who recommends the use of alternative perturbation techniques for the local characterization of optimal policy.

[6]The same problem can also result in incorrect welfare rankings of alternative simple policies, as discussed by Kim and Kim (2003, 2006).

[7]See also Levine *et al.* (2007) for another discussion of how our method compares to that of Fleming and Magill.

icy. In section 2, we then show how the method can be applied to a general class of dynamic optimization problems with forward-looking constraints. Section 3 discusses the general algebraic form of the first- and second-order conditions for optimality in the LQ approximate problem. Section 4 shows how the quadratic objective for stabilization policy derived in section 2 can also be used to compute welfare comparisons between alternative sub-optimal policies, in the case that the stochastic disturbances are small enough. Finally, section 5 discusses applications of the general method described here and concludes.

1 Pitfalls of Naive LQ Approximation

Here we explain why naive LQ approximation is generally inadequate, in the context of a simple static optimization problem that allows us to the explain the issues in terms of simple multivariate calculus. We then compare a variety of possible responses to the problem, including the one that we favor.

1.1 Static Analysis

Suppose that we wish to find the policy $y(\xi)$ that maximizes an objective $U(y;\xi)$, where y is an n-vector of endogenous variables and ξ is a vector of exogenous disturbances; we assume that U is at least twice continuously differentiable with respect to the arguments y. Suppose furthermore that the possible outcomes y that can be achieved by policy in any state of the world ξ are those values consistent with the structural equations

$$F(y;\xi) = 0, \tag{1.1}$$

where F is a vector of m functions (for some $m < n$), again each at least twice continuously differentiable. We assume that $m < n$ so that there is at least one direction in which it is possible for the outcome y to be varied by policy. We might suppose that y is determined by equations (1.1) together with an additional set of $n - m$ equations of the form

$$G(y;i,\xi) = 0, \tag{1.2}$$

where i is a vector of $n-m$ instrument settings (or control variables); but the nature of the additional equations (1.2) does not matter for our conclusions below, as long as the derivative matrices

$$\begin{bmatrix} D_yF \\ D_yG \end{bmatrix}, \qquad D_iG$$

are of full rank when the partial derivatives are evaluated at the point around which we conduct our local analysis.

Now let $\bar{y}$ be the outcome under an optimal policy in the case that $\xi = 0$; that is, it maximizes $U(y; 0)$ subject to the constraints $F(y; 0) = 0$.[8] A second-order Taylor series expansion of U, computed at values $(\bar{y}; 0)$ of the arguments, is then given by

$$\begin{aligned} U(y;\xi) &= \bar{U} + D_yU \cdot \tilde{y} + D_\xi U \cdot \xi + \frac{1}{2}\tilde{y}' D^2_{yy}U \cdot \tilde{y} + \\ &\quad \frac{1}{2}\xi' D^2_{\xi\xi}U \cdot \xi + \tilde{y}' D^2_{y\xi}U \cdot \xi + \mathcal{O}(||\xi||^3) \\ &= D_yU \cdot \tilde{y} + \frac{1}{2}\tilde{y}' D^2_{yy}U \cdot \tilde{y} + \tilde{y}' D^2_{y\xi}U \cdot \xi + \text{t.i.p.} + \mathcal{O}(||\xi||^3), \end{aligned} \tag{1.3}$$

where we introduce the notation $\tilde{y} \equiv y - \bar{y}$, $\bar{U} \equiv U(\bar{y}; 0)$, and the several matrices of partial derivatives are each evaluated at $(\bar{y}; 0)$. The expression "t.i.p." refers to terms that are independent of the policy chosen (such as the constant term and terms that depend only on the exogenous disturbances); the form of these terms is irrelevant in obtaining a correct ranking of alternative policies. Finally, $||\xi||$ is a bound on the vector of disturbances ξ. In stating that the residual is of order $\mathcal{O}(||\xi||^3)$ in the amplitude of the disturbances, we assume that $y - \bar{y} = \mathcal{O}(||\xi||)$. This condition will hold, in the case of any policy that makes $y(\xi)$ continuously differentiable,[9] as long as $y(0) = \bar{y}$. We shall restrict our analysis to policies that satisfy the latter property, *i.e.*, that bring about $\bar{y}$ in the case that there are no disturbances.[10]

A naive LQ approximation of this problem can then be obtained by replacing the exact objective $U(y; \xi)$ by the quadratic objective

$$U^Q(y;\xi) \equiv D_yU \cdot \tilde{y} + \frac{1}{2}\tilde{y}' D^2_{yy}U \cdot \tilde{y} + \tilde{y}' D^2_{y\xi}U \cdot \xi, \tag{1.4}$$

[8]Note that we must compute our local approximations to the objective and constraints around this optimal point if there is to be any hope that consideration of these local approximations alone can correctly identify the optimal policy rule even in the case that ξ is small.

[9]In the case that $y(\xi)$ is determined by a vector of instrument settings through structural relations of the form (1.2), $y(\xi)$ will be continuously differentiable near $\xi = 0$ as long as the rank conditions stated in the previous paragraph are satisfied, and the policy rule $i(\xi)$ is itself continuously differentiable.

[10]Note that by assumption, the optimal policy rule belongs to this class of rules.

and replacing the exact constraints (1.1) by their linearized form,

$$D_y F \cdot \tilde{y} + D_\xi F \cdot \xi = 0. \tag{1.5}$$

The question that we wish to consider is whether the solution to this problem — that is, the policy $y^{LQ}(\xi)$ that maximizes $U^Q(y;\xi)$ subject to the constraints (1.5) — represents at least a correct local linear approximation to the true optimal policy $y^{opt}(\xi)$. That is, we wish to determine whether

$$y^{opt}(\xi) = y^{LQ}(\xi) + \mathcal{O}(||\xi||^2) \tag{1.6}$$

in the case of small enough disturbances.

In fact, the regularity conditions stated thus far do not suffice to guarantee this. The policy that maximizes the naive quadratic objective (1.4) subject to the linearized constraints (1.5) satisfies linear first-order conditions

$$D_y U + \tilde{y}' D^2_{yy} U + \xi' D^2_{\xi y} U + \lambda' D_y F = 0, \tag{1.7}$$

where λ (a function of ξ) is the vector of Lagrange multipliers associated with the constraints. The naive LQ-optimal policy $y^{LQ}(\xi)$ is then obtained by solving the system of equations consisting of (1.5) and (1.7) for y and λ as linear functions of ξ.

The solution $y^{opt}(\xi)$ to the exact policy problem instead satisfies the nonlinear first-order conditions[11]

$$D_y U(y;\xi) + \lambda' D_y F(y;\xi) = 0 \tag{1.8}$$

along with (1.1). A correct local approximation to the solution to these equations can be obtained (using the implicit function theorem) by linearizing equations (1.1) and (1.8) around the unperturbed solution $y(0) = \bar{y}$. The linearization of equations (1.1) is given by (1.5), as above, but the linearization of the first-order conditions (1.8) is given by

$$D_y U + \tilde{y}' D^2_{yy} U + \xi' D^2_{\xi y} U + \lambda' D_y F + \bar{\lambda}_I [\tilde{y}' D^2_{yy} F^I + \xi' D^2_{\xi y} F^I] = 0, \tag{1.9}$$

where $\bar{\lambda} \equiv \lambda(0)$ is the vector of multipliers when there are no shocks. Here we use tensor notation as in Judd (1998, chap. 14), omitting the summation sign Σ_I; the

[11] Here we assume that the solution to the first-order conditions is indeed the optimum, though this need not be true if the constraint set is non-convex.

index I ranges over the m constraints. Hence the correct linear approximation to $y^{opt}(\xi)$ is obtained by solving the system of equations consisting of (1.5) and (1.9) for y and λ as linear functions of ξ. Because the two final terms on the left-hand side of (1.9) are missing in (1.7), the naive method will generally yield incorrect coefficients for the linear policy rule.

The problem is that a linear approximation of the structural equations (1.5) suffices to indicate the possible ways in which it is possible for y to vary in response to ξ, *to first order* in the amplitude of the disturbances ξ, but this is not generally a sufficiently accurate characterization of outcomes under a given policy to allow an approximate evaluation of the objective U that is accurate to *second* order. In general, second-order contributions to the solution for $y(\xi)$ under a given policy rule make second-order contributions to the level of U associated with that rule; and even when $||\xi||$ is arbitrarily small, these second-order contributions to U need not be negligible relative to the other second-order contributions that are taken account of when one evaluates U^Q using a local linear approximation to $y(\xi)$.[12]

In fact, in the case of any given outcome $y(\xi)$ associated with a (sufficiently differentiable) policy, a second-order Taylor series expansion of $U(y(\xi);\xi)$ can be written in the form

$$U(y(\xi);\xi) = U^Q(y^L(\xi);\xi) + D_j U[\xi' D^2_{\xi\xi} y^j \cdot \xi] + \text{t.i.p.} + \mathcal{O}(||\xi||^3), \tag{1.10}$$

where U^Q is again the naive quadratic objective defined in (1.4),

$$y^L(\xi) \equiv \bar{y} + D_\xi y \cdot \xi$$

is a local linear approximation to $y(\xi)$, and in the second term on the right-hand side, we again use tensor notation. Here we have simplified using the fact that the derivatives $D_\xi y$ must satisfy

$$D_y F \cdot D_\xi y + D_\xi F = 0,$$

in order for y^L to represent a solution to the linearized structural relations (1.5). Estimation of the level of welfare associated with the given policy using $U^Q(y^L)$ omits the second-order contributions from the second term on the right-hand side of (1.10). These are second-order contributions to U resulting from second-order terms

[12]See Woodford (2002; 2003, sec. 6.1) and Sutherland (2002) for further discussion.

in the Taylor expansion of $y(\xi)$, that exist to the extent that the gradient vector $D_y U$ has non-zero elements. When these additional terms are non-zero, alternative policies cannot be correctly ranked, even to second order in the amplitude of the disturbances, simply on the basis of a local linear characterization of equilibrium outcomes under those policies.

1.2 Responses to the Problem

Several approaches have been taken in the literature to computing a correct local linear approximation to optimal policy, that (at least under certain circumstances) avoid the problem just expounded with a naive LQ approximation. We briefly discuss some of these before presenting our own proposed solution.

(1) The naive LQ approach yields a correct local characterization of optimal policy in the case that the constraints (1.1) are *exactly linear*). If they are, the use of the linearized equations (1.5) involves no error, and the problem discussed above does not arise. In the case of our static example above, linear constraints imply that

$$D^2_{yy}F^I, \ D^2_{\xi y}F^I = 0$$

for each I, so that equations (1.9) are equivalent to (1.7). Thus the problem with naive LQ approximation is not that the objective functions in optimal policy problems are not exactly quadratic, but rather that the constraints are almost never exactly linear.

Even in the case of a policy problem with nonlinear constraints, it may be possible to obtain a problem with purely linear constraints through a suitable change of variables. This is the approach used in Kydland and Prescott (1982) to obtain a valid LQ approximation, and expounded more generally by Diaz-Gimenez (1999). The (nonlinear) production function is substituted into the utility function to express utility as a function of the paths of hours, capital, and investment spending; the only remaining constraint is the exactly linear relation between investment spending and the dynamics of the capital stock. After this transformation of their planning problem, a second-order Taylor series expansion of the derived objective function yields an LQ planning problem, the solution to which is a correct linear approximation to the solution to the original planning problem. However, the circumstances under which a transformation of this kind can be found are fairly special.[13]

[13]Kydland and Prescott's "time-to-build" approach to modelling capital adjustment costs is nec-

(2) The naive LQ approach also yields a correct local characterization of optimal policy in the case that one expands around a point $\bar{y}$ at which the gradient vector $D_y U(\bar{y}; 0) = 0$. In this case the second term on the right-hand side of (1.10) is equal to zero under any policy, and $U^Q(y^L)$ correctly ranks alternative policies, to second order. Similarly, since in this case the constraints (1.1) do not bind in the absence of shocks, $\bar{\lambda} = 0$, and again conditions (1.9) reduce to (1.7). This is why an LQ approximation can be used to characterize optimal policy in the model of Rotemberg and Woodford (1997).[14]

In some cases, an appropriate change of variables may result in this condition holding. In the case of Rotemberg and Woodford, the gradient vector would be non-zero if one were to expand in terms of consumption and hours, the "direct" arguments of the utility function. But they use the (nonlinear) production function to solve for hours of each variety as a function of sectoral output, and the market-clearing relation to solve for consumption of each differentiated good as a function of output, obtaining an expression for utility as a function of the quantities produced of the various goods; and the gradient with respect to each of these quantities is zero, in the case that they consider. But even with the change of variables, the method is applicable only if the flexible-price equilibrium allocation of resources is efficient, which need not be the case, owing for example to market power or tax distortions (Benigno and Woodford, 2005a). This last observation is itself an important practical limitation, and in more complex examples it may not be easy to find a suitable change of variables.

(3) A correct local linear approximation to optimal policy can often be obtained by deriving the exact first-order conditions for (Ramsey) optimal policy using exact specifications of the objective and constraints, and then log-linearizing the non-linear stochastic difference equations obtained in this way, as illustrated in the derivation of equations (1.9) above. This method has been used extensively in the recent literature on optimal monetary and fiscal policy by authors such as King and Wolman (1999), Khan *et al.* (2003) and Schmitt-Grohé and Uribe (2004b). The method will generally yield a correct result as long as the optimal equilibrium, in the case of small enough

essary in order for the constraint to be exactly linear in their case, and hence important for the validity of the numerical method that they use to characterize equilibrium dynamics, though they do not comment on this. Standard convex adjustment costs, for example, would result in a nonlinear constraint.

[14] The conditions for the validity of this approach are further discussed in Woodford (2002).

exogenous shocks, remains forever near a deterministic steady state, around which first-order conditions are log-linearized.[15] It is also straightforward to obtain higher-order local characterizations of optimal policy, through a higher-order perturbation expansion of the first-order conditions.

A disadvantage of this approach, however, is that while it allows a solution for optimal policy, it does not provide a convenient way of ranking sub-optimal policies. An LQ approximation, if valid (as in either of the two cases just described), also provides a simple way of evaluating arbitrary policies, as long as they are not *too far* from optimal: one obtains an approximate characterization of the outcome under the policy by solving the linearized model equations (constraints), and then evaluates the quadratic loss function under the resulting linear dynamics. (The method should correctly rank policies, in the case of small enough shocks, as long as they are consistent with the steady state around which the local approximations are computed – or more generally, as long as they are *close enough* to consistency with it.) This is important, insofar as in models of a complexity that would allow them to be used in quantitative policy analysis, the fully optimal (Ramsey) policy is almost certainly too complex to represent a practical policy proposal, and the welfare losses associated with a simpler policy may be quite small. The comparative evaluation of simple policy rules, within families of rules too restrictive to include the optimal policy, is accordingly a prime goal of quantitative analyses of stabilization policy.

Another disadvantage is that solution of a local linear approximation to the first-order conditions does not guarantee that the solution is even locally an (approximate) optimum, as second-order conditions for the optimal policy problem may fail, as discussed further below in section 3, and in the context of a specific example in Benigno and Woodford (2005a). In the case of a valid LQ approximation, this issue is automatically settled (*i.e.*, a *local* optimum is guaranteed) if the quadratic loss function is convex, which requires only that one check an algebraic property of the weighting matrix.[16]

---

[15] In general, the equilibrium resulting from the optimal Ramsey policy is time-invariant, even in the absence of stochastic disturbances, only if one adds certain constraints on initial outcomes to the standard, "unconstrained" Ramsey problem. These are discussed further in section 2.1 below. This issue must be confronted by any local approximation method that characterizes optimal policy using linear equations with constant coefficients.

[16] Of course, one could check the second-order conditions for an optimum as part of the perturbation analysis of the exact Ramsey problem; but this seems seldom to be done in the literature

(4) Alternatively, substitution of an approximate solution into the naive quadratic objective U^Q will yield a correct ranking of alternative policies in the case of small enough disturbances, if we evaluate (1.3) using a *second-order* approximation to the equilibrium evolution $y(\xi)$ under any given policy rule, rather than a mere linear (or log-linear approximation). A second-order approximation to $y(\xi)$ can be computed by applying perturbation techniques to the system of equations consisting of (1.1) and (1.2), where the latter equation(s) specify the policy that is to be evaluated. Methods for executing computations of this kind in the case of general classes of forward-looking equation systems are now widely available,[17] and have been used in many recent numerical analyses of optimal policy (*e.g.,* Schmitt-Grohé and Uribe, 2004c).

This approach, however, has the disadvantage that it does not make it easy to find even an approximate characterization of fully optimal policy, as one has to compute a second-order approximation to the equilibrium dynamics implied by each candidate policy rule individually. One can approximate the optimal rule within a particular parametric family, by searching over a grid of parameter values, at each element of which one evaluates welfare; in practice, in such studies attention is restricted to low-dimensional families of simple rules. An LQ approach, when valid, instead allows one to determine which form of rule is optimal. And while the fully optimal rule is not likely to be of interest as a practical policy proposal, as noted above, computing it is nonetheless valuable as a source of insight into which types of simple rules are most likely to be nearly optimal.

Hence there would remain important advantages of an LQ approach, were a valid approximation of this form possible outside the restrictive cases already mentioned. Here we show how a valid LQ approximation can be derived, for a much more general class of policy problems.

(5) In the approach that we recommend, a quadratic loss function is derived that differs (in general) from U^Q, but that nonetheless represents a valid second-order approximation to U, in the case of the outcomes associated with any possible policy.

on Ramsey policy, and would in any event involve computing essentially the same matrices as are required to derive our LQ approximation, as is discussed further below.

[17] See, *e.g.*, Jin and Judd (2002), Kim *et al.* (2003), and Schmitt-Grohé and Uribe (2004a). The DYNARE project at CEPREMAP has been especially important in making these techniques widely available to macroeconomic researchers.

That is, we seek a quadratic function $\hat{U}(y;\xi)$ with the property that

$$U(y;\xi) = \hat{U}(y;\xi) + \mathcal{O}(||\xi||^3) \tag{1.11}$$

in the case of any values of the arguments satisfying (1.1) and such that $y - \bar{y} = \mathcal{O}(||\xi||)$. The fact that we require (1.11) to hold *only* for values of y that can be achieved by some policy, rather than for all values of y near enough to $\bar{y}$, means that $\hat{U}$ need not coincide with U^Q, despite Taylor's theorem. Among the variety of possible quadratic approximations $\hat{U}$ with this property, we furthermore seek one that is *purely quadratic, i.e.,* with zero coefficients on the linear terms. Then $D_y\hat{U} = 0$, and (1.11) can be evaluated to second-order accuracy using only a first-order accurate approximation to $y(\xi)$ under the policy rule of interest. Hence the LQ problem of maximizing the quadratic objective $\hat{U}(y;\xi)$ subject to the linear constraints (1.5) represents a valid local approximation to the original policy problem, and the linear policy that solves this LQ problem represents a correct local linear approximation of the optimal policy y^{opt}.

The key to finding an approximate objective with these properties is to use a second-order Taylor series approximation to the constraints (1.1) to replace the linear terms in (1.3) with purely quadratic terms;[18] while the resulting function is not even locally equivalent to U^Q, it is equivalent in the case of all outcomes consistent with equations (1.1) that are near enough to $\bar{y}$. While this method (like the one just discussed) relies upon computing a second-order approximation to the model structural relations, the second-order approximation need be used only *once,* in determining the coefficients of the quadratic objective $\hat{U}$, rather than having to be used again each time one seeks to evaluate the welfare associated with yet another candidate policy.

We can illustrate the method in the case of the static problem considered above. A second-order approximation to the structural relations (1.1), of the same form as the approximation (1.3), implies that

$$D_yF^I \cdot \tilde{y} = -\frac{1}{2}\tilde{y}'D^2_{yy}F^I \cdot \tilde{y} - \tilde{y}'D^2_{y\xi}F^I \cdot \xi + \text{t.i.p.} + \mathcal{O}(||\xi||^3)$$

in the case of any $(y;\xi)$ satisfying (1.1). The fact that $\bar{y}$ is an optimal policy when

[18] A similar method is used by Sutherland (2002) to compute correct second-order approximations to welfare under alternative policies. However, his second-order approximation is computed for a particular parametric class of policies, while we derive a quadratic loss function that yields a correct welfare measure for *any* feasible policy.

the disturbances are zero implies that

$$D_y U = -\bar{\lambda}' D_y F, \tag{1.12}$$

where $\bar{\lambda}$ is a vector of Lagrange multipliers associated with the constraints (1.1) in the case of zero disturbances.It then follows that

$$\begin{aligned} D_y U \cdot \tilde{y} &= -\bar{\lambda}_I D_y F^I \cdot \tilde{y} \\ &= \frac{1}{2}\bar{\lambda}_I \tilde{y}' D^2_{yy} F^I \cdot \tilde{y} + \bar{\lambda}_I \tilde{y}' D^2_{y\xi} F^I \cdot \xi + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \end{aligned}$$

We can then use this expression to substitute for the term $D_y U \cdot \tilde{y}$ in (1.3), yielding

$$\begin{aligned} U(y;\xi) &= \frac{1}{2}\tilde{y}'[D^2_{yy}U \\ &\quad + \bar{\lambda}_I D^2_{yy} F^I] \cdot \tilde{y} + \tilde{y}'[D^2_{y\xi}U + \bar{\lambda}_I D^2_{y\xi} F^I] \cdot \xi + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \end{aligned}$$

This is an approximation of the form (1.11), where

$$\hat{U}(y;\xi) \equiv \frac{1}{2}\tilde{y}'[D^2_{yy}U + \bar{\lambda}_I D^2_{yy} F^I] \cdot \tilde{y} + \tilde{y}'[D^2_{y\xi}U + \bar{\lambda}_I D^2_{y\xi} F^I] \cdot \xi. \tag{1.13}$$

Use of the corrected quadratic objective (1.13) solves the problems associated with the use of U^Q discussed above. For example, the policy that maximizes (1.13) subject to the linearized constraints (1.5) satisfies linear first-order conditions of precisely the form (1.9). Hence this linear policy will represent a correct linear approximation to the optimal policy $y^{opt}(\xi)$. The objective (1.13) can also be used to correctly rank alternative policies (none of which need be fully optimal), as long as these policies imply that $y(0) = \bar{y}$. [19]

We have remarked above that an advantage of an LQ approximation (when valid) is that it makes it straightforward to verify that the solution to the LQ problem represents at least a local welfare maximum, by checking the second-order conditions for optimality. In our static example, the quadratic objective (1.13) is strictly concave in y if and only if the matrix of coefficients

$$D^2_{yy}U + \bar{\lambda}_I D^2_{yy} F^I \tag{1.14}$$

[19] Kim and Kim (2006) illustrate how the method expounded here can be used, for example, to correctly rank alternative policies with regard to international risk-sharing, in an example where naive LQ analysis sometimes gives an incorrect ranking.

is negative definite. In this case, the solution to the (linear) first-order conditions represents a global maximum of $\hat{U}$. Because our approximation is valid only locally, this only implies that the solution to the LQ problem approximates a *local* welfare maximum of the exact problem. Of course, under method (3) above, it would also have been possible to verify that the solution to the first-order conditions (1.8) represents a local maximum — and hence that the solution to the linearized conditions approximates a local maximum — by checking for local concavity in y of the Lagrangian

$$\mathcal{L}(y;\xi;\lambda) \equiv U(y;\xi) + \lambda F(y;\xi)$$

associated with the exact policy problem. This would involve checking for negative definiteness of the matrix $\mathcal{L}_{yy}(\bar{y};0;\bar{\lambda})$;[20] but this is just the matrix (1.14). Thus in order to check the second-order conditions under this method, one would have to compute the coefficients of the LQ objective function in any event. Recognizing that these define a quadratic approximation to the policy objective has the advantage of not only allowing one to compute a linear approximation to the solution to the first-order conditions for optimal policy and to verify the second-order conditions, but also providing a criterion with which to rank suboptimal policies.

The type of correct LQ approximation that we discuss here is not unknown to the economics literature; in an important early application of this method, Magill (1977) derives a correct LQ approximation to a multi-sector stochastic optimal growth model (in which, unlike the case treated by Kydland and Prescott, the constraints are not linear), using results due to Fleming (1971) in the literature on optimal control. These results are not directly applicable to the class of problems of interest to us (and frequently encountered in the literature on optimal stabilization policy), however, for two reasons: we work in discrete time, and we allow for forward-looking constraints (the equilibrium relations of a macro model derived from optimizing private-sector behavior), rather than assuming purely backward-looking evolution equations as in the standard (engineering) theory of optimal control. However, as we show here, a straightforward extension of the method to the kind of problems frequently encountered in the literature on optimal stabilization is possible, allowing a valid LQ

[20] Note that strict negative definiteness also implies that the matrix must be non-singular; this is the condition required for the first-order conditions (1.8) to have a determinate solution. Hence if one checks the second-order conditions, determinacy of the solution is guaranteed, as one would expect if each solution must be a local maximum in this case.

approximation of a fairly general class of discrete-time optimal policy problems.

2 LQ Approximation of a Problem with Forward-Looking Constraints

We now consider a general dynamic optimal policy problem. Suppose that the policy authority wishes to choose the evolution of a state vector $\{y_t\}$ for $t \geq t_0$ to maximize an objective of the form

$$V_{t_0} \equiv E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \pi(y_t, \xi_t), \tag{2.1}$$

where $0 < \beta < 1$ is a discount factor, the period objective $\pi(y, \xi)$ is a concave function of y, and ξ_t is a vector of exogenous disturbances. The evolution of the endogenous states must satisfy a system of backward-looking structural relations

$$F(y_t, \xi_t; y_{t-1}) = 0 \tag{2.2}$$

and a system of forward-looking structural relations

$$E_t g(y_t, \xi_t; y_{t+1}) = 0, \tag{2.3}$$

that both must hold for each $t \geq t_0$, given the vector of initial conditions y_{t_0-1}.

Conditions of the form (2.2) allow current endogenous variables to depend on lagged states; for example, these relations could include a technological relation between the capital stock carried into the next period, current investment expenditure, and the capital stock carried into the current period.[21] Conditions of the form (2.3) instead allow current endogenous variables to depend on current expectations regarding future states; for example, these relations could include an Euler equation for the optimal timing of consumer expenditure, relating current consumption to expected consumption in the next period and the expected rate of return on saving.[22] While

[21] The next period's capital stock and the current investment expenditure would both be elements of y_t; the vector ξ_t could include a random disturbance to investment adjustment costs.

[22] Current consumption and the current period ex-post return on saving in the previous period would both be elements of y_t; the vector ξ_t could include a random disturbance to the impatience to consume. Note that without loss of generality we may suppose that the vector ξ_t includes all information available in period t regarding future exogenous disturbances.

the most general notation would allow both leads and lags in all of the structural equations, supposing that there are equations of these two types will make clearer the different types of complications arising from the two distinct types of intertemporal linkages. We shall suppose that the number n_F of constraints of the first type each period plus the number n_g of constraints of the second type is less than the number n_y of endogenous state variables each period, so that there is at least one dimension along which policy can continuously vary the outcome y_t each period, even the past and expected future evolution of the endogenous variables. A t_0-optimal commitment (the standard Ramsey policy problem) is then the state-contingent evolution $\{y_t\}$ consistent with equations (2.2)–(2.3) for all $t \geq t_0$ that maximizes (2.1).

2.1 A Recursive Policy Problem

As is well-known, the presence of the forward-looking constraints (2.3) implies that a t_0-optimal commitment is not generally time-consistent. If, however, we suppose that a policy to apply from period t_0 onward must be chosen subject to an additional set of constraints on the acceptable values of y_{t_0}, it is possible for the resulting policy problem to have a recursive structure. As discussed in Benigno and Woodford (2003, 2005a), we wish to choose initial pre-commitments regarding y_{t_0} that are *self-consistent,* in the sense that the policy that is chosen subject to these constraints would also satisfy constraints of exactly the same form in all later periods as well. The required initial pre-commitments are of the form

$$g(y_{t_0-1}, \xi_{t_0-1}; y_{t_0}) = \bar{g}_{t_0}, \tag{2.4}$$

where $\bar{g}_{t_0}$ may depend on the exogenous state at date t_0. Note that we assume the existence of a pre-commitment only about those aspects of y_{t_0} the anticipation of which back in period $t_0 - 1$ should have been relevant to equilibrium determination then; there is no need for any stronger form of commitment in order to render optimal policy time-consistent.

We are thus interested in characterizing the state-contingent policy $\{y_t\}$ for $t \geq t_0$ that maximizes (2.1) subject to constraints (2.2) – (2.4). Such a policy is *optimal from a timeless perspective* if $\bar{g}_{t_0}$ is chosen, as a function of predetermined or exogenous states at t_0, according to a self-consistent rule.[23] This means that the initial pre-

[23]See Giannoni and Woodford (2002), Woodford (2003, chap. 7), or Benigno and Woodford (2005a) for further discussion.

commitment is determined by past conditions through a function

$$\bar{g}_{t_0} = \bar{g}(\xi_{t_0}, \mathbf{y_{t_0-1}}), \tag{2.5}$$

where $\mathbf{y_t}$ is an extended state vector;[24] this function has the property that under optimal policy, given this initial pre-commitment, the state-contingent evolution of the economy will satisfy

$$g(y_{t-1}, \xi_{t-1}; y_t) = \bar{g}(\xi_t, \mathbf{y_{t-1}}) \tag{2.6}$$

in each possible state of the world at each date $t \geq t_0$ as well. Thus the initial constraint is of a form that one would optimally commit oneself to satisfy at all (subsequent) dates.

Let $V(\bar{g}_{t_0}; y_{t_0-1}, \xi_{t_0}, \xi_{t_0-1})$ be the maximum achievable value of the objective (2.1) in this problem.[25] Then the infinite-horizon problem just defined is equivalent to a sequence of one-period decision problems in which, in each period $t \geq t_0$, a value of y_t is chosen and state-contingent one-period-ahead pre-commitments $\bar{g}_{t+1}(\xi_{t+1})$ (for each of the possible states ξ_{t+1} in the following period) are chosen so as to maximize

$$\pi(y_t, \xi_t) + \beta E_t V(\bar{g}_{t+1}; y_t, \xi_{t+1}, \xi_t), \tag{2.7}$$

subject to the constraints

$$F(y_t,\xi_t; y_{t-1}) = 0,$$

$$g(y_{t-1}, \xi_{t-1}; y_t) = \bar{g}_t,$$

$$E_t \bar{g}_{t+1} = 0,$$

given the values of $\bar{g}_t, y_{t-1}, \xi_{t-1}$, and ξ_t, all of which are predetermined and/or exogenous in period t. It is this recursive policy problem that we wish to study; note that it is only when we consider this problem (as opposed to the unconstrained Ramsey

[24] The extended state vector may include both endogenous and exogenous variables, the values of which are realized in period t or earlier. More specific assumptions about the nature of the extended state vector are made below; see the discussion of equation (2.8).

[25] We assume, to economize on notation, that the exogenous state vector ξ_t evolves in accordance with a Markov process. Hence ξ_t summarizes not only all of the disturbances that affect the structural relations at date t, but all information at date t about the subsequent evolution of the exogenous disturbances. This is important in order for a time-invariant value function to exist with the arguments indicated.

problem) that it is possible, in general, to obtain a deterministic steady state as an optimum in the case of suitable initial conditions, and hence only in this case that we can hope to approximate the optimal policy problem around such a steady state.[26]

The solution to the recursive policy problem just defined involves values for the endogenous variables y_t given by a policy function of the form

$$y_t = y^*(\bar{g}_t, y_{t-1}, \xi_t, \xi_{t-1}),$$

and a choice of the following period's pre-commitment $\bar{g}_{t+1}$ of the form

$$\bar{g}_{t+1} = g^*(\xi_{t+1}; \bar{g}_t, y_{t-1}, \xi_t, \xi_{t-1}),$$

where y^* and g^* are time-invariant functions. Let us suppose furthermore that the evolution of the extended state vector depends only on the evolution of the two vectors $\{y_t, \xi_t\}$, through a recursion of the form

$$\mathbf{y_t} = \psi(\xi_t, y_t, \mathbf{y_{t-1}}); \tag{2.8}$$

this system of identities *defines* the extended state vector, the elements of which consist essentially of linear combinations of current and lagged elements of the vectors y_t and ξ_t. (To simplify notation, we shall suppose that the current values y_t and ξ_t are among the elements of $\mathbf{y}_t$.) The initial pre-commitment (2.5) is then self-consistent if

$$g^*(\xi_{t+1}; \bar{g}(\xi_t, \mathbf{y_{t-1}}), y_{t-1}, \xi_t, \xi_{t-1}) = \bar{g}(\xi_{t+1}, \psi(\xi_t, y^*(\bar{g}_t, y_{t-1}, \xi_t, \xi_{t-1}), \mathbf{y_{t-1}})) \tag{2.9}$$

for all possible values of ξ_{t+1}, ξ_t, and $\mathbf{y_{t-1}}$.[27] Note that this implies that equation (2.6) is satisfied at all times.

2.2 A Correct LQ Local Approximation

As in the static problem treated in the previous section, our method involves a local approximation to both the objective and the constraints, near an optimal policy for

[26] In the literature on Ramsey policy, one sometimes sees approximate characterizations of optimal policy computed by log-linearizing around a steady state that Ramsey policy approaches asymptotically in the absence of random disturbances. But in such a case, there is no guarantee that the approximate characterization would be accurate even in the case of arbitrarily small disturbances, as Ramsey policy need not be near the steady state except asymptotically.

[27] Both sides of this equation involve only the elements of ξ_{t+1}, ξ_t, and $\mathbf{y_{t-1}}$, on the understanding that y_{t-1} and ξ_{t-1} are both elements of $\mathbf{y_{t-1}}$.

the case of zero disturbances. We furthermore assume both an initial state y_{t_0-1} and initial pre-commitments $\bar{g}_{t_0}$ such that the optimal policy in the case of zero disturbances is a steady state, *i.e.*, such that $y_t = \bar{y}$ for all t, for some vector $\bar{y}$. (More precisely, our calculations below assume that both y_{t_0-1} and $\bar{g}_{t_0-1}$ are *close enough* to being consistent with this steady state.) In order to define this steady state, we must consider the nature of optimal policy in the exact problem just defined.

The first-order conditions for the exact policy problem can obtained by differentiating a Lagrangian of the form

$$\mathcal{L}_{t_0} = V_{t_0} + E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\lambda_t' F(y_t, \xi_t; y_{t-1}) + \beta^{-1} \varphi_{t-1}' g(y_{t-1}, \xi_{t-1}; y_t) \right], \qquad (2.10)$$

where λ_t and φ_t are Lagrange multipliers associated with constraints (2.2) and (2.3) respectively, for any date $t \geq t_0$, and we use the notation $\beta^{-1}\varphi_{t_0-1}$ for the Lagrange multiplier associated with the additional constraint (2.4). This last notational choice allows the first-order conditions to be expressed in the same way for all periods. Optimality requires that the joint evolution of the processes $\{y_t, \xi_t, \lambda_t, \varphi_t\}$ satisfy

$$\begin{aligned} D_y\pi(y_t, \xi_t) + {\lambda_t}' D_y F(y_t, \xi_t; y_{t-1}) + \beta E_t {\lambda_{t+1}}' D_{\check{y}} F(y_{t+1}, \xi_{t+1}; y_t) & \\ + E_t {\varphi_t}' D_y g(y_t, \xi_t; y_{t+1}) + \beta^{-1} {\varphi_{t-1}}' D_{\hat{y}} g(y_{t-1}, \xi_{t-1}; y_t) &= 0 \end{aligned} \qquad (2.11)$$

at each date $t \geq t_0$, where D_y denotes the vector of partial derivatives of any of the functions with respect to the elements of y_t, while $D_{\hat{y}}$ means the vector of partial derivatives with respect to the elements of y_{t+1} and $D_{\check{y}}$ means the vector of partial derivatives with respect to the elements of y_{t-1}.

An *optimal steady state* is then described by a collection of vectors $(\bar{y}, \bar{\lambda}, \bar{\varphi})$ satisfying

$$\begin{aligned} D_y\pi(\bar{y}, 0) + \bar{\lambda}' D_y F(\bar{y}, 0; \bar{y}) + \beta \bar{\lambda}' D_{\check{y}} F(\bar{y}, 0; \bar{y}) & \\ + \bar{\varphi}' D_y g(\bar{y}, 0; \bar{y}) + \beta^{-1} \bar{\varphi}' D_{\hat{y}} g(\bar{y}, 0; \bar{y}) &= 0, \end{aligned} \qquad (2.12)$$

$$F(\bar{y}, 0; \bar{y}) = 0, \qquad (2.13)$$

$$g(\bar{y}, 0; \bar{y}) = 0. \qquad (2.14)$$

We shall suppose that such a steady state exists, and assume (in the policy problem with random disturbances) an initial state y_{t_0-1} near $\bar{y}$ — more precisely, such that $y_{t_0-1} - \bar{y} = \mathcal{O}(||\xi||)$ — and an initial pre-commitment such that $\bar{g}_{t_0} = \mathcal{O}(||\xi||)$ as well.[28] Once the optimal steady state has been computed, we make no further use of conditions (2.11); our proposed method does not require that we directly seek to solve these equations.

Instead, we now consider local approximations to the objective and constraints near an optimal steady state. We can compute a second-order Taylor expansion of the period objective function π, obtaining an expression of exactly the form (1.3). Substituting this into (2.1), we obtain the approximate objective

$$V_{t_0} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[D_y\pi \cdot \tilde{y}_t + \frac{1}{2}\tilde{y}_t' D_{yy}^2\pi \cdot \tilde{y}_t + \tilde{y}_t' D_{y\xi}^2\pi \cdot \xi_t \right] + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \quad (2.15)$$

This would be used as the quadratic objective in what we have called the "naive" LQ approximation. Under our alternative approach, we must substitute purely quadratic terms for the linear terms $D_y\pi \cdot \tilde{y}_t$ in this sum.

A similar second-order Taylor series approximation can be written for each of the functions F^k. It follows that

$$\begin{aligned}
\sum_{t=t_0}^{\infty} \beta^{t-t_0} \bar{\lambda}' F(y_t, \xi_t; y_{t-1}) \;=\;& \sum_{t=t_0}^{\infty} \beta^{t-t_0} \Big\{ \bar{\lambda}' [D_y F \cdot \tilde{y}_t + D_{\check{y}} F \cdot \tilde{y}_{t-1}] \\
& + \bar{\lambda}_k \Big[\frac{1}{2} \tilde{y}_t' D_{yy}^2 F^k \cdot \tilde{y}_t + \tilde{y}_t' D_{y\xi}^2 F^k \cdot \xi_t + \tilde{y}_{t-1}' D_{\check{y}\xi}^2 F^k \cdot \xi_t \\
& + \frac{1}{2} \tilde{y}_{t-1}' D_{\check{y}\check{y}}^2 F^k \cdot \tilde{y}_{t-1} + \tilde{y}_t' D_{y\check{y}}^2 F^k \cdot \tilde{y}_{t-1} \Big] \Big\} \\
& + \text{t.i.p.} + \mathcal{O}(||\xi||^3) \\
=\;& \sum_{t=t_0}^{\infty} \beta^{t-t_0} \Big\{ \bar{\lambda}' [D_y F + \beta D_{\check{y}} F] \cdot \tilde{y}_t \\
& + \frac{1}{2} \bar{\lambda}_k \big[\tilde{y}_t' D_{yy}^2 F^k \cdot \tilde{y}_t + 2 \tilde{y}_t' D_{y\xi}^2 F^k \cdot \xi_t + 2\beta \tilde{y}_t' D_{\check{y}\xi}^2 F^k \cdot \xi_{t+1} \\
& + \beta \tilde{y}_t' D_{\check{y}\check{y}}^2 F^k \cdot \tilde{y}_t + 2 \tilde{y}_t' D_{y\check{y}}^2 F^k \cdot \tilde{y}_{t-1} \big] \Big\} \\
& + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \qquad (2.16)
\end{aligned}$$

Using a similar Taylor series approximation of each of the functions g^i, we corre-

[28] Note that the steady-state value of $\bar{g}$ is equal to $g(\bar{y}, 0; \bar{y}) = 0$.

spondingly obtain

$$
\begin{aligned}
\sum_{t=t_0}^{\infty} \beta^{t-t_0-1} \bar{\varphi}' g(y_{t-1}, \xi_{t-1}; y_t) &= \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \bar{\varphi}' [D_y g + \beta^{-1} D_{\hat{y}} g] \cdot \tilde{y}_t \right. \\
&\quad + \frac{1}{2} \bar{\varphi}_i \left[\tilde{y}_t' D^2_{yy} g^i \cdot \tilde{y}_t + 2 \tilde{y}_t' D^2_{y\xi} g^i \cdot \xi_t + 2\beta^{-1} \tilde{y}_t' D^2_{\hat{y}\xi} g^i \cdot \xi_{t-1} \right. \\
&\quad \left. \left. + \beta^{-1} \tilde{y}_t' D^2_{\hat{y}\hat{y}} g^i \cdot \tilde{y}_t + 2\beta^{-1} \tilde{y}_t' D^2_{\hat{y}y} g^i \cdot \tilde{y}_{t-1} \right] \right\} \\
&\quad + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \qquad (2.17)
\end{aligned}
$$

It then follows from constraints (2.2)–(2.4) that in the case of any admissible policy,[29]

$$
\begin{aligned}
\beta^{-1} \bar{\varphi}' \bar{g}_{t_0} &= E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} [\bar{\lambda}' F(y_t, \xi_t; y_{t-1}) + \beta^{-1} \bar{\varphi}' g(y_{t-1}, \xi_{t-1}; y_t)] \\
&= E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ [\bar{\lambda}' (D_y F + \beta D_{\check{y}} F) + \bar{\varphi}' (D_y g + \beta^{-1} D_{\hat{y}} g)] \cdot \tilde{y}_t \right. \\
&\quad + \frac{1}{2} \bar{\lambda}_k \left[\tilde{y}_t' D^2_{yy} F^k \cdot \tilde{y}_t + 2 \tilde{y}_t' D^2_{y\xi} F^k \cdot \xi_t + 2\beta \tilde{y}_t' D^2_{\check{y}\xi} F^k \cdot \xi_{t+1} \right. \\
&\quad \left. + \beta \tilde{y}_t' D^2_{\check{y}\check{y}} F^k \cdot \tilde{y}_t + 2 \tilde{y}_t' D^2_{y\check{y}} F^k \cdot \tilde{y}_{t-1} \right] \\
&\quad + \frac{1}{2} \bar{\varphi}_i \left[\tilde{y}_t' D^2_{yy} g^i \cdot \tilde{y}_t + 2 \tilde{y}_t' D^2_{y\xi} g^i \cdot \xi_t + 2\beta^{-1} \tilde{y}_t' D^2_{\hat{y}\xi} g^i \cdot \xi_{t-1} \right. \\
&\quad \left. \left. + \beta^{-1} \tilde{y}_t' D^2_{\hat{y}\hat{y}} g^i \cdot \tilde{y}_t + 2\beta^{-1} \tilde{y}_t' D^2_{\hat{y}y} g^i \cdot \tilde{y}_{t-1} \right] \right\} \\
&\quad + \text{t.i.p.} + \mathcal{O}(||\xi||^3), \qquad (2.18)
\end{aligned}
$$

where we have used (2.16) and (2.17) to substitute for the F and g terms respectively. We can write this more compactly in the form

$$
\begin{aligned}
\beta^{-1} \bar{\varphi}' \bar{g}_{t_0} &= E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \Phi \cdot \tilde{y}_t + \frac{1}{2} \left[\tilde{y}_t' H \cdot \tilde{y}_t + 2 \tilde{y}_t' R \tilde{y}_{t-1} + 2 \tilde{y}_t' Z(L) \xi_{t+1} \right] \right\} \\
&\quad + \text{t.i.p.} + \mathcal{O}(||\xi||^3), \qquad (2.19)
\end{aligned}
$$

where

$$
\Phi \equiv \bar{\lambda}' [D_y F + \beta D_{\check{y}} F] + \bar{\varphi}' [D_y g + \beta^{-1} D_{\hat{y}} g],
$$

$$
H \equiv \bar{\lambda}_k [D^2_{yy} F^k + \beta D^2_{\check{y}\check{y}} F^k] + \bar{\varphi}_i [D^2_{yy} g^i + \beta^{-1} D^2_{\hat{y}\hat{y}} g^i],
$$

[29] Note that we here include (2.4) among the constraints that a policy must satisfy. We shall call any evolution that satisfies (2.2)–(2.3) a "feasible" policy. Under this weaker assumption, the left-hand sides of (2.18) and (2.19) must instead be replaced by $\beta^{-1} \bar{\varphi}' g(y_{t_0-1}, \xi_{t_0-1}; y_{t_0})$.

$$R \equiv \bar{\lambda}_k D^2_{y\tilde{y}} F^k + \bar{\varphi}_i \beta^{-1} D^2_{\hat{y}y} g^i,$$

$$Z(L) \equiv \beta\bar{\lambda}_k D^2_{\tilde{y}\xi} F^k + (\bar{\lambda}_k D^2_{y\xi} F^k + \bar{\varphi}_i D^2_{y\xi} g^i) \cdot L + \beta^{-1}\bar{\varphi}_i D^2_{\hat{y}\xi} g^i \cdot L^2.$$

Using (2.12), we furthermore observe that[30]

$$\Phi = -D_y \pi.$$

With this substitution in (2.19), we obtain an expression that can be solved for

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} D_y \pi \cdot \tilde{y}_t,$$

which can in turn be used to substitute for the linear terms in (2.15). We thus obtain an alternative quadratic approximation to (2.1),[31]

$$V_{t_0} = \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\tilde{y}_t' Q \cdot \tilde{y}_t + 2\tilde{y}_t' R \tilde{y}_{t-1} + 2\tilde{y}_t' B(L) \xi_{t+1} \right] + \text{t.i.p.} + \mathcal{O}(||\xi||^3), \quad (2.20)$$

where now

$$Q \equiv D^2_{yy} \pi + H,$$

$$B(L) \equiv Z(L) + D^2_{y\xi} \pi \cdot L. \quad (2.21)$$

Since (2.20) involves no linear terms, it can be evaluated (up to a residual of order $\mathcal{O}(||\xi||^3)$) using only a linear approximation to the evolution of $\tilde{y}_t$ under a given policy rule.

It follows that a correct LQ approximation to the original problem is given by the problem of choosing a state-contingent evolution $\{\tilde{y}_t\}$ for $t \geq t_0$ to maximize the objective

$$V^Q_{t_0}(\tilde{y}; \xi) \equiv \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\tilde{y}_t' A(L) \tilde{y}_t + 2\tilde{y}_t' B(L) \xi_{t+1} \right] \quad (2.22)$$

subject to the constraints that

$$C(L)\tilde{y}_t = f_t, \quad (2.23)$$

[30] This is the point at which our calculations rely on the assumption that the steady state around which we compute our local approximations is optimal.

[31] Here we include $\bar{g}_{t_0}$ among the "terms independent of policy." If we consider also policies that are not necessarily consistent with the initial pre-commitment, the left-hand side of (2.20) is more generally equal to $V_{t_0} + \beta^{-1}\bar{\varphi}' g(y_{t_0-1}, \xi_{t_0-1}; y_{t_0})$. This generalization of (2.20) is used in the derivation of equation (4.3) below.

$$E_t D(L)\tilde{y}_{t+1} = h_t \tag{2.24}$$

for all $t \geq t_0$, and the additional initial constraint that

$$D(L)\tilde{y}_{t_0} = \tilde{h}_{t_0}, \tag{2.25}$$

where now

$$A(L) \equiv Q + 2R \cdot L, \tag{2.26}$$

$$C(L) \equiv D_y F + D_{\tilde{y}} F \cdot L, \tag{2.27}$$

$$f_t \equiv -D_\xi F \cdot \xi_t,$$

$$D(L) \equiv D_{\hat{y}} g + D_y g \cdot L, \tag{2.28}$$

$$h_t \equiv -D_\xi g \cdot \xi_t, \tag{2.29}$$

$$\tilde{h}_{t_0} \equiv h_{t_0-1} + \bar{g}_{t_0}.$$

2.3 An Equivalent Lagrangian Approach

In the case that the objective (2.22) is concave,[32] the first-order conditions associated with the LQ problem just defined characterize the solution to that problem. Here we show that these linear equations also correspond to a local linear approximation to the first-order conditions associated with the exact problem, *i.e.*, the modified Ramsey policy problem defined in section 2.1, and hence that the solution to the LQ problem represents a local linear approximation to optimal policy from a timeless perspective.[33]

As already noted, the first-order conditions for the exact policy problem are obtained by differentiating the Lagrangian $\mathcal{L}_{t_0}$ defined in (2.10). This yields the system of first-order conditions (2.11). The linearization of these first-order conditions around the optimal steady state is in turn the set of linear equations that would be obtained by differentiating a quadratic approximation to $\mathcal{L}_{t_0}$ around that same steady state. Hence we are interested in computing such a local approximation, for the case in which $y_t - \bar{y}$, $\lambda_t - \bar{\lambda}$, and $\varphi_t - \bar{\varphi}$ are each of order $\mathcal{O}(||\xi||)$ for all t. (Here the

[32]The algebraic conditions under which this is so are discussed in the next section.

[33]See also Levine *et al.* (2007) for a similar discussion of the equivalence between our approach and the Lagrangian approach.

steady-state values of the Lagrange multipliers $\bar{\lambda}, \bar{\varphi}$ are again given by the solution to equations (2.12) – (2.14).)

We may furthermore write the Lagrangian in the form

$$\mathcal{L}_{t_0} = \bar{\mathcal{L}}_{t_0} + \tilde{\mathcal{L}}_{t_0},$$

where

$$\bar{\mathcal{L}}_{t_0} = V_{t_0} + E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\bar{\lambda}' F(y_t, \xi_t; y_{t-1}) + \beta^{-1} \bar{\varphi}' g(y_{t-1}, \xi_{t-1}; y_t) \right],$$

$$\tilde{\mathcal{L}}_{t_0} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\tilde{\lambda}_t' F(y_t, \xi_t; y_{t-1}) + \beta^{-1} \tilde{\varphi}_{t-1}' g(y_{t-1}, \xi_{t-1}; y_t) \right],$$

$$\tilde{\lambda}_t \equiv \lambda_t - \bar{\lambda}, \qquad \tilde{\varphi}_t \equiv \varphi_t - \bar{\varphi}.$$

We can then use equations (2.15) and (2.18) to show that the local quadratic approximation to $\bar{\mathcal{L}}_{t_0}$ is given by[34]

$$\bar{\mathcal{L}}_{t_0} = V_{t_0}^Q + \text{t.i.p.} + \mathcal{O}(||\xi||^3).$$

In addition, the fact that $\tilde{\lambda}_t, \tilde{\varphi}_t$ are both of order $\mathcal{O}(||\xi||)$ means that a local quadratic approximation to the other term is given by

$$\tilde{\mathcal{L}}_{t_0} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\tilde{\lambda}_t' \tilde{F}(y_t, \xi_t; y_{t-1}) + \beta^{-1} \tilde{\varphi}_{t-1}' \tilde{g}(y_{t-1}, \xi_{t-1}; y_t) \right] + \mathcal{O}(||\xi||^3),$$

where $\tilde{F}$ and $\tilde{g}$ are local linear approximations to the functions F and g respectively.

Hence the local quadratic approximation to the complete Lagrangian is given by

$$\begin{aligned} \mathcal{L}_{t_0} &= V_{t_0}^Q + E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\tilde{\lambda}_t' \tilde{F}(y_t, \xi_t; y_{t-1}) + \beta^{-1} \tilde{\varphi}_{t-1}' \tilde{g}(y_{t-1}, \xi_{t-1}; y_t) \right] \\ &\quad + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \end{aligned} \tag{2.30}$$

But this is identical (up to terms independent of policy) to the Lagrangian for the LQ problem of maximizing $V_{t_0}^Q$ subject to the linearized constraints. Hence the first-order conditions obtained from this approximate Lagrangian (which coincide with the local linear approximation to the first-order conditions for the exact problem) are identical to the first-order conditions for the LQ problem, and their solutions are identical as well.

[34] It is worth noting that this equality holds in the case of all feasible policies, whether or not the policy is consistent with the initial pre-commitment (2.4). This is important for our discussion of the welfare evaluation of suboptimal policies in section 4.

3 Characterizing Optimal Policy

We now study necessary and sufficient conditions for a policy to solve the LQ problem of maximizing (2.22) subject to constraints (2.23) – (2.25). Let $\mathcal{H}$ be the Hilbert space of (real-valued) stochastic processes $\{\tilde{y}_t\}$ such that

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \tilde{y}_t' \tilde{y}_t < \infty. \tag{3.1}$$

We are interested in solutions to the LQ problem that satisfy the bound (3.1) because it guarantees that the objective V^Q is well-defined (and is generically required for it to be so). Of course, our LQ approximation to the original problem is only guaranteed to be accurate in the case that $\tilde{y}_t$ is always sufficiently small; hence a solution to the LQ problem in which $\tilde{y}_t$ grows without bound, but at a slow enough rate for (3.1) to be satisfied, need not correspond (even approximately) to any optimum (or local optimum) of the exact problem. In this section, however, we take the LQ problem at face value, and discuss the conditions under which it has a solution, despite the fact that we should in general only be interested in bounded solutions.

3.1 A Lagrangian Approach

The Lagrangian for this problem is given by

$$\mathcal{L}_{t_0}^Q = \frac{1}{2}\left\{ E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left[\tilde{y}_t' A(L) \tilde{y}_t + 2\tilde{y}_t' B(L) \xi_{t+1} + 2\tilde{\lambda}_t' C(L) \tilde{y}_t \right.\right.$$
$$\left.\left. + 2\beta^{-1} \tilde{\varphi}_{t-1}' D(L) \tilde{y}_t \right] \right\}.$$

(Note that this is just (2.30), omitting the terms independent of policy and those of third or higher order.) Differentiation of the Lagrangian then yields a system of linear first-order conditions

$$\frac{1}{2} E_t\{[A(L) + A'(\beta L^{-1})]\tilde{y}_t\} + E_t[B(L)\xi_{t+1}]$$
$$+ E_t[C'(\beta L^{-1})\tilde{\lambda}_t] + \beta^{-1} D'(\beta L^{-1}) \tilde{\varphi}_{t-1} = 0 \tag{3.2}$$

that must hold for each $t \geq t_0$ under an optimal policy. (Here we use the notation M' for the transpose of a matrix M.) These conditions, together with (2.23) – (2.25), form a linear system to be solved for the joint evolution of the processes $\{\tilde{y}_t, \tilde{\lambda}_t, \tilde{\varphi}_t\}$ given the exogenous disturbance processes $\{\xi_t\}$ and the initial conditions $\tilde{y}_{t_0-1}$ and the initial pre-commitment $\bar{g}_{t_0}$ (or $\hat{h}_{t_0}$). This type of system of linear stochastic difference equations is easy to solve using standard methods.[35]

The first-order conditions (3.2) are easily shown to be *necessary* for optimality, but they are not generally *sufficient* for optimality as well; one must also verify that second-order conditions for optimality are satisfied. (In the case of an LQ problem, satisfaction of the second-order conditions implies global, and not just local, optimality; so we need not check any further conditions. But because our LQ problem is only a local approximation to the original policy problem, a global optimum of the LQ problem still may only correspond to a local optimum of the exact problem.) We next consider these additional conditions.

Let us consider the subspace $\mathcal{H}_1 \subset \mathcal{H}$ of processes $\hat{y} \in \mathcal{H}$ that satisfy the additional constraints

$$C(L)\hat{y}_t = 0 \tag{3.3}$$

$$E_t D(L)\hat{y}_{t+1} = 0 \tag{3.4}$$

for each date $t \geq t_0$, along with the initial commitments

$$D(L)\hat{y}_{t_0} = 0, \tag{3.5}$$

where we define $\hat{y}_{t_0-1} \equiv 0$ in writing (3.3) for period $t = t_0$ and in writing (3.5). This subspace is of interest because if a process $\tilde{y} \in \mathcal{H}$ satisfies constraints (2.23) – (2.25), another process $y \in \mathcal{H}$ with $y_{t_0-1} = \tilde{y}_{t_0-1}$ satisfies those constraints as well if and only if $y - \tilde{y} \in \mathcal{H}_1$. We may now state our first main result.

Proposition 1 *For $\{\tilde{y}_t\} \in \mathcal{H}$ to maximize the quadratic form (2.22), subject to the constraints (2.23) – (2.25) given initial conditions $\tilde{y}_{t_0-1}$ and $\bar{g}_{t_0}$, it is necessary and sufficient that (i) there exist Lagrange multiplier processes*[36] *$\tilde{\varphi}, \tilde{\lambda} \in \mathcal{H}$ such that the processes $\{\tilde{y}_t, \tilde{\varphi}_t, \tilde{\lambda}_t\}$ satisfy (3.2) for each $t \geq t_0$; and (ii)*

$$V^Q(\hat{y}) \equiv V^Q_{t_0}(\hat{y}; 0) = \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} [\hat{y}_t' A(L) \hat{y}_t] \leq 0 \tag{3.6}$$

[35]See, for example, Giannoni and Woodford (2002) for discussion of the solution of an equation system of this form using an eigenvector-decomposition method.

[36]Note that $\tilde{\varphi}_t$ is also assumed to be defined for $t = t_0 - 1$.

for all processes $\hat{y} \in \mathcal{H}_1$, where in evaluating (3.6) we define $\hat{y}_{t_0-1} \equiv 0$. A process $\{\tilde{y}_t\}$ with these properties is furthermore uniquely optimal if and only if

$$V^Q(\hat{y}) < 0 \tag{3.7}$$

for all processes $\hat{y} \in \mathcal{H}_1$ that are non-zero almost surely.

The proof is given in the Appendix. The case in which the stronger condition (3.7) holds — *i.e.*, the quadratic form $V^Q(\hat{y})$ is negative definite on the subspace $\mathcal{H}_1$ — is the one of primary interest to us, since it is in this case that we know that the process $\{\tilde{y}_t\}$ represents at least a local welfare maximum in the exact problem. In this case we can also show that pure randomization of policy reduces the welfare objective (2.22), and hence is locally welfare-reducing in the exact problem as well, as is discussed further in Benigno and Woodford (2005a).

3.2 A Dynamic Programming Approach

We can furthermore establish a useful characterization of the algebraic conditions under which the second-order conditions (3.7) are satisfied. These are most easily developed by considering the recursive formulation of our optimal policy problem presented in section 2.1.[37] Let us suppose that the exogenous state vector ξ_t evolves according to a linear law of motion

$$\xi_{t+1} = \Gamma\, \xi_t + \epsilon_{t+1}, \tag{3.8}$$

where Γ is a matrix, all of the eigenvalues of which have modulus less than $\beta^{-1/2}$, and $\{\epsilon_t\}$ is an i.i.d. vector-valued random sequence, drawn each period from a distribution with mean zero and a variance-covariance matrix Σ.[38] In this case, our LQ

[37] This section has been improved by the suggestions of Paul Levine and Joe Pearlman.

[38] These assumptions ensure that the process $\{\xi_t\}$ satisfies a bound of the form (3.1). If we further wish to ensure that the disturbances are bounded, so that our local approximations can be expected to be accurate in the event of small enough disturbances, we may assume further that all eigenvalues of Γ have a modulus less than 1, and that ϵ_{t+1} is drawn from a distribution with bounded support. We may assume that, like the other structural relations in this section, (3.8) is merely a local linear approximation. Finally, note that the assumption of a law of motion of the form (3.8) allows for disturbances with arbitrarily complex forms of serial correlation, simply by adding elements to the vector ξ_t reflecting past exogenous states.

approximate policy problem has a recursive formulation, in which the continuation problem from any period t forward depends on the extended state vector

$$\mathbf{z}_t \equiv \begin{bmatrix} \tilde{y}_{t-1} \\ \tilde{h}_t \\ \xi_t \\ \xi_{t-1} \end{bmatrix}. \tag{3.9}$$

Let $\bar{V}^Q(\mathbf{z}_t)$ denote the maximum attainable value of the continuation objective V_t^Q, if the process $\{\tilde{y}_\tau\}$ from date t onward is chosen to satisfy constraints (2.23)–(2.24) for all $\tau \geq t$, an initial precommitment of the form

$$D(L)\tilde{y}_t = \tilde{h}_t, \tag{3.10}$$

and the bound (3.1). As usual in an LQ problem of this form, it can be shown that the value function is a quadratic function of the extended state vector,

$$\bar{V}^Q(\mathbf{z}_t) = \frac{1}{2}\mathbf{z}_t' P \mathbf{z}_t, \tag{3.11}$$

where P is a symmetric matrix to be determined. In characterizing the solution to the problem, it is useful to introduce notation for partitions of the matrix P. Let P_{ij} (for $i, j = 1, 2, 3, 4$) be the 16 blocks obtained when P is partitioned in both directions conformably with the partition of $\mathbf{z}_t$ in (3.9), and let

$$\mathbf{P}_i \equiv [P_{i1}\ P_{i2}\ P_{i3}\ P_{i4}]$$

(for $i = 1, 2, 3, 4$) be the four blocks obtained when P is partitioned only vertically.

In the recursive formulation of the approximate LQ problem, in each period t, $\tilde{y}_t$ is chosen, and a precommitment $\tilde{h}_{t+1}(\xi_{t+1})$ is chosen for each possible state in the period $t+1$ continuation, so as to maximize

$$\frac{1}{2}\tilde{y}_t' A(L)\tilde{y}_t + E_t[\tilde{y}_t' B(L)\xi_{t+1}] + \beta E_t \bar{V}^Q(\mathbf{z}_{t+1}), \tag{3.12}$$

subject to the constraints that $\tilde{y}_t$ satisfy (2.23) and (3.10), and that the choices of $\{\tilde{h}_{t+1}(\xi_{t+1})\}$ satisfy

$$E_t\tilde{h}_{t+1} = h_t. \tag{3.13}$$

To simplify the discussion, we shall further assume that

$$\text{rank} \begin{bmatrix} C_0 \\ D_0 \end{bmatrix} = n_F + n_g, \tag{3.14}$$

where here and below we write lag polynomials in the form $X(L) = \sum_j X_j L^j$. This condition implies that the constraints (2.23) and (3.10) include neither any redundant constraints nor any constraints that are inconsistent in the case of a generic state $\mathbf{z}_t$.

The first-order conditions for the optimal choice of $\tilde{y}_t$ in this single-period problem are of the form

$$[A_0 + (1/2)A_1 L]\tilde{y}_t + E_t[B(L)\xi_{t+1}] + \beta\mathbf{P}_1\, E_t\mathbf{z}_{t+1} + C_0'\tilde{\lambda}_t + D_0'\tilde{\psi}_t = 0, \qquad (3.15)$$

where $\tilde{\lambda}_t, \tilde{\psi}_t$ are the Lagrange multipliers associated with constraints (2.23) and (3.10) respectively. Condition (3.15) together with the constraints (2.23) and (3.10) constitute a system of $n = n_y + n_F + n_g$ linear equations to solve for $\tilde{y}_t, \tilde{\lambda}_t$, and $\tilde{\psi}_t$ as functions of $\mathbf{z}_t$. This system can be written in the matrix form $M\, y_t^\dagger = -G\,\mathbf{z}_t$, where

$$M \equiv \begin{bmatrix} A_0 + \beta P_{11} & C_0' & D_0' \\ C_0 & 0 & 0 \\ D_0 & 0 & 0 \end{bmatrix}, \qquad y_t^\dagger \equiv \begin{bmatrix} \tilde{y}_t \\ \tilde{\lambda}_t \\ \tilde{\psi}_t \end{bmatrix}, \qquad (3.16)$$

and G is a matrix of coefficients, the first two columns of which (of particular interest here) are

$$G_1 \equiv \begin{bmatrix} (1/2)A_1 \\ C_1 \\ D_1 \end{bmatrix}, \qquad G_2 \equiv \begin{bmatrix} 0 \\ 0 \\ -I \end{bmatrix}.$$

This has a determinate solution if and only if M is non-singular. This is evidently a necessary condition for strict concavity of the policy problem, and we shall assume that it holds in the remainder of our discussion.[39] Given this assumption, the unique solution is

$$y_t^\dagger = -M^{-1}\, G\,\mathbf{z}_t. \qquad (3.17)$$

The first-order conditions for the optimal choice of the precommitments $\{\tilde{h}_{t+1}(\xi_{t+1})\}$ are that

$$\beta\mathbf{P}_2\,\mathbf{z}_{t+1} = -\tilde{\varphi}_t \qquad (3.18)$$

in each possible state ξ_{t+1} that can succeed the given state in period t, where $\tilde{\varphi}_t$ is the Lagrange multiplier associated with constraint (3.13); note that the value of $\tilde{\varphi}_t$

[39] We are actually only interested in whether there exists a unique solution for $\tilde{y}_t$. However, condition (3.14) implies that there can be no vector $y^\dagger \neq 0$ such that $My^\dagger = 0$, unless it involves $\tilde{y} \neq 0$. Thus if M is singular, there are necessary multiple solutions for $\tilde{y}_t$ if there are any solutions at all, and not just multiple solutions for the Lagrange multipliers.

depends only on the state in period t. The fact that the left-hand side of (3.18) must be the same in each state ξ_{t+1} implies that

$$P_{22}\,[\tilde{h}_{t+1} - h_t] + P_{23}\,\epsilon_{t+1} = 0$$

in each state. This allows a determinate solution for $\tilde{h}_{t+1}$ if and only if P_{22} is non-singular; this too is evidently a necessary condition for concavity, and is assumed from here on.[40] Under this assumption, (3.18) together with (3.10) implies that

$$\tilde{h}_{t+1} = h_t - P_{22}^{-1}\,P_{23}\,\epsilon_{t+1}. \tag{3.19}$$

We can also solve uniquely for the Lagrange multiplier,

$$\begin{aligned} \tilde{\varphi}_t &= -\beta \mathbf{P}_2\, E_t \mathbf{z}_{t+1} \\ &= -\beta P_{21}\,\tilde{y}_t - \beta P_{22} h_t - \beta [P_{23}\Gamma + P_{24}]\,\xi_t. \end{aligned} \tag{3.20}$$

Equations (3.17) and (3.19) completely describe the optimal dynamics of the variables $\{\tilde{y}_t, \tilde{h}_t\}$, starting from some initial conditions $(\tilde{y}_{t_0-1}, \tilde{h}_{t_0})$, given the evolution of the exogenous states $\{\xi_t\}$. The system consisting of these solutions for $\tilde{y}_t$ and $\tilde{h}_{t+1}(\xi_{t+1})$, together with the law of motion (3.8), can be written in the form

$$\mathbf{z}_{t+1} = \Phi\,\mathbf{z}_t + \Psi\,\epsilon_{t+1}, \tag{3.21}$$

for certain matrices Φ and Ψ. If we partition Φ in the same way as P, it follows from the form of the solutions obtained above that $\Phi_{ij} = 0$ for all $i \geq 2, j \leq 2$. From this (together with our assumption about the eigenvalues of Γ) it follows that all eigenvalues of Φ have modulus less than $\beta^{-1/2}$ if and only if all eigenvalues of

$$\Phi_{11} \;\equiv\; [-I\,0\,0]\;\; M^{-1} G_1 \tag{3.22}$$

have this property. Hence there exists a determinate solution to the first-order conditions for optimal policy, *i.e.*, a unique solution satisfying the bound (3.1), if and only if M and P_{22} are non-singular matrices, and all eigenvalues of Φ_{11} have modulus less than $\beta^{-1/2}$.

[40] If P_{22} is singular, it is obvious that there are multiple solutions for $\tilde{h}_{t+1}(\xi_{t+1})$ consistent with the first-order conditions, but one might wonder if these correspond to multiple state-contingent evolutions $\{\tilde{y}_t\}$. In fact they do, for a single state-contingent evolution $\{\tilde{y}_t\}$ is consistent with only one process $\{\tilde{h}_t\}$, which can be determined from (3.10).

Note that the solution (3.21) involves elements of the matrix P. We can solve for those elements of P in the following way. It follows from the assumed representation (3.11) for the value function that the vector of partial derivatives with respect to $\tilde{y}_{t-1}$ will equal

$$\bar{V}_1^Q = \mathbf{P}_1 \, \mathbf{z}_t.$$

On the other hand, application of the envelope theorem to the problem (3.12) implies that

$$\bar{V}_1^Q = G_1' \, y_t^{\dagger} = -G_1' \, M^{-1} \, G \, \mathbf{z}_t. \tag{3.23}$$

Equating the corresponding coefficients in these two representations, we observe that

$$P_{1j} = -G_1' \, M^{-1} \, G_j$$

for $j = 1, 2, 3, 4$. A similar argument implies that

$$P_{2j} = -G_2' \, M^{-1} \, G_j \tag{3.24}$$

for $j = 1, 2, 3, 4$.

These expressions involve the matrix M, which depends on P_{11}; but the system

$$P_{11} = -G_1' \, M(P_{11})^{-1} \, G_1 \tag{3.25}$$

is a set of n_y^2 equations to solve for the n_y^2 elements of P_{11}.[41] Once we have solved for P_{11}, we know the matrix M, and can solve for the other elements of P. In particular, we can solve for

$$P_{22} = -G_2' \, M^{-1} \, G_2, \tag{3.26}$$

and check whether it is non-singular, as required in (3.19). The other elements of P can be solved for using the same method.[42]

Thus far, we have discussed only the implications of the first-order conditions for the single-period optimization problem. Again, the question arises whether a solution to the first-order conditions corresponds to a maximum of (3.12). The second-order conditions for a finite-dimensional optimization problem are well-known. First, the objective is strictly concave in $\tilde{y}_t$ if and only if the matrix $A_0 + \beta P_{11}$ is such that

$$\tilde{y}' \, [A_0 + \beta P_{11}] \, \tilde{y} < 0$$

[41] Actually, because P_{11} is symmetric, and the system (3.25) has the same symmetry, we need only solve a system of $n(n+1)/2$ equations for $n(n+1)/2$ independent quantities.

[42] Details of the algebra are provided in a note on computational issues available from the authors.

for all $\tilde{y} \neq 0$ such that

$$C_0\,\tilde{y} = 0, \qquad D_0\,\tilde{y} = 0.$$

Using a result of Debreu (1952),[43] we can state algebraic conditions on these matrices that are easily checked. For each r such that $n_F+n_g+1 \leq r \leq n_y$, let M_r be the lower-right square block of M of size $n_F + n_g + r$.[44] Then the concavity condition stated above holds if and only if $\det M_r$ has the same sign as $(-1)^r$, for each $n_F + n_g + 1 \leq r \leq n_y$. Note that in the case that policy is *unidimensional* — meaning that there is a single instrument to set each period, which suffices to determine the evolution of the endogenous variables, so that $n_F + n_g = n_y - 1$ — then this requirement reduces to the single condition that the determinant of M have the same sign as $(-1)^{n_y}$.

Second, in each possible state ξ_{t+1} in the following period, the continuation objective $\bar{V}^Q(\mathbf{z}_{t+1})$ is a concave function of $\tilde{h}_{t+1}(\xi_{t+1})$ if and only if the submatrix P_{22} is *negative definite, i.e.*, such that $\tilde{h}' P_{22} \tilde{h} < 0$ for all $\tilde{h} \neq 0$. This condition is also straightforward to check using the Debreu theorem: the principal minors of P_{22} must have alternating signs.

These two conditions are obviously necessary for strict concavity of the single-period problem, and hence for strict concavity of the infinite-horizon optimal policy problem. In fact, they are also sufficient, yielding the following result.

Proposition 2 *Suppose that the exogenous disturbances have a law of motion of the form (3.8), where Γ is a matrix the eigenvalues of which all have modulus less than $\beta^{-1/2}$, and that the constraints satisfy the rank condition (3.14), where $n_F + n_g < n_y$. Then the LQ policy problem has a determinate solution, given by (3.21), if and only if (i) there exists a solution P_{11} to equations (3.25) such that for each of the minors of the matrix M defined in (3.16), $\det M_r$ has the same sign as $(-1)^r$, for each $n_F + n_g + 1 \leq r \leq n_y$; (ii) the eigenvalues of the matrix Φ_{11} defined in (3.22) all have modulus less than $\beta^{-1/2}$; and (iii) the matrix P_{22} defined in (3.26) is negative definite, i.e., is such that its rth principle minor has the same sign as $(-1)^r$, for each $1 \leq r \leq n_g$.*

[43] See also Theorem 1.E.17 of Takayama (1985).

[44] Given (3.14), we can order the elements of $\tilde{y}_t$ so that the left $(n_F + n_g) \times (n_F + n_g)$ block of the matrix in (3.14) is non-singular, and we assume that this has been done when forming these submatrices.

The proof of this proposition is also given in the Appendix. Note that the conditions stated in the proposition are necessary and sufficient both for the existence of a determinate solution to the first-order conditions, and for the quadratic form $V^Q(\psi)$ to satisfy the strict concavity condition (3.7). In the case that either condition (i) or (iii) is violated, there may exist a determinate solution to the first-order conditions, but it will not represent an optimum, owing to violation of the second-order conditions.

The fact that condition (iii) is needed in addition to conditions (i)–(ii) in order to ensure that we have a concave problem indicates an important respect in which the theory of LQ optimization with forward-looking constraints is not a trivial generalization of the standard theory for backward-looking problems, since conditions (i)–(ii) are sufficient in a backward-looking problem of the kind treated by Magill (1977). [45] It also shows that the second-order conditions for a stochastic problem are more complex than they would be in the case of a deterministic policy problem (again, unlike what is true of purely backward-looking LQ problems). For in a deterministic version of our problem with forward-looking constraints, conditions (i)–(ii) would also be sufficient for concavity, and thus for the solution to the first-order conditions to represent an optimum.

In a deterministic version of the problem — where we not only assume that $\xi_t = 0$ each period, but we restrict our attention to policies under which the evolution of the variables $\{\tilde{y}_t\}$ is purely deterministic (and hence perfectly forecastable), so that we seek to characterize the optimal *perfect foresight equilibrium,* without addressing the question whether this is also optimal among the larger set of possible *rational-expectations equilibria.*[46] — the constraints on possible equilibria are the purely backward-looking constraints (2.23) and

$$D(L)\tilde{y}_t = \tilde{h}_t \tag{3.27}$$

for each $t \geq t_0$, where we specify $\tilde{h}_t = h_{t-1} = 0$ for all $t \geq t_0 + 1$. This is a purely

[45] See Levine *et al.* for a derivation of the second-order conditions for a backward-looking, deterministic LQ problem, using what is essentially a discrete-time version of the approach of Magill. In some cases, conditions (i)–(ii) are both necessary and sufficient for concavity, even in the presence of forward-looking constraints. The problem treated in Benigno and Woodford (2005a) is an example of this kind. Note that in that paper an alternative, frequency-domain characterization of the conditions for concavity is used, that is discussed more generally in Benigno and Woodford (2006b).

[46] Additional equilibria can be attained, by randomization of policy, even in the case that there are no exogenous random disturbances. This may or may not allow an increase in welfare relative to the optimal deterministic policy.

backward-looking problem, so that the standard second-order conditions apply. And it should be obvious that, as there is no longer a choice of $\tilde{h}_{t+1}(\xi_{t+1})$ to be made each period, our argument above for the necessity of condition (iii) would not apply.

But conditions (i)–(ii) are not generally a sufficient condition to guarantee that (3.7) is satisfied, in the presence of forward-looking constraints (2.24), if policy randomization is allowed.[47] Because constraints (2.24) need hold only in expected value, random policy may be able to vary the paths of the endogenous variables (in some states of the world) in directions that would not be possible in the corresponding deterministic problem, and this makes the algebraic conditions required for (3.7) to hold more stringent. Specifically, the value function for the continuation problem must be a strictly concave function of the state-contingent pre-commitment $\tilde{h}_{t+1}$ made for the following period, or it is possible to randomize $\tilde{h}_{t+1}$ (requiring a corresponding randomization of subsequent policy) without changing the fact that constraint (2.24) is satisfied in period t. Hence condition (iii) is necessary in the stochastic case.[48] It can also easily be shown that condition (iii) is not implied in general by conditions (i)–(ii).

A simple example may clarify this point. Suppose that y_t has two elements, and that the only constraint on what policy can achieve is a single, forward-looking constraint

$$E_t[\delta \tilde{y}_{1,t} - \tilde{y}_{1,t+1}] = 0 \tag{3.28}$$

[47] Our remarks here apply even in the case that the "fundamental" disturbances $\{\xi_t\}$ are purely deterministic; what matters is whether policy may be contingent upon random events. As is discussed further in Benigno and Woodford (2005a), when the second-order conditions fail to hold, policy randomization can be welfare-improving, even when the random variations in policy are unrelated to any variation in fundamentals.

[48] Levine *et al.* (2007) provide a different argument for a condition similar to our condition (iii) as a necessary condition for optimality in a model with a forward-looking constraint, which does not require a consideration of stochastic policy. They consider Ramsey-optimal policy rather than optimality from a timeless perspective; that is, they assume no initial precommitment (2.25). In this case, the deterministic optimal policy problem is like the one considered above, except that (3.27) need hold only in periods $t \geq t_0 + 1$; the optimal policy is then the same as in the backward-looking problem just discussed, except that instead of taking $\tilde{h}_{t_0}$ as given, one is free to choose $\tilde{h}_{t_0}$ so as to maximize (2.22). This latter problem has a solution only if the value function $\bar{V}_{t_0}^{Q}$ is bounded above, for a given vector $\tilde{y}_{t_0-1}$, and this is true in general only if it is a strictly concave function of $\tilde{h}_{t_0}$. The validity of this argument, however, depends on considering an exact LQ problem, rather than an LQ local approximation to a problem that may have different global behavior.

for all $t \geq t_0$, where $\delta < \beta^{-1/2}$. (The path of $\{\tilde{y}_{2,t}\}$ can be freely chosen, subject to the bound (3.1).) An initial pre-commitment specifies the value that $\tilde{y}_{1,t_0}$ must have. In the corresponding deterministic problem, constraint (3.28) implies that one must have

$$\tilde{y}_{1,t+1} = \delta \tilde{y}_{1,t}$$

for each $t \geq t_0$, and this, together with the pre-commitment, uniquely determines the entire path of the sequence $\{\tilde{y}_{1,t}\}$ that must be brought about by deterministic policy. Hence the second-order condition for the deterministic problem requires only that the objective be a concave function of the path of $\{\tilde{y}_{2,t}\}$. But if random policies are considered, it is also possible for $\{\tilde{y}_{1,t}\}$ to evolve in accordance with any law of motion

$$\tilde{y}_{1,t+1} = \delta \tilde{y}_{1,t} + \epsilon_{t+1},$$

where $\{\epsilon_t\}$ is any martingale difference sequence with a suitable bound on its asymptotic variance; in this simple example, the set of possible evolutions $\{\tilde{y}_{1,t}\}$ is independent of the evolution chosen for $\{\tilde{y}_{2,t}\}$. Whether randomization of the path of $\{\tilde{y}_{1,t}\}$ can increase the value of the policy objective obviously depends on terms in the objective involving the path of $\{\tilde{y}_{1,t}\}$ (including cross terms), and not just the terms involving the path of $\{\tilde{y}_{2,t}\}$. Hence the conditions required for a concave optimization problem are more stringent in this case.[49]

4 Welfare Evaluation of Alternative Policy Rules

We have argued that another advantage of our approach is that it can be used not only to derive a linear approximation to a fully optimal policy commitment, but also to compute approximate welfare comparisons between alternative rules (neither of which may be fully optimal), that will correctly rank these rules in the case that random disturbances are small enough. Because empirically realistic models are inevitably fairly complex, a fully optimal policy rule is likely to be too complex to represent a realistic policy proposal; hence comparisons among alternative simple (though suboptimal) rules are of considerable practical interest. Here we discuss how this can be done.

[49] In the Appendix, we illustrate the application of the conditions in Proposition 2 to this example.

We do not propose to simply evaluate (a local approximation to) expected discounted utility V_{t_0} under a candidate policy rule, because the optimal policy locally characterized above (*i.e.*, optimal policy "from a timeless perspective") does not maximize this objective; hence ranking rules according to this criterion would lead to the embarrassing conclusion that there exist policies better than the optimal policy. (We could, of course, define "optimal policy" as the policy that maximizes V_{t_0}; but this would result in a time-inconsistent policy recommendation, as noted earlier.) Thus we wish to use a criterion that ranks rules according to how close they come to solving the recursive policy problem defined in section 2.1, rather than how close they come to maximizing V_{t_0}.

Of course, if we restrict our attention to policies that necessarily satisfy the initial pre-commitment (2.4), there is no problem; our optimal rule will be the one that maximizes V_{t_0}, or (in the case of small enough shocks) the one that maximizes $V^Q_{t_0}$. But *simple* policy rules are unlikely to precisely satisfy (2.4); thus in order to be able to select the best rule from some simple class, we need an alternative criterion, one that is defined for *all* policies that are close enough to being optimal, in a sense that is to be defined. At the same time, we wish it to be a criterion the maximization of which implies that one has solved the constrained optimization problem defined in section 2.1.

4.1 A Lagrangian Approach

Our Lagrangian characterization of optimal policy suggests such a criterion. The timelessly optimal policy from date t_0 onward — that is, the policy that maximizes V_{t_0} subject to the initial constraint (2.4) in addition to the feasibility constraints (2.2)–(2.3) — is also the policy that maximizes the Lagrangian

$$V^{mod}_{t_0} \equiv V_{t_0} + \beta^{-1}\varphi'_{t_0-1} g(y_{t_0-1}, \xi_{t_0-1}; y_{t_0}), \tag{4.1}$$

where φ_{t_0-1} is the vector of Lagrange multipliers associated with the initial constraint (2.4). This is a function that coincides (up to a constant) with the objective V_{t_0} in the case of policies satisfying the constraint (2.4), but that is defined more generally, and that is maximized over the broader class of feasible policies by the timelessly optimal policy. Hence an appropriate criterion to use in ranking alternative policies is the value of $V^{mod}_{t_0}$ associated with each one. This criterion penalizes policies that fail to satisfy the initial pre-commitment (2.4), by exactly the amount by which a previously *anticipated* deviation of that kind would have reduced the expected utility of the representative household.

In the case of any policy that satisfies the feasibility constraints (2.2)–(2.3) for all $t \geq t_0$, we observe that

$$\begin{aligned} V_{t_0}^{mod} &= \bar{\mathcal{L}}_{t_0} + \beta^{-1}\tilde{\varphi}'_{t_0-1} g(y_{t_0-1}, \xi_{t_0-1}; y_{t_0}) \\ &= V_{t_0}^{Q} + \beta^{-1}\tilde{\varphi}'_{t_0-1}\tilde{g}(y_{t_0-1}, \xi_{t_0-1}; y_{t_0}) + \text{t.i.p.} + \mathcal{O}(||\xi||^3) \\ &= V_{t_0}^{Q} + \beta^{-1}\tilde{\varphi}'_{t_0-1} D_{\hat{y}} g \cdot \tilde{y}_{t_0} + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \end{aligned}$$

This suggests that in the case of small enough shocks, the ranking of alternative policies in terms of $V_{t_0}^{mod}$ will correspond to the ranking in terms of the welfare measure

$$W_{t_0} \equiv V_{t_0}^{Q} + \beta^{-1}\tilde{\varphi}'_{t_0-1} D_{\hat{y}} g \cdot \tilde{y}_{t_0}. \tag{4.2}$$

Note that in this derivation we have assumed that $\tilde{y}_t = \mathcal{O}(||\xi||)$. This will be true in the equilibrium associated with any (sufficiently differentiable) policy rule that is *consistent with the optimal steady state* in the absence of random disturbances. We shall restrict attention to policy rules of this kind. Note that while this is an important restriction, it does not preclude consideration of extremely simple rules; and it is a property of the simple rules of greatest interest, *i.e.*, those that come closest to being optimal among rules of that degree of complexity.

In expression (4.1), and hence in (4.2), φ_{t_0-1} is the Lagrange multiplier associated with constraint (2.4) under the optimal policy. However, in order to evaluate W_{t_0} to second-order accuracy, it suffices to have a first-order approximation to this multiplier. Such an approximation is given by the multiplier $\tilde{\varphi}_{t_0-1}$ associated with the constraint (2.25) of the LQ problem. Thus we need only solve the LQ problem, as discussed in the previous section — obtaining a value for $\tilde{\varphi}_{t_0-1}$ along with our solution for the optimal evolution $\{y_t\}$ — in order to determine the value of W_{t_0}.

Moreover, we observe that in the characterization given in the previous section of the solution to the LQ problem, $\tilde{\varphi}_{t_0-1} = \mathcal{O}(||\xi||)$.[50] Thus a solution for the equilibrium evolution $\{\tilde{y}_t\}$ under a given policy that is accurate to first order suffices to evaluate the second term in (4.2) to second-order accuracy. Hence W_{t_0} inherits this

[50] This follows from solution (3.17) for the Lagrange multiplier associated with the initial pre-commitment.

property of $V_{t_0}^Q$, and it suffices to compute a linear approximation to the equilibrium dynamics $\{\tilde{y}_t\}$ under each candidate policy rule in order to evaluate W_{t_0} to second-order accuracy. We can therefore obtain an approximation solution for $\{\tilde{y}_t\}$ under a given policy by solving the linearized structural equations (2.23)–(2.24), together with the policy rule, and use this solution in evaluating W_{t_0}. In this way welfare comparisons among alternative policies are possible, to second-order accuracy, using linear approximations to the model structural relations and a quadratic welfare objective.

Moreover, we can evaluate W_{t_0} to second-order accuracy using only a linear approximation to the policy rule. This has important computational advantages. For example, if we wish to find the optimal policy rule from among the family of simple rules of the form $i_t = \phi(y_t)$, where i_t is a policy instrument, and we are content to evaluate $V_{t_0}^{mod}$ to second-order accuracy, then it suffices to search over the family of linear policy rules[51]

$$\tilde{\imath}_t = f'\tilde{y}_t,$$

parameterized by the vector of coefficients f. There are no possible second-order (or larger) welfare gains resulting from nonlinearities in the policy rule.

It is important to note that these conclusions obtain *only* because we evaluate welfare taking into account the welfare losses that would result from a violation of the initial pre-commitment if it were to have been anticipated. Some would prefer to evaluate alternative simple policy rules by computing the expected value of V_{t_0} (rather than $V_{t_0}^{mod}$) associated with each rule (*e.g.*, Schmitt-Grohé and Uribe, 2004c). As noted above, this alternative criterion is one under which the optimal rule from a timeless perspective can be dominated by other rules, a point stressed by Blake (2001) and Jensen and McCallum (2002), among others. The alternative criterion is also one that cannot be evaluated to second-order accuracy using only a first-order solution for the equilibrium evolution under a given policy. For a general feasible policy — consistent with the optimal steady state, but not necessarily consistent

[51]Here we restrict attention to rules that are consistent with the optimal steady state, so that the intercept term is zero when the rule is expressed in terms of deviations from steady-state values. Note that a rule without this property will result in lower welfare, in the case of any small enough disturbances.

with the initial pre-commitment (2.4) — we can show that[52]

$$V_{t_0} = V_{t_0}^Q - \beta^{-1}\bar{\varphi}' D_{\hat{y}} g \cdot \tilde{y}_{t_0} + \text{t.i.p.} + \mathcal{O}(||\xi||^3). \tag{4.3}$$

The first term on the right-hand side of this expression is purely quadratic (has zero linear terms), but this is not true of the second term, if the initial pre-commitment is binding under the optimal policy. Evaluation of the second term to second-order accuracy requires a second-order approximation to the evolution $\{y_t\}$ under the policy of interest; there is thus no alternative to the use of higher-order perturbation solution methods as illustrated by Schmitt-Grohé and Uribe, and nonlinear terms in the policy rule generally matter for welfare.[53]

In expression (4.2), the value of the multiplier $\tilde{\varphi}_{t_0-1}$ depends on the economy's initial state and on the value of the initial pre-commitment $\bar{g}_{t_0}$. However, we wish to be able to rank alternative rules for an economy in which no such commitment may exist prior to the adoption of the policy rule. We can avoid having to make reference to any historically given pre-commitment by assuming a self-consistent constraint of the form (2.5).

If we define a new extended state vector

$$\hat{\mathbf{z}}_t \equiv \begin{bmatrix} \tilde{y}_{t-1} \\ \hat{h}(\xi_t, \xi_{t-1}) \\ \xi_t \\ \xi_{t-1} \end{bmatrix},$$

where[54]

$$\hat{h}(\xi_t, \xi_{t-1}) \equiv h_{t-1} - P_{22}^{-1} P_{23}(\xi_t - \Gamma \xi_{t-1}),$$

then it follows from (3.19) that under the solution to the recursive policy problem, $\mathbf{z}_t = \hat{\mathbf{z}}_t$ for each $t \geq t_0 + 1$. (However, $\hat{\mathbf{z}}_t$, unlike $\mathbf{z}_t$, is a function solely of $\tilde{y}_{t-1}$ and the

[52] Here we use the more general form of (2.20) mentioned in footnote 31.

[53] Thus welfare comparisons of the kind proposed by Blake (2001), Jensen and McCallum (2002), or Sauer (2006), in which the implications of a policy rule are computed using the structural equations of a canonical log-linearized New Keynesian model and welfare is evaluated using the canonical quadratic loss function, cannot be justified as representing a quadratic approximation to the expected utility of the representative household in a micro-founded model with Calvo price adjustment. The welfare criterion proposed here can instead be computed using the usual log-linearized structural equations, as is discussed further in Benigno and Woodford (2005a, sec. 5).

[54] Here it should be recalled that h_{t-1} is a linear function of ξ_{t-1}, defined in (2.29).

history of the exogenous disturbances.) Hence

$$\tilde{h}_{t_0} = \hat{h}(\xi_t, \xi_{t-1}) \tag{4.4}$$

is a self-consistent constraint of the form (2.5).

If we assume an initial pre-commitment specified in this way, it also follows from (3.17) that

$$\tilde{\psi}_t = [0 \;\; 0 \;\; -I] \; M^{-1} G \; \hat{\mathbf{z}}_t \tag{4.5}$$

is the Lagrange multiplier associated with the pre-commitment each period in the recursive problem. Moreover, because the only constraint on the way in which $\tilde{h}_{t+1}(\xi_{t+1})$ can be chosen for the following period is given by the expected-value constraint (3.13), the first-order conditions for optimal policy imply that $\tilde{\psi}_t = E_{t-1}\tilde{\psi}_t$ for each $t \geq t_0 + 1$,[55] and hence that

$$\begin{aligned} \tilde{\psi}_t &= [0 \;\; 0 \;\; -I] \; M^{-1} G \; E_{t-1}\hat{\mathbf{z}}_t \\ &= \tilde{\psi}(\tilde{y}_{t-1}, \xi_{t-1}) \equiv [0 \;\; 0 \;\; -I] \; M^{-1} G \begin{bmatrix} \tilde{y}_{t-1} \\ h_{t-1} \\ \Gamma\xi_{t-1} \\ \xi_{t-1} \end{bmatrix}. \end{aligned}$$

Consistency of this result with (4.5) implies that the right-hand-side of (4.5) must be equivalent to $\tilde{\psi}(\tilde{y}_{t-1}, \xi_{t-1})$; that is, that the coefficients multiplying $\tilde{y}_{t-1}$, ξ_t, and ξ_{t-1} must be the same in both expressions. But since (4.5) must hold at $t = t_0$ as well, in the case of an initial pre-commitment (4.4), and not only for $t \geq t_0 + 1$, it follows that under such a pre-commitment,

$$\tilde{\psi}_t = \tilde{\psi}(\tilde{y}_{t-1}, \xi_{t-1})$$

for all $t \geq t_0$. In the case that $t = t_0$, the multiplier $\tilde{\psi}_{t_0}$ associated with the initial pre-commitment is the one that is denoted $\beta^{-1}\tilde{\varphi}_{t_0-1}$ in (2.30) and in (4.2). Thus we can write

$$\tilde{\varphi}_{t_0-1} = \varphi^*(\mathbf{y_{t_0-1}}) \equiv \beta\tilde{\psi}(\tilde{y}_{t_0-1}, \xi_{t_0-1}). \tag{4.6}$$

[55]In fact, one can show that $\tilde{\psi}_t = \beta^{-1}\tilde{\varphi}_{t-1}$ for each $t \geq t_0 + 1$. This follows from differentiation of the value function $V^Q(\mathbf{z}_{t+1})$ with respect to $\tilde{h}_{t+1}$ using the envelope theorem, and comparison of the result with (3.18).

Then we can write[56]

$$W_{t_0} = W(\tilde{y};\ \xi_{t_0}, \mathbf{y_{t_0-1}}) \equiv V^Q_{t_0} + \beta^{-1}\varphi^*(\mathbf{y_{t_0-1}})' D_{\tilde{y}} g \cdot \tilde{y}_{t_0}. \tag{4.7}$$

This gives us an expression for our welfare measure purely in terms of the history and subsequent evolution of the extended state vector.[57]

4.2 A Time-Invariant Criterion for Ranking Alternative Rules

Let us suppose that we are interested in evaluating a policy rule r that implies an equilibrium evolution of the endogenous variables of the form[58]

$$y_t = \phi_r(\xi_t, \mathbf{y_{t-1}}).$$

This (together with the law of motion for the exogenous disturbances) then implies a law of motion for the complete extended state vector

$$\mathbf{y}_t = \psi_r(\xi_t, \mathbf{y_{t-1}}). \tag{4.8}$$

Using this law of motion, we can evaluate (4.7), obtaining

$$W_{t_0} = W_r(\xi_{t_0}, \mathbf{y_{t_0-1}}).$$

We can do this for any rule r of the assumed type, and hence we can define an optimization problem

$$\max_{r \in \mathcal{R}} W_r(\xi_{t_0}, \mathbf{y_{t_0-1}}) \tag{4.9}$$

in order to determine the optimal rule from among the members of some family of rules $\mathcal{R}$.

[56]In writing the function $W(\cdot)$, and others that follow, we suppress the argument ξ, as the evolution of the exogenous disturbances is the same in the case of each of the alternative policies under consideration.

[57]Note that it is possible to solve for the initial Lagrange multipliers $\varphi^*(\mathbf{y_{t_0-1}})$ using only the values of $\tilde{y}_{t_0-1}$ and of ξ_{t_0-1}. It is not necessary to simulate the optimal equilibrium dynamics over a lengthy "estimation period" prior to the date t_0 at which the new policy is to commence, as proposed by Juillard and Pelgrin (2006).

[58]This assumption that y_t depends only on the state variables indicated is without loss of generality, as we can extend the vector $\mathbf{y_t}$ if necessary in order for this to be so.

However, the solution to problem (4.9) may well depend on the initial conditions $\mathbf{y_{t_0-1}}$ and ξ_{t_0} for which W_{t_0} is evaluated.[59] This leads to the possibility of an unappealing degree of arbitrariness of the choice that would be recommended from within some family of simple rules, as well as time inconsistency of the policy recommendation: a rule chosen at date t_0 on the ground that it solves problem (4.9) need not be found to also solve the corresponding problem at some later date, though the calculation at date t_0 assumes that rule r is to be followed forever. One way of avoiding this might be to assume that one should choose the rule that would be judged best in the case of initial conditions consistent with the optimal steady state, whether the economy's actual initial state is that one or not;[60] that is, one would choose the rule that solves the problem

$$\max_{r \in \mathcal{R}} W_r(0, \bar{\mathbf{y}}).$$

This choice would not be time-inconsistent, but the choice is still an arbitrary one. In particular, the decision to evaluate W_r assuming initial conditions consistent with the steady state — when in fact the state of the economy will fluctuate on both sides of the steady-state position — favors rules r for which W_r is a less concave function of the initial condition.

The criterion that we find most appealing is accordingly to integrate over a distribution of possible initial conditions, rather than evaluating W_r at the economy's actual state at the time of the choice, or at any other single state (such as the optimal steady state). Suppose that in the case of the optimal policy rule r^*, the law of motion (4.8) implies that the evolution of the extended state vector $\{\mathbf{y_t}\}$ is *stationary*.[61] In this case, there exists a well-defined invariant (or unconditional) probability distribution μ for the possible values of $\mathbf{y_t}$ under the optimal policy.[62] Then we can define the optimal policy rule within some class of simple rules $\mathcal{R}$ as the one that solves the problem

$$\max_{r \in \mathcal{R}} E_\mu[\bar{W}_r(\mathbf{y_t})], \tag{4.10}$$

[59]This is not a problem if the family of rules $\mathcal{R}$ includes a fully optimal rule r^*, since the same rule r^* solves the problem (2.7) for all possible values of the initial conditions. But the result can easily depend on the initial conditions if we restrict attention to a family of suboptimal rules.

[60]This approach is proposed by Schmitt-Grohé and Uribe (2004c), though they use V_{t_0} rather than $V_{t_0}^{mod}$ as the criterion to be maximized.

[61]Benigno and Woodford (2005a) provide an example of an optimal monetary stabilization policy problem in which this is case.

[62]We discuss the computation of the relevant properties of this invariant measure in the Appendix.

where[63]

$$\bar{W}_r(\mathbf{y_t}) \equiv E_t W_r(\xi_{t+1}, \mathbf{y_t}). \tag{4.11}$$

Because of the linearity of our approximate characterization of optimal policy, the calculations required in order to evaluate $E_\mu[W_r]$ to second-order accuracy are straightforward; these are illustrated in Benigno and Woodford (2005a, sec. 5).

The most important case in which the method just described cannot be applied is when some of the elements of $\{\mathbf{y_t}\}$ possess unit roots, though all elements are at least difference-stationary (and some of the non-stationary elements may be cointegrated). Note that it is possible for even the equilibrium under optimal policy to have this property, consistent with our assumption of the bound (3.1).[64] There is a question in such a case whether our local approximation to the problem should remain an accurate approximation, but this is not a problem in the case that random disturbances occur in only a *finite* number of periods, so LQ problems of this kind may be of practical interest.

Let us suppose that those elements which possess unit roots are pure random walks (*i.e.*, with zero drift).[65] We can in such a case decompose the extended state vector as

$$\mathbf{y_t} = \mathbf{y_t}^{tr} + \mathbf{y_t}^{cyc},$$

where

$$\mathbf{y_t}^{tr} \equiv \lim_{T\to\infty} E_t \mathbf{y_T}$$

is the Beveridge-Nelson (1981) "trend" component, and the "cyclical" component $\mathbf{y_t}^{cyc}$ will still be a stationary process. Moreover, the evolution of the cyclical component as a function of the exogenous disturbances under the optimal policy will be independent of the assumed initial value of the trend component (though not of the

[63] Recall that we assume that the exogenous disturbance process $\{\xi_t\}$ is Markovian, and that ξ_t is included among the elements of $\mathbf{y_t}$. Hence $\mathbf{y_t}$ contains all relevant elements of the period t information set for the calculation of this conditional expectation.

[64] Benigno and Woodford (2003) provide an example of an optimal stabilization policy problem in which the LQ approximate problem has this property. In this example, the unit root is associated with the dynamics of the level of real public debt, which display a unit root under optimal policy for the same reason as in the classic analysis of optimal tax smoothing by Barro (1979) and Sargent (1987, chap. XV).

[65] We may suppose that any deterministic trend under optimal policy has been eliminated by local expansion around a deterministic solution with constant trend growth, so that there is zero trend in the state variables $\{\tilde{y}_t\}$ expressed as deviations from that deterministic solution.

initial value of the cyclical component). It follows that we can define an invariant distribution μ for the possible values of $\mathbf{y_t}^{cyc}$ under the optimal policy, that is independent of the assumed value for the trend component. Then for any assumed initial value for the trend component $\mathbf{y_{t_0-1}}^{tr}$, we can define the optimal policy rule within the class $\mathcal{R}$ as the one that solves the problem

$$\max_{r\in\mathcal{R}} \Omega_r(\mathbf{y_{t_0-1}}^{tr}) \equiv E_\mu[\bar{W}_r(\mathbf{y_{t_0-1}})], \tag{4.12}$$

a generalization of (4.10).[66]

It might seem in this case that our criterion is again dependent on initial conditions, just as with the criterion (4.9) proposed first. The following result shows that this is not the case.

Lemma 3 *Suppose that under optimal policy, the extended state vector* $\mathbf{y_t}$ *consists entirely of components that are either (i) stationary, or (ii) pure random walks. Suppose also that the class of policy rules* $\mathcal{R}$ *is such that each rule in the class implies convergence to the same long-run values of the state variables as under optimal policy, in the absence of stochastic disturbances, so that the initial value of the trend component* $\mathbf{y}_{t_0-1}^{tr}$ *is the same regardless of the rule* r *that is considered. Then for any rule* $r \in \mathcal{R}$*, the objective* $\Omega_r(\mathbf{y_{t_0-1}}^{tr})$ *defined in (4.12) can be decomposed into two parts,*

$$\Omega_r(\mathbf{y_{t_0-1}}^{tr}) = \Omega^1(\mathbf{y_{t_0-1}}^{tr}) + \Omega^2_r, \tag{4.13}$$

where the first component is the same for all rules in this class, while the second component is independent of the initial condition $\mathbf{y_{t_0-1}}^{tr}$*.*

Hence the criterion (4.12) establishes the same ranking of alternative rules, regardless of the initial condition. The proof of this result is given in the Appendix.

5 Applications

The approach expounded here has already proven fruitful in a number of applications to problems of optimal monetary and fiscal policy. Benigno and Woodford (2005a)

[66] In the case that all elements of $\mathbf{y_t}$ are stationary, $\mathbf{y_t}^{tr}$ is simply a constant, and all variations in $\mathbf{y_t}$ correspond to variations in $\mathbf{y_t}^{cyc}$. In this case, (4.12) is equivalent to the previous criterion (4.10).

use this method to derive an LQ approximation to the problem of optimal monetary stabilization policy in a DSGE model with monopolistic competition, Calvo-style staggered price-setting, and a variety of exogenous disturbances to preferences, technology, and fiscal policy. Unlike the method used by Rotemberg and Woodford (1997) and Woodford (2002), the present method is applicable even in the case of (possibly substantial) distortions even in the absence of shocks, owing to market power or distorting taxes. The quadratic stabilization objective obtained is of the form

$$-\frac{1}{2}E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\left[q_\pi\pi_t^2+q_y(\hat{Y}_t-\hat{Y}_t^*)^2\right], \tag{5.1}$$

where π_t is the inflation rate between periods $t-1$ and t, $\hat{Y}_t$ is the log deviation of aggregate real output from trend, $\hat{Y}_t^*$ is a target level of output that depends purely on the exogenous real disturbances, $0<\beta<1$ is the representative household's discount factor, and the weights q_π, q_y are functions of model parameters (both positive if steady-state distortions are not severe). The single linear constraint corresponds to the familiar "new Keynesian Phillips curve,"

$$\pi_t=\kappa[\hat{Y}_t-\hat{Y}_t^*]+\beta E_t\pi_{t+1}+u_t, \tag{5.2}$$

where $\kappa>0$ is a function of model parameters and the "cost-push" term u_t is a linear function of the various exogenous real disturbances.

The resulting LQ problem is of a form that has already been extensively studied in the literature on optimal monetary stabilization policy,[67] and so the ways in which the parameterization of the objective and constraint shape the character of optimal policy is well understood once the problem is stated in this form. The analysis in Benigno and Woodford (2005a), however, explains the microeconomic determinants of these factors. For example, it provides an interpretation of the "cost-push" disturbances that play a crucial role in familiar discussions of the tradeoffs between inflation and output stabilization, and shows that the cost-push effects of most types of shocks are larger the more distorted is the economy's steady state; and it explains the relative weight that should be assigned to the output-gap stabilization objective, showing that this need not be positive in the case of a sufficiently distorted economy. (Indeed, if distortions are severe, the quadratic objective can fail to be concave, so that a small

[67] See, *e.g.*, , Clarida *et al.* (1999) and Woodford (2003, chap. 7).

amount of policy randomization can be welfare-improving.) Benigno and Woodford (2005b) extend the analysis to the case in which both wages and prices are sticky, obtaining a generalization of (5.1) in which a third quadratic loss term appears, proportional to squared deviations of nominal wage inflation from zero. This shows that the analysis by Erceg *et al.* (2000) of the tradeoff between stabilization of wage inflation and price inflation applies also to economies with distorted steady states, though the policy tradeoffs are complicated by the presence of cost-push terms that do not appear in those authors' analysis of the case of an undistorted steady state.

An important limitation of the LQ method of Rotemberg and Woodford (1997), that restricts attention to cases in which the utility gradient is zero in the steady state, is that it cannot easily be applied to analyses of optimal policy for open economies; for in an open economy, domestic production and consumption cannot be equated, and the marginal utility associated with a change in either individually will inevitably be non-zero in any reasonable case. The method proposed here instead allows LQ analyses of optimal policy also in the case of open economies.

Benigno and Benigno (2006) analyze policy coordination between two national monetary authorities which each seek to maximize the welfare of their own country's representative household, and show that it is possible to locally characterize each authority's aims by a quadratic stabilization objective. Previous LQ analyses of policy coordination have often assumed an objective of the form (5.1) for each national authority, but with the nation's own inflation rate and output being the arguments in each case. Benigno and Benigno instead show that household utility maximization would correspond to a quadratic objective for each authority with terms penalizing fluctuations in *both* domestic and foreign inflation (but with different weights on the two terms for the distinct national authorities), and similarly with terms penalizing fluctuations in both domestic and foreign output (again with different weights in the case of the two authorities). They also show that each authority's stabilization objective should contain a term penalizing departures of the terms of trade from a "target" level (that depends on exogenous disturbances), and show how both the weight placed on this additional objective and the nature of variation in the terms of trade "target" depend on underlying micro-foundations. De Paoli (2004) similarly shows how the analysis of Benigno and Woodford (2005a) can be extended to a small open economy, requiring the addition of a terms-of-trade (or real-exchange-rate) stabilization objective to the two terms shown in (5.1).

Another advantage of the fact that the present method applies to economies with a distorted steady state is that it can be used to analyze optimal tax smoothing when only distorting taxes are available as sources of government revenue, after the fashion of Barro (1979) and Sargent (1987, chap. XV), and allows the theory of tax smoothing to be integrated with the theory of monetary stabilization policy. Benigno and Woodford (2003) extend the analysis of Benigno and Woodford (2005a) to the case of an economy with only distorting taxes, and show that the problem of choosing jointly optimal monetary and fiscal policies can also be treated within an LQ framework that nests standard analyses of tax smoothing (with flexible prices, so that real effects of monetary policy are ignored) and of monetary policy (with lump-sum taxes, so that fiscal effects of monetary policy can be ignored) as special cases. Notably, they find that allowing for tax distortions introduces no additional stabilization goals into the quadratic objective (5.1). Instead, the benefits of tax smoothing are represented by the penalty on squared departures of equilibrium output from its "target" level; tax variations can increase the average size of this term, because of the effects of the level of distorting taxes on equilibrium output (which occur due to a "cost-push" effect of tax rates in the generalized version of the constraint (5.2)). Benigno and De Paoli (2005) extend this analysis to treat optimal monetary and fiscal policy in a small open economy, while Ferrero (2005) analyzes optimal monetary and fiscal policy in a monetary union with separate national fiscal authorities.

All of the analyses just mentioned involve fairly simple DSGE models, in which it is possible to derive the coefficients of the LQ approximate policy problem by hand. In the case of larger (and more realistic) models of the kind that are now being estimated for use in practical policy analysis, such calculations are likely to be tedious. Nonetheless, it is an advantage of our method that it is straightforward to apply it even to fairly complex models and fairly general specifications of disturbances. Altissimo *et al.* (2005) describe computer code that executes the calculations explained above, for a general nonlinear problem with an arbitrary number of state variables, and demonstrate its application to two important extensions of the work described above, an analysis of optimal monetary policy in the presence of non-trivial frictions of the kind that result in a transactions demand for money, and an analysis of optimal monetary policy for the empirical model of Smets and Wouters. Curdia (2007) illustrates the application of the methods proposed here to another fairly complex model, namely, a model of "sudden stops" in a small emerging-market economy; in particular, the method explained in section 4 is used to evaluate alternative simple policy rules for such a setting. We believe that it should similarly be practical to apply these methods to a wide variety of other models of interest to policy institutions.

A Appendix: Proofs and Derivations

A.1 Proposition 1

Recall that $\mathcal{H}$ is the Hilbert space of (real-valued) stochastic processes $\{\tilde{y}_t\}$ such that

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \tilde{y}_t' \tilde{y}_t < \infty, \tag{A.1}$$

and $\mathcal{H}_1 \subset \mathcal{H}$ is the subspace of sequences $\hat{y} \in \mathcal{H}$ that satisfy the additional constraints

$$C(L)\hat{y}_t = 0 \tag{A.2}$$

$$E_t D(L)\hat{y}_{t+1} = 0 \tag{A.3}$$

for each date $t \geq t_0$, along with the initial commitments

$$D(L)\hat{y}_{t_0} = 0, \tag{A.4}$$

where we define $\hat{y}_{t_0-1} \equiv 0$ in writing (A.2) for period $t = t_0$ and in writing (A.4).

Proposition 1 *For $\{\tilde{y}_t\} \in \mathcal{H}$ to maximize the quadratic form (2.22), subject to the constraints (2.23) – (2.25) given initial conditions $\tilde{y}_{t_0-1}$ and $\bar{g}_{t_0}$, it is necessary and sufficient that (i) there exist Lagrange multiplier processes[68] $\tilde{\varphi}, \tilde{\lambda} \in \mathcal{H}$ such that the processes $\{\tilde{y}_t, \tilde{\varphi}_t, \tilde{\lambda}_t\}$ satisfy (3.2) for each $t \geq t_0$; and (ii)*

$$V^Q(\hat{y}) \equiv V^Q_{t_0}(\hat{y};0) = \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} [\hat{y}_t' A(L) \hat{y}_t] \leq 0 \tag{A.5}$$

for all processes $\hat{y} \in \mathcal{H}_1$, where in evaluating (A.5) we define $\hat{y}_{t_0-1} \equiv 0$. A process $\{\tilde{y}_t\}$ with these properties is furthermore uniquely *optimal if and only if*

$$V^Q(\hat{y}) < 0 \tag{A.6}$$

for all processes $\hat{y} \in \mathcal{H}_1$ that are non-zero almost surely.

PROOF: We have already remarked on the necessity of the first-order conditions (i). To prove the necessity of the second-order condition (ii) as well, let $\{\tilde{y}_t\} \in \mathcal{H}$, and consider the the perturbed process

$$y_t = \tilde{y}_t + \hat{y}_t \tag{A.7}$$

[68] Note that $\tilde{\varphi}_t$ is also assumed to be defined for $t = t_0 - 1$.

for all $t \geq t_0 - 1$, where $\{\hat{y}_t\}$ belongs to $\mathcal{H}_1$ and we define $\hat{y}_{t_0-1} \equiv 0$. This construction guarantees that if the process $\{\tilde{y}_t\}$ satisfies the constraints (2.23) – (2.25), so does the process $\{y_t\}$.

We note that

$$\begin{aligned} V^Q_{t_0}(y;\xi) &= V^Q_{t_0}(\tilde{y};\xi) + \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} [\hat{y}'_t A(L) \tilde{y}_t + \tilde{y}'_t A(L) \hat{y}_t + 2\hat{y}'_t B(L) \xi_{t+1}] \\ & \quad + \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} [\hat{y}'_t A(L) \hat{y}_t]. \end{aligned}$$

The second term on the right-hand side is furthermore equal to

$$\begin{aligned} & \frac{1}{2} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \hat{y}'_t \cdot \left\{ [A(L) + A'(\beta L^{-1})] \tilde{y}_t + 2B(L)\xi_{t+1} \right\} \\ = & -E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \hat{y}'_t \cdot \left\{ C'(\beta L^{-1}) \tilde{\lambda}_t + \beta^{-1} D'(\beta L^{-1}) \tilde{\varphi}_{t-1} \right\} \\ = & -E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \tilde{\lambda}'_t C(L) \hat{y}_t + \beta^{-1} \tilde{\varphi}'_{t-1} D(L) \hat{y}_t \right\}, \end{aligned}$$

where we use the first-order conditions (3.2) to establish the first equality, and conditions (3.3) – (3.5) to establish the final equality.

Thus for any feasible process $\tilde{y}$ and any perturbation (A.7) defined by a process $\hat{y}$ belonging to $\mathcal{H}_1$,

$$V^Q_{t_0}(y;\xi) = V^Q_{t_0}(\tilde{y};\xi) + V^Q(\hat{y}). \tag{A.8}$$

It follows that if there were to exist any $\hat{y} \in \mathcal{H}_1$ for which $V^Q(\hat{y}) > 0$, the plan $\tilde{y}$ could not be optimal. But as this is true regardless of what plan $\tilde{y}$ may be, (A.5) is necessary for optimality. Furthermore, if there were to exist a non-zero $\hat{y}$ for which $V^Q(\hat{y}) = 0$, it would be possible to construct a perturbation y (not equal to $\tilde{y}$ almost surely at all dates) that would achieve an equally high level of welfare. Hence the stronger version of the second-order conditions (A.6) must hold for all $\hat{y}$ not equal to zero almost surely, in order for $\{\tilde{y}_t\}$ to be a unique optimum.

One easily sees from the same calculation that these conditions are also sufficient for an optimum. Let $\{\tilde{y}_t\}$ be a process consistent with the constraints of the LQ problem. Then any alternative process $\{y_t\}$ that is also consistent with those constraints can be written in the form (A.7), where $\hat{y}$ is some element of $\mathcal{H}_1$. If the first-order conditions (3.2) are satisfied by the process $\{\tilde{y}_t\}$, we can again establish (A.8). Condition (A.5) then implies that no alternative process is preferable to $\{\tilde{y}_t\}$, while (A.6) would imply that $\{\tilde{y}_t\}$ is superior to any alternative that is not equal to $\tilde{y}$ almost surely.

A.2 Proposition 2

Proposition 2 *Suppose that the exogenous disturbances have a law of motion of the form (3.8), where* Γ *is a matrix the eigenvalues of which all have modulus less than* $\beta^{-1/2}$*, and that the constraints satisfy the rank condition (3.14), where* $n_F + n_g < n_y$*. Then the LQ policy problem has a determinate solution, given by (3.21), if and only if (i) there exists a solution* P_{11} *to equations (3.25) such that for each of the minors of the matrix* M *defined in (3.16),* $\det M_r$ *has the same sign as* $(-1)^r$*, for each* $n_F + n_g + 1 \leq r \leq n_y$*; (ii) the eigenvalues of the matrix* Φ_{11} *defined in (3.22) all have modulus less than* $\beta^{-1/2}$*; and (iii) the matrix* P_{22} *defined in (3.26) is negative definite, i.e., is such that its* r*th principle minor has the same sign as* $(-1)^r$*, for each* $1 \leq r \leq n_g$*.*

PROOF: (1) The discussion in the text has already established the necessity of each of conditions (i)–(iii), so it remains only to show that they are also sufficient for the solution (3.21) to represent a solution to the original infinite-horizon optimal policy problem. We shall do this by establishing that conditions (i)–(iii) imply that the sufficient conditions of Proposition 1 are satisfied by this solution.

We begin by establishing that the processes $\{\tilde{y}_t, \tilde{\lambda}_t, \tilde{\varphi}_t\}$ associated with the solution (3.21) satisfy the first-order conditions (3.2) for the infinite-horizon problem. We have already shown in the text that under conditions (i)–(iii), there exists a determinate solution (3.21) for the dynamics of $\{\mathbf{z}_t\}$, that it satisfies the bound (3.1) along with the constraints (2.23)–(2.25), and that associated with it are a unique system of Lagrange multipliers $\{\tilde{\lambda}_t, \tilde{\psi}_t, \tilde{\varphi}_t\}$, the solution for which has also been explained in the text. We wish to show that these processes must satisfy (3.2) for each $t \geq t_0$.

By construction, the processes $\{y_t^\dagger\}$ satisfy the first-order conditions (3.15) for each $t \geq t_0$. Moreover, it follows from (3.23) that

$$\mathbf{P}_1 E_t \mathbf{z}_{t+1} = G_1' E_t y_{t+1}^\dagger.$$

Substituting this into (3.15), we obtain

$$\frac{1}{2}E_t\{[A(L) + A'(\beta L^{-1})]\tilde{y}_t\} + E_t[B(L)\xi_{t+1}] \\ +E_t[C'(\beta L^{-1})\tilde{\lambda}_t] + E_t[D'(\beta L^{-1})\tilde{\psi}_t] \;\; = \;\; 0 \qquad \text{(A.9)}$$

for each $t \geq t_0$.

Differentiating $\bar{V}^Q(\mathbf{z}_t)$ with respect to $\tilde{h}_t$, and using the envelope theorem as in the derivation of (3.23), we obtain $\bar{V}_2^Q = -\tilde{\psi}_t$, from which we conclude that

$$\mathbf{P}_2\,\mathbf{z}_t = -\tilde{\psi}_t$$

for each $t \geq t_0$. Comparison with first-order condition (3.18) for the optimal choice of $\tilde{h}_{t+1}$ in the recursive policy problem indicates that

$$\tilde{\psi}_t = \beta^{-1}\tilde{\varphi}_{t-1} \qquad \text{(A.10)}$$

for each $t \geq t_0+1$. We may assume (as a definition of $\tilde{\varphi}_{t_0-1}$[69]) that (A.10) holds when $t = t_0$ as well. Then use of (A.10) to substitute for the process $\{\tilde{\psi}_t\}$ in (A.9) yields (3.2), which accordingly must hold for each $t \geq t_0$. Hence the processes constructed to satisfy the first-order conditions of the recursive policy problem must satisfy the first-order conditions for the infinite-horizon policy problem characterized in section 3.1 as well.

(2) It remains to show that conditions (i)–(iii) also imply that the strict concavity condition (A.6) is satisfied. Let us consider an arbitrary process $\tilde{y} \in \mathcal{H}_1$, and associated with it define the process $\tilde{h}$ by

$$\tilde{h}_t = D(L)\tilde{y}_t \qquad \text{(A.11)}$$

for each $t \geq t_0 + 1$, and by the stipulation that $\tilde{h}_{t_0} = 0$. We thus obtain a pair of processes satisfying

$$C(L)\tilde{y}_t = 0, \qquad \text{(A.12)}$$

$$D(L)\tilde{y}_t = \tilde{h}_t, \qquad \text{(A.13)}$$

$$E_t\tilde{h}_{t+1} = 0 \qquad \text{(A.14)}$$

[69]Note that $\tilde{\varphi}_{t_0-1}$ has no other meaning in the analysis of the recursive policy problem presented in section 3.2.

for all $t \geq t_0$. These are furthermore an example of a process $\{\mathbf{z}_t\}$ consistent with the constraints of the recursive policy problem, in the case that $\xi_t = 0$ at all times and the initial precommitment is given by $\tilde{h}_{t_0} = 0$.

We note that the analysis given in the text of the single-period problem of maximizing (3.12), applied to the special case in which $\xi_t = 0$ at all times,[70] implies that for any values of $\tilde{y}_{t-1}$ and $\tilde{h}_t$, the maximum possible attainable value of the objective

$$\frac{1}{2}\tilde{y}_t' A(L)\tilde{y}_t + \frac{\beta}{2}E_t[\mathbf{z}_{t+1}' P\mathbf{z}_{t+1}]$$

consistent with constraints (A.12)–(A.14) is equal to

$$\frac{1}{2}\mathbf{z}_t' P\mathbf{z}_t;$$

and this value is attained only if

$$\mathbf{z}_{t+1} = \Phi\ \mathbf{z}_t$$

with certainty, which is to say, only if

$$\tilde{y}_t = \Phi_{11}\,\tilde{y}_{t-1} + \Phi_{12}\,\tilde{h}_t \tag{A.15}$$

and

$$\tilde{h}_{t+1} = 0 \tag{A.16}$$

in each possible state in period $t+1$.

Thus the fact that the processes $\{\tilde{y}_t, \tilde{h}_t\}$ satisfy (A.12)–(A.14) for all $t \geq t_0$ implies that

$$\frac{1}{2}\tilde{y}_t' A(L)\tilde{y}_t + \frac{\beta}{2}E_t[\mathbf{z}_{t+1}' P\mathbf{z}_{t+1}] \leq \frac{1}{2}\mathbf{z}_t' P\mathbf{z}_t$$

for all $t \geq t_0$, and that the inequality is strict unless (A.15)–(A.16) hold. Now if conditions (A.15)–(A.16) hold for all $t \geq t_0$, $\tilde{y}_t = 0$ at all times. Thus in the case that $\tilde{y}_t$ is not equal to zero almost surely for all t, there must be at least one date t_1 such that at least one of these conditions is violated with positive probability when $t = t_1$. In that case, there must be some $k > 0$ such that

$$E_{t_0}\left\{\frac{1}{2}\tilde{y}_{t_1}' A(L)\tilde{y}_{t_1} + \frac{\beta}{2}\mathbf{z}_{t_1+1}' P\mathbf{z}_{t_1+1}\right\}\frac{1}{2} \leq E_{t_0}\mathbf{z}_{t_1}' P\mathbf{z}_{t_1} - k.$$

[70] It follows from the usual principle of certainty equivalence for LQ problems that the matrices characterizing the solution to this problem do not depend on the value of the variance-covariance matrix Σ for the disturbances. In fact, it is easily observed that the derivations given in the text would apply equally to a problem in which $\xi_t = 0$ at all times.

It then follows, by summing these inequalities (appropriately discounted) for successive periods, that

$$E_{t_0}\sum_{t=t_0}^{T}\beta^{t-t_0}\frac{1}{2}\tilde{y}_t' A(L)\tilde{y}_t + \frac{\beta^{T+1-t_0}}{2}E_{t_0}\mathbf{z}_{T+1}' P\mathbf{z}_{T+1} \leq \frac{1}{2}\mathbf{z}_{t_0}' P\mathbf{z}_{t_0} - k = -k, \qquad \text{(A.17)}$$

for all $T \geq t_1$.

As we have stipulated that the process $\tilde{y}$ is an element of $\mathcal{H}_1$, and thus satisfies the bound (3.1), we necessarily have

$$\lim_{T\to\infty}\beta^{T+1}E_{t_0}\mathbf{z}_{T+1}' P\mathbf{z}_{T+1} = 0.$$

(Note that it follows from (A.11) that the elements of $\tilde{h}$ cannot grow asymptotically at a faster rate than do the elements of $\tilde{y}$.) It then follows from (A.17) that

$$\limsup_{T\to\infty} E_{t_0}\sum_{t=t_0}^{T}\beta^{t-t_0}\frac{1}{2}\tilde{y}_t' A(L)\tilde{y}_t \leq -k. \qquad \text{(A.18)}$$

But since it follows from the assumption that $\tilde{y}$ satisfies (3.1) that the series in (A.18) has a limit, this limit must be no greater than $-k$. Hence $\tilde{y}$ satisfies (A.6), and all of the sufficient conditions of Proposition 1 have been verified. This establishes the proposition.

Example: Suppose that y_t has two elements, that the objective of policy is to maximize

$$V_{t_0}^{Q}(\tilde{y}) \equiv \frac{1}{2}E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\tilde{y}_t' A\tilde{y}_t, \qquad \text{(A.19)}$$

where A is a symmetric 2×2 matrix, and that the only constraint on what policy can achieve is a single, forward-looking constraint

$$E_t[\delta\tilde{y}_{1,t} - \tilde{y}_{1,t+1}] = 0 \qquad \text{(A.20)}$$

for all $t \geq t_0$, where $|\delta| < \beta^{-1/2}$. There are no exogenous disturbances, but the expectations appear because we wish to consider the possibility of (arbitrarily) randomized policies. We assume an initial pre-commitment of the form

$$\tilde{y}_{1,t_0} = \delta\tilde{y}_{1,t_0-1} + \tilde{h}_{t_0}, \qquad \text{(A.21)}$$

for some quantity $\tilde{h}_{t_0}$.

In the case that policy is restricted to be deterministic, the constraint completely determines the path of $\{\tilde{y}_{1t}\}$; the only (perfect foresight) sequence consistent with the initial pre-commitment and the forward-looking constraint is the one in which

$$\tilde{y}_{1,t} = [\delta \tilde{y}_{1,t_0-1} + \tilde{h}_{t_0}]\delta^{t-t_0}$$

for all $t \geq t_0$. The problem then reduces to the choice of a sequence $\{\tilde{y}_{2,t}\}$, constrained only by the bound (3.1), so as to maximize the objective. This is obviously a concave problem if and only if $\tilde{y}' A \tilde{y}$ is a concave function of $\tilde{y}_2$ for a given value of $\tilde{y}_1$. This in turn is true if and only if $A_{22} < 0$; the other elements of A are irrelevant.

If instead we allow random policies, the condition just derived is no longer sufficient for concavity (though still necessary). One can show that the problem is concave if and only if A is a negative definite matrix. This is obviously a sufficient condition (as it implies that (A.19) is concave for arbitrary sequences). To show that it is also necessary, suppose instead that it is not true. Then there exists a vector $v \neq 0$ such that $v' A v \geq 0$. Now let $\{\bar{y}_t\}$ be any process satisfying the constraints (3.1), (A.20), and (A.21), and consider the alternative process $\{\tilde{y}_t\}$ generated by the law of motion

$$\tilde{y}_t = \bar{y}_t + \delta(\tilde{y}_{t-1} - \bar{y}_{t-1}) + v\epsilon_t$$

for each $t \geq t_0 + 1$, starting from the initial condition (A.21), where $\{\epsilon_t\}$ is a (scalar-valued) martingale-difference sequence satisfying the bound (3.1). One can easily show that the process $\{\tilde{y}_t\}$ satisfies (3.1), (A.20), and (A.21) as well; moreover, the value of the objective in the case of this process satisfies

$$\begin{aligned} V_{t_0}^Q(\tilde{y}) &= V_{t_0}^Q(\bar{y}) + (1 - \beta\delta^2)^{-1}\, v' A v\, E_{t_0} \sum_{t=t_0+1}^{\infty} \beta^t \epsilon_t^2 \\ &\geq V_{t_0}^Q(\bar{y}). \end{aligned}$$

Since we can construct an alternative policy that is at least as good in the case of *any* policy, there is no uniquely optimal policy in such a case; and in addition, we have shown that arbitrary randomization of policy is possible without welfare loss.

Let us examine how these results compare with the conditions stated in Proposition 2. In this example, condition (3.25) states that

$$P_{11} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

where

$$\alpha = -\delta^2 \, [M^{-1}]_{33}.$$

This form for P_{11} implies in turn that M is invertible as long as $A_{22} \neq 0$, and that in that case,

$$[M^{-1}]_{33} = -\alpha\beta - \frac{|A|}{A_{22}}.$$

Hence we obtain a unique solution,

$$\alpha = \frac{\delta^2}{1-\beta\delta^2}\frac{|A|}{A_{22}}.$$

Since $n_F = 0, n_g = 1, n_y = 2$, condition (i) of the proposition holds if and only if $\det M_2 = \det M > 0$, and under the above solution for P_{11}, $\det M = -A_{22}$; hence condition (i) reduces to the requirement that $A_{22} < 0$.

This solution for P_{11}, and hence for M, also implies that

$$\Phi_{11} = \begin{bmatrix} \delta & 0 \\ -\delta A_{21}/A_{22} & 0 \end{bmatrix}.$$

Hence the eigenvalues of Φ_{11} are 0 and δ. Thus under our assumption about δ, condition (ii) is necessarily satisfied, as long as $A_{22} \neq 0$ (so that Φ_{11} exists). We observe that both conditions (i) and (ii) hold if and only if $A_{22} < 0$, which is just the concavity condition derived above for the deterministic policy problem.

The solution for P_{11} similarly implies that

$$P_{22} = -G_2' M^{-1} G_2 = -[M^{-1}]_{33} = \frac{1}{1-\beta\delta^2}\frac{|A|}{A_{22}}.$$

Since the numerator in this last expression is positive, condition (iii) holds (in addition to the other two conditions) if and only if we also have $\det A > 0$. Since A is negative definite if and only if $A_{22} < 0$ and $\det A > 0$, we can alternatively state that condition (iii) holds (in addition to the other two) if and only if A is also negative definite. This is the additional condition derived above for concavity in the case of stochastic policies.

A.3 Lemma 3

Lemma 3 *Suppose that under optimal policy, the extended state vector* $\mathbf{y_t}$ *consists entirely of components that are either (i) stationary, or (ii) pure random walks. Suppose also that the class of policy rules* $\mathcal{R}$ *is such that each rule in the class implies* *convergence to the same long-run values of the state variables as under optimal policy, in the absence of stochastic disturbances, so that the initial value of the trend component* $\mathbf{y}_{t_0-1}^{tr}$ *is the same regardless of the rule* r *that is considered. Then for any rule* $r \in \mathcal{R}$*, the objective*

$$\Omega_r(\mathbf{y_{t_0-1}}^{tr}) \equiv E_\mu[\bar{W}_r(\mathbf{y_{t_0-1}})], \tag{A.22}$$

can be decomposed into two parts,

$$\Omega_r(\mathbf{y_{t_0-1}}^{tr}) = \Omega^1(\mathbf{y_{t_0-1}}^{tr}) + \Omega_r^2, \tag{A.23}$$

where the first component is the same for all rules in this class, while the second component is independent of the initial condition $\mathbf{y_{t_0-1}}^{tr}$*.*

Proof: We restrict attention to a class of rules $\mathcal{R}$ with the property that each rule in the class implies convergence to the same long-run values of the state variables as under optimal policy, in the absence of stochastic disturbances. Because we analyze the dynamics under a given policy using a linearized version of the structural relations, certainty-equivalence obtains, and it follows that the limiting behavior (as $T \to \infty$) of the long-run forecast $E_{t_0}[\mathbf{y_T}]$ must also be the same under any rule $r \in \mathcal{R}$, given the initial conditions $\mathbf{y_{t_0-1}}$. Thus given these initial conditions, the decomposition of the initial extended state vector into components $\mathbf{y_{t_0-1}}^{tr}$ and $\mathbf{y_{t_0-1}}^{cyc}$ is the same under any rule $r \in \mathcal{R}$.

Let us consider the decomposition

$$\tilde{y}_t = \bar{y}_t + \hat{y}_t,$$

where $\{\bar{y}_t\}$ is the deterministic sequence

$$\bar{y}_t \equiv E_{t_0-1}\tilde{y}_t$$

and $\hat{y}_t$ is the component of $\tilde{y}_t$ that is unforecastable as of date $t_0 - 1$. Then if we evaluate

$$\bar{W}(\tilde{y};\ \mathbf{y}_{t_0-1}) \equiv E_{t_0-1}W(\tilde{y};\ \xi_{t_0}, \mathbf{y_{t_0-1}}),$$

where W is the quadratic form defined in (4.7), under the evolution implied by any rule r, we find that

$$\bar{W}(\tilde{y};\ \mathbf{y}_{t_0-1}) = \bar{W}(\bar{y};\ \mathbf{y}_{t_0-1}) + \bar{W}(\hat{y};\ \mathbf{y}_{t_0-1}). \tag{A.24}$$

Here all the cross terms in the quadratic form have conditional expectation zero because $\bar{y}$ is deterministic while $\hat{y}$ is unforecastable.

Moreover, under any rule r, the value of $\hat{y}_t$ is a linear function of the sequence of unexpected shocks between periods t_0 and t, that is independent of the initial state. (This independence follows from the linearity of the law of motion (4.8), under the linear approximation that we use to solve for the equilibrium dynamics under a given policy rule.) Hence the second term on the right-hand side of (A.24),[71]

$$\bar{W}(\hat{y};\ \mathbf{y}_{t_0-1}) = E_{t_0-1} V^Q_{t_0}(\hat{y}),$$

is independent of the initial state $\mathbf{y}_{t_0-1}$ as well. Let $\bar{W}^2_r$ denote the value of this expression associated with a given rule r.

Instead, the value of $\bar{y}_t$ will be a linear function of $\mathbf{y}_{t_0-1}$, again as a result of the linearity of (4.8). And in our LQ problem with a self-consistent initial precommitment, the function (4.6) is linear as well. It follows that the first term on the right-hand side of (A.24) is a quadratic function of $\mathbf{y}_{t_0-1}$,

$$\bar{W}(\bar{y};\ \mathbf{y}_{t_0-1}) = \mathbf{y}'_{t_0-1} \Xi_r \mathbf{y}_{t_0-1},$$

where the subscript r indicates that the matrix of coefficients Ξ_r can depend on the policy rule that is chosen. Then substituting $\mathbf{y}^{tr}_{t_0-1} + \mathbf{y}^{cyc}_{t_0-1}$ for $\mathbf{y}_{t_0-1}$ in the above expression, and integrating over possible initial values of the cyclical component, for a given initial value of the trend component, we observe that

$$E_\mu[\bar{W}(\bar{y};\ \mathbf{y}_{t_0-1})] = \mathbf{y}^{tr\prime}_{t_0-1} \Xi_r \mathbf{y}^{tr}_{t_0-1} + E_\mu[\mathbf{y}^{cyc\prime} \Xi_r \mathbf{y}^{cyc}], \tag{A.25}$$

using the fact that $E_\mu[\mathbf{y}^{cyc}] = 0$.

Finally, we observe that under any rule r, the linearity of the law of motion (4.8) implies that conditional forecasts of the evolution of the endogenous variables take the form

$$E_{t_0-1} y_T = y^{tr}_{t_0-1} + B_{T+1-t_0} \mathbf{y_{t_0-1}}^{cyc},$$

where the sequence of matrices $\{B_j\}$ may depend on the rule r, but the first term on the right-hand side is the same for all rules in the class $\mathcal{R}$. Using this solution for the sequence $\bar{y}$ to evaluate $\bar{W}(\bar{y};\ \mathbf{y}_{t_0-1})$, we find that the first term in (A.25) must

[71] Here the expected value of the second term on the right-hand side of (4.7) vanishes because of the unforecastability of $\hat{y}_{t_0}$.

be a quadratic function of $y^{tr}_{t_0-1}$ that is the same for all rules r, that can be denoted $y^{tr\prime}_{t_0-1}\bar{\Xi}y^{tr}_{t_0-1}$. Thus if we integrate (A.24) over the invariant distribution μ, we obtain

$$E_\mu[\bar{W}_r(\mathbf{y_{t_0-1}})] = y^{tr\prime}_{t_0-1}\bar{\Xi}y^{tr}_{t_0-1} + E_\mu[\mathbf{y}^{cyc\prime}\Xi_r\mathbf{y}^{cyc}] + \bar{W}^2_r,$$

which is precisely a decomposition of the asserted form (A.23). This proves that the criterion (A.22) establishes the same ranking of alternative rules, regardless of the initial condition.

A.4 Computing the Invariant Measure μ

We need to know the invariant distribution μ over possible initial conditions under optimal policy, in order to compute the proposed welfare criterion (4.12). Because $\bar{W}_r(\cdot)$ is a quadratic function, we only need to compute the unconditional mean and variance-covariance matrix of $\mathbf{y_t}^{cyc}$ under optimal policy.

Substituting (3.19) for the pre-commitment $\tilde{h}_{t+1}$ in the solution (3.17) for the optimal choice of $\tilde{y}_{t+1}$, we observe that under the solution to the recursive policy problem (and hence under the solution to the original problem as well), $\tilde{y}_{t+1}$ is a linear function of $\tilde{y}_t, \xi_{t+1}$, and ξ_t, for each $t \geq t_0$. This solution together with the process (3.8) for the exogenous disturbances imply a law of motion of the form

$$\mathbf{y}_{t+1} = \bar{\Phi}\,\mathbf{y}_t + \bar{\Psi}\,\epsilon_{t+1} \tag{A.26}$$

for the extended state vector

$$\mathbf{y}_t \equiv \begin{bmatrix} \tilde{y}_t \\ \xi_t \end{bmatrix}. \tag{A.27}$$

Under this law of motion, the trend component of the extended state vector is given by $\mathbf{y_t}^{tr} = \Pi\mathbf{y_t}$, where Π is the matrix[72]

$$\Pi \equiv \lim_{j\to\infty} \bar{\Phi}^j,$$

and the cyclical component is correspondingly given by $\mathbf{y_t}^{cyc} = [I - \Pi]\mathbf{y_t}$. It then follows that the law of motion for the cyclical component is

$$\mathbf{y_{t+1}}^{cyc} = \bar{\Phi}\,\mathbf{y_t}^{cyc} + [I - \Pi]\bar{\Psi}\,\epsilon_{t+1}. \tag{A.28}$$

[72]Under the assumption (made in the text) that the extended state vector is difference-stationary, this limit must be well-defined.

We note furthermore that (A.28) describes a jointly stationary set of processes, since the matrix $\bar{\Phi}$ is stable on the subspace of vectors $\mathbf{v}$ of the form $\mathbf{v} = [I - \Pi]\mathbf{y}$ for some vector $\mathbf{y}$.[73] Hence there exist a well-defined vector of unconditional means $\mathbf{E}$ and an unconditional variance-covariance matrix $\mathbf{V}$. The unconditional means are all zero, while the matrix V is given by the solution to the linear equation system

$$\mathbf{V} = \bar{\Phi}\mathbf{V}\bar{\Phi}' + [I - \Pi]\bar{\Psi}\Sigma\bar{\Psi}'[I - \Pi'].$$

In the case of some policy rules, it may be necessary to include additional lags of $\tilde{y}_t$ or ξ_t in the extended state vector $\mathbf{y}_t$, in order for the equilibrium dynamics under the rule r to have a representation of the form (4.8). However, in this case, the additional elements of $\mathbf{y}_t^{cyc}$ will all be lags of elements of the vector considered above. Hence the law of motion (A.28) can be used to derive the relevant unconditional moments in this case as well (though we omit the algebra).

[73] When restricted to this subspace, the operator $\bar{\Phi}$ has eigenvalues consisting of those eigenvalues of $\bar{\Phi}$ that are less than one in modulus; these are in turn a subset of the eigenvalues of Φ that are less than one in modulus (some zero eigenvalues have been dropped).

References

[1] Altissimo, Filippo, Vasco Curdia, and Diego Rodriguez Palenzuela (2005), "Linear-Quadratic Approximation to Optimal Policy: An Algorithm and Two Applications," paper presented at the conference "Quantitative Analysis of Stabilization Policies," Columbia University, September.

[2] Amman, Hans M., and David A. Kendrick (1999), "The Duali/Dualpc Software for Optimal Control Models: User's Guide," Center for Applied Research in Economics, University of Texas.

[3] Backus, David, and John Driffill (1986), "The Consistency of Optimal Policy in Stochastic Rational Expectations Models," CEPR Discussion Paper No. 124.

[4] Barro, Robert J. (1979), "On the Determination of Public Debt," *Journal of Political Economy* 87: 940-971.

[5] Benigno, Gianluca, and Pierpaolo Benigno (2006), "Designing Targeting Rules for International Monetary Policy Cooperation," *Journal of Monetary Economics,* 53(3), pages 473-506.

[6] Benigno, Gianluca and Bianca De Paoli (2005), "Optimal Monetary and Fiscal Policy for a Small Open Economy," unpublished manuscript, London School of Economics.

[7] Benigno, Pierpaolo, and Michael Woodford (2003), "Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach," in M. Gertler and K. Rogoff, eds., *NBER Macroeconomics Annual 2003,* Cambridge, MA: MIT Press.

[8] Benigno, Pierpaolo, and Michael Woodford (2005a), "Inflation Stabilization And Welfare: The Case Of A Distorted Steady State," *Journal of the European Economic Association,* vol. 3(6), pages 1185-1236, December.

[9] Benigno, Pierpaolo, and Michael Woodford (2005b), "Optimal Stabilization Policy when Wages and Prices are Sticky: The Case of a Distorted Steady State," in J. Faust, A. Orphanides, and D. Reifschneider (eds.) *Models and Monetary Policy,* Board of Governors of the Federal Reserve System: Washington, pp. 127-180.

[10] Benigno, Pierpaolo, and Michael Woodford (2006a), "Optimal Taxation in an RBC Model: A Linear Quadratic Approach," *Journal of Economic Dynamics and Control,* 30(9-10): 1445-1489.

[11] Benigno, Pierpaolo, and Michael Woodford (2006b), "Linear-Quadratic Approximation of Optimal Policy Problems," NBER Working Paper no. 12672, November.

[12] Bertsekas, Dimitri P. (1976), *Dynamic Programming and Stochastic Control*, New York: Academic Press.

[13] Blake, Andrew P. (2001), "A 'Timeless Perspective' on Optimality in Forward-Looking Rational Expectations Models," NIESR Discussion Papers 188, National Institute of Economic and Social Research.

[14] Chow, Gregory C. (1975), *Analysis and Control of Dynamic Economic Systems* John Wiley & Sons, New York.

[15] Clarida, Richard, Jordi Gali and Mark Gertler (1999), "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature* 37: 1661-1707.

[16] Curdia, Vasco, "Optimal Monetary Policy under Sudden Stops," unpublished, Federal Reserve Bank of New York, April 2007.

[17] Debreu, Gerard, "Definite and Semidefinite Quadratic Forms," *Econometrica* 20: 295-300 (1952).

[18] De Paoli, Bianca (2004), "Monetary Policy and Welfare in a Small Open Economy," unpublished manuscript, London School of Economics.

[19] Diaz-Gimenez, Javier, "Linear-Quadratic Approximations: An Introduction," in Ramon Marimon and Andrew Scott, editors, *Computational Methods for the Study of Dynamic Economies,* Oxford: Oxford University Press, 1999.

[20] Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics* 46: 281-313.

[21] Ferrero, Andrea (2005), "Fiscal and Monetary Rules for a Currency Union," unpublished manuscript, New York University.

[22] Fleming, Wendell H. (1971), "Stochastic Control for Small Noise Intensities," *SIAM Journal of Control* 9: 473-517.

[23] Gerali, Andrea and Francesco Lippi (2005), "Solving Dynamic Linear-Quadratic Problems with Forward-Looking Variables and Imperfect Information using Matlab," unpublished, Bank of Italy, November.

[24] Giannoni, Marc, and Michael Woodford (2002), "Optimal Interest-Rate Rules: I. General Theory,"NBER Working Paper no. 9419, December.

[25] Hansen, Lars P., and Thomas J. Sargent (2004), *Recursive Models of Dynamic Linear Economies*, unpublished, University of Chicago, August 2004.

[26] Jensen, Christian, and Bennet C. McCallum (2002), "The Non-Optimality of Proposed Monetary Policy Rules Under Timeless-Perspective Commitment," NBER Working Paper No. 8882.

[27] Jin, Hehui, and Kenneth L. Judd (2002), "Perturbation Methods for General Dynamic Stochastic Models," unpublished, Stanford University.

[28] Judd, Kenneth L. (1998), *Numerical Methods In Economomics*, Cambridge, Mass.: MIT Press.

[29] Juillard, Michel, and Florian Pelgrin, "Computing Optimal Policy in a Timeless Perspective: An Application to a Small Open Economy," unpublished, University of Paris VIII, September 2006.

[30] Kalchbrenner, J.H., and Peter A. Tinsley (1975), "On the Use of Optimal Control in the Design of Monetary Policy," Special Studies Paper No. 76, Federal Reserve Board.

[31] Kendrick, David A. (1981), *Stochastic Control for Economic Models,* New York: McGraw-Hill.

[32] Kendrick, David A. (2005), "Stochastic Control for Economic Models: Past, Present and the Paths Ahead," *Journal of Economic Dynamics and Control* 29: 3-30.

[33] Khan, Aubhik, Robert G. King, and Alexander L. Wolman (2003), "Optimal Monetary Policy," *Review of Economic Studies* 70(4): 825-860.

[34] Kim, Jinill, and Sunghyun Kim (2003), "Spurious Welfare Reversal in International Business Cycle Models", *Journal of International Economics,* Volume 60, Issue 2, Pages 471-500 .

[35] Kim, Jinill, and Sunghyun Kim (2006), "Two Pitfalls of Linearization Methods," unpublished, Federal Reserve Board, April. (Forthcoming, *Journal of Money, Credit and Banking.*)

[36] Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A. Sims (2003), "Calculating and Using Second Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models," unpublished manuscript, University of Virginia, June.

[37] King, Robert G., and Alexander L. Wolman (1999), "What Should the Monetary Authority Do When Prices are Sticky?" in J.B. Taylor, ed., *Monetary Policy Rules,* Chicago: University of Chicago Press.

[38] Kydland, Finn E., and Edward C. Prescott (1982), "Time to Build and Aggregate Fluctuations," *Econometrica,* vol. 50(6), pages 1345-70, November.

[39] Kwakernaak, Huibert, and Raphael Sivan (1972), *Linear Optimal Control Systems*, New York: Wiley.

[40] LeRoy, Stephen F., and Roger N. Waud (1977), "Applications of the Kalman Filter in Short-run Monetary Control," *International Economic Review* 18, 195–207.

[41] Levine, Paul, Joseph Pearlman, and Richard Pierse, "Linear-Quadratic Approximation, External Habit, and Targeting Rules," ECB Working Paper no. 759, June 2007.

[42] Magill, Michael J.P. (1977), "A Local Analysis of N-Sector Capital Accumulation under Uncertainty," *Journal of Economic Theory* 15: 211-218.

[43] Rotemberg, Julio J., and Michael Woodford (1997), "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual* 12: 297-346.

[44] Sargent, Thomas J. (1987), *Macroeconomic Theory*, 2d edition, New York: Academic Press.

[45] Sauer, Stephan (2006), "Discretion Rather than Rules? When is Discretionary Policy-Making Better Than the Timeless Perspective?," unpublished manuscript, Ludwig-Maximilians-University Munich, July.

[46] Schmitt-Grohé, Stephanie, and Martin Uribe (2004a), "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, vol. 28, pp. 755-775.

[47] Schmitt-Grohé, Stephanie, and Martin Uribe (2004b), "Optimal Fiscal and Monetary Policy under Sticky Prices," *Journal of Economic Theory* vol. 114, pp. 198-230.

[48] Schmitt-Grohé, Stephanie, and Martin Uribe (2004c), "Optimal Simple and Implementable Monetary and Fiscal Rules," NBER Working Paper no. 10253, January.

[49] Söderlind, Paul (1999), "Solution and estimation of RE macromodels with optimal policy," *European Economic Review*, vol. 43(4-6), pages 813-823, April.

[50] Sutherland, Alan (2002), "A Simple Second-Order Solution Method for Dynamic General Equilibrium Models," CEPR discussion paper no. 3554, July.

[51] Takayama, Akira, *Mathematical Economics*, Cambridge: Cambridge University Press, 2d ed., 1985.

[52] Woodford, Michael (2002), "Inflation Stabilization and Welfare," *Contributions to Macroeconomics* 2(1), Article 1. [www.bepress.com]

[53] Woodford, Michael (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton: Princeton University Press.