Macroeconomic Dynamics, **2**, 1998, 156–182. Printed in the United States of America.

CALCULATION, ADAPTATION AND RATIONAL EXPECTATIONS

GEORGE W. EVANS
University of Oregon

GAREY RAMEY
University of California at San Diego

We propose an active cognition approach to bounded rationality, in which agents use a calculation algorithm to improve on the forecasts provided by a purely adaptive learning rule such as least-squares learning. Agents' choices of calculation intensity depend on their estimates of the benefits of improved forecasts relative to calculation costs. Using an asset-pricing model, we show how more rapid adjustment to rational expectations and forward-looking behavior arise naturally when there are large anticipated structural changes such as policy shifts. We also give illustrative applications in which the severity of asset price bubbles and the intensity of hyperinflationary episodes are related to the cognitive ability of the agents.

Keywords: Expectations, Bounded Rationality, Active Cognition, Calculation Algorithm, Adaptation

1. INTRODUCTION

Although the rational-expectations approach has been the dominant methodology in economics for over two decades, economists still have scant understanding of the processes by which agents might acquire rational expectations. A recent outpouring of work, however, has explored the idea that rational expectations may emerge from *passive adaptation* on the part of agents. The typical model in this vein endows agents with an adaptive learning rule that revises expectations in a given way each time a new data point becomes available.[1] A drawback of the passive adaptation approach is that agents are not allowed to engage in any inferential process beyond the updating procedure that is imposed by their learning rule. Importantly, publicly available information concerning abrupt structural changes or future policies can elicit only gradual responses that occur after the fact, because expectation revision is driven solely by the ex-post accumulation of data. In essence, passive adaptation restricts agents to highly inertial, backward-looking behavior, which rules out the possibility of fast-adjusting, forward-looking behavior that may be realistic in many contexts.

We are indebted to John Conlisk, Vince Crawford, Wouter den Haan, Roger Farmer, Joel Sobel and numerous seminar participants. This research was supported by NSF Grant SES-9210405. Address correspondence to: Professor George W. Evans, Department of Economics, 1285 University of Oregon, Eugene, OR 97403-1285, USA; e-mail: gevans@oregon.uoregon.edu.

© 1998 Cambridge University Press 1365-1005/98 $9.50

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

We propose a new approach to modeling expectation formation, in which expectations emerge from *active cognition* by agents. The key hypothesis in our expectation-formation framework is that agents possess a *calculation algorithm* that they use to formulate expectations. The algorithms that we study, which are based on a true model of the economy, incorporate information about present and future economic structure, and can be used with variable intensity within a given period. Thus, agents are given an explicit capacity to adjust expectations rapidly, and also to anticipate future events. Further, the algorithms are able to approximate rational expectations arbitrarily closely if calculation intensity and the forecast horizon are sufficiently great. In our framework, the cognitive effort exerted by agents determines the extent and speed of adjustment toward rational expectations.

Adaptive behavior also plays a role in our framework. Agents possess *input functions* that specify the initial inputs that are fed into the algorithm on its first iteration in any period. The input functions map observed economic data, as well as calculations made in past periods, to expectations that serve to initialize calculation. If agents choose not to calculate, then expectations are determined by the input functions; thus, the input functions constitute the passive adaptive component of our framework. As a consequence, our approach may be viewed as building a calculation capacity on top of an adaptive learning scheme.

In the active cognition approach to bounded rationality, a central issue is the determination of the intensity of cognitive effort. In our past work on expectation calculation [Evans and Ramey (1992, 1995)], we resolved this issue by positing that agents choose their intensity of calculation *optimally*; that is, they make the calculation decision that gives the optimal balance between their *actual* losses from forecast errors and the costs of calculation.[2] We break new ground by developing an alternative approach to calculation decisions, which utilizes the algorithm itself to assess the gains in forecast accuracy obtaining from calculation, rather than endowing agents with "correct guesses" of these gains. The key idea is that agents perform a *diagnostic calculation* as part of each iteration of the calculation algorithm. The diagnostic calculation computes a *target expectation* that is used to evaluate how the algorithm has performed on its previous iterations. Agents cease calculating when the reduction in losses from forecast errors, as assessed relative to the target expectation, fails to make up for the costs of calculating.

We continue our practice of developing our calculation framework in the context of a specific economic model. Here, we consider a version of the classical overlapping-generations model, in which there is a single durable asset. Young agents' demands for the asset are determined by the expected rate of return. We endow the agents with calculation algorithms that are well suited for computing expected returns in dynamic macroeconomic settings; in particular, agents may use one of two versions of the extended path algorithm of Fair and Taylor (1983), or the parameterized-expectations algorithm of Marcet (1991). A very broad class of adaptive learning rules gives suitable specifications of the input function. In particular, we develop an application that generates initial inputs using an autoregressive

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

forecast rule, the parameters of which are estimated from the observed data via recursive least squares. Calculation decisions are made according to a simple *adaptive sequential calculation decision rule* (CDR) under which calculation is halted as soon as the preceding iteration is assessed to have yielded a negative net benefit.

We verify analytically that, in a stationary environment and with suitably restricted versions of the three calculation algorithms, the adaptive sequential CDR generates expectations that give an arbitrarily close approximation to rational expectations, when the forecast horizon of the algorithm is sufficiently long and calculation costs are sufficiently low.[3] We show further, however, that the speed of adjustment to rational expectations, and the degree to which agents anticipate future structural changes in a nonstationary environment, depend on the level of calculation costs. In a numerical simulation that relates the path of expected returns to the costs of calculation for an anticipated policy change, the solution path shifts smoothly from the fully adaptive solution to the fully rational solution as calculation costs fall to zero. In this way, our framework fills in the gap between adaptive and rational expectations.

To illustrate potential applications of our approach, we assess the implications of expectation calculation for asset pricing when the adaptive forecasting rule leads agents to forecast asset price bubbles. The characteristics of bubbles are sensitive to the level of calculation costs: As calculation costs fall, bubbles become smaller and break sooner, and the crashes that ensue are less intense. In the limit, the solution path approaches the steady-state rational-expectations equilibrium (REE). Thus, our expectation-formation framework yields a theory in which financial-market instability is linked explicitly to the cognitive abilities of traders.

In a second application, we develop an extension of the overlapping-generations model in which the government may issue money to finance persistent deficits, giving an overlapping-generations version of the Cagan-Fisher hyperinflation model [see Evans and Ramey (1995)]. The government's ability to temporarily run an unsustainably-high deficit depends critically on calculation costs: For lower levels of calculation costs, agents' expectations adjust faster during the high-deficit phase, giving rise to a hyperinflationary episode that terminates in collapse of the currency. Thus, the costs and benefits of policy choices, as well as their long-run effects, can depend critically on agents' cognitive abilities as expressed by calculation costs.

In addition to developing a new approach for determining calculation decisions, the current paper builds on our two previous papers by allowing calculation algorithms to have an arbitrarily long forecast horizon, and also input functions may utilize observed data. We restrict attention to a deterministic environment; fully incorporating stochastic elements into our calculation framework remains a high priority for future work. In addition, we continue our practice of assuming that agents possess an accurate structural model of the economy with which they can perform calculations. We also plan to relax this assumption in future work. There exist few other papers on active-cognition approaches to modeling bounded rationality in economic contexts; see Conlisk (1988) and de Palma et al. (1994)

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

for models in which cognitive effort serves to reduce stochastic errors in one-agent decision models. Conlisk (1996) gives methodological arguments in favor of bounded-rationality approaches, along with an extensive survey of the literature.

The sequel proceeds as follows. Section 2 introduces the overlapping-generations model that gives our economic structure, and Section 3 develops the calculation framework. Analytical results are given in Section 4, a numerical example of an anticipated policy change is presented in Section 5, applications to asset price bubbles and deficit policy are developed in Section 6, and conclusions are presented in Section 7. Proofs of the analytical results are collected in the Appendix.

2. ECONOMIC MODEL

In this section, we develop a structural economic model that will form the base on which we construct our calculation framework. We consider a deterministic version of the classical overlapping-generations model in which there is a single durable asset. Each generation consists of a unit mass of atomistic agents that live for two periods. Let c_1 and c_2 denote consumption in the first and second periods for each generation. Utility is given by $V(c_1) + W(c_2)$, where we assume $V', W' > 0$, $V'', W'' < 0$ and $\lim_{c_1 \to 0} V'(c_1) = \lim_{c_2 \to 0} W'(c_2) = \infty$. Let $y_1 > 0$ denote first-period endowment income, and assume that second-period endowment income is zero.

There is a single durable asset, the total supply of which is G in every period. The asset yields a dividend of $d > 0$ per unit in each period. Let q_t denote the price of the asset in period t. The one-period gross rate of return on the asset between periods t and $t+1$, written r_{t+1}, is given by

$$r_{t+1} = \frac{d + q_{t+1}}{q_t}.$$

The period-1 consumption of each generation is given by $c_1(r^e_{t+1})$, which solves

$$\max_{c_1} \left\{ V(c_1) + W\left((y_1 - c_1) r^e_{t+1}\right) \right\}, \tag{1}$$

where r^e_{t+1} gives agents' expectation of r_{t+1} held in period t. Throughout this paper, we make the assumption that the substitution effect strictly dominates the income effect, so that $c_1(r^e_{t+1})$ is a strictly decreasing and differentiable function that satisfies $0 < c_1(r^e_{t+1}) < y_1$ for every positive r^e_{t+1}. It is also convenient to impose the added condition that $c_1(r^e_{t+1}) \to 0$ as $r^e_{t+1} \to \infty$ and $c_1(r^e_{t+1}) \to y_1$ as $r^e_{t+1} \to 0$.

The level of q_t that clears the asset market, written $F(r^e_{t+1})$, is determined by

$$F\left(r^e_{t+1}\right) = \frac{y_1 - c_1\left(r^e_{t+1}\right)}{G}. \tag{2}$$

Observe that $F(r^e_{t+1})$ is strictly positive, strictly increasing, and differentiable, and that F maps the positive reals onto $(0, y_1/G)$. For a given pair of expected returns

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

r^e_{t+1}, r^e_{t+2}, the realized return r_{t+1} is expressed as the function $T(r^e_{t+1}, r^e_{t+2})$, which we call the *T-mapping*. In particular, the T-mapping is defined by

$$T\left(r^e_{t+1}, r^e_{t+2}\right) = \frac{d + F\left(r^e_{t+2}\right)}{F\left(r^e_{t+1}\right)}. \tag{3}$$

The T-mapping is strictly decreasing in r^e_{t+1}, strictly increasing in r^e_{t+2}, and differentiable. This map provides the basis for the calculation algorithms discussed below.

In a REE, agents' expected returns are in agreement with the realized returns generated by those expectations. Because our model is nonstochastic, rational expectations amount to perfect foresight and such a solution will satisfy $r_{t+1} = T(r_{t+1}, r_{t+2})$. Let $r_{t+1} = S(r_{t+2})$ denote the *perfect-foresight solution* implicitly defined by this equation. $S(r_{t+2})$ gives the level of r_{t+1} that is associated with an equilibrium in period t for given r_{t+2}:

$$S(r_{t+2}) = \frac{d + F\left(r_{t+2}\right)}{F(S(r_{t+2}))}.$$

It is straightforward to verify that there exists a single value $\hat{r}$ that satisfies $\hat{r} = S(\hat{r})$. Thus, $\hat{r}$ is a unique steady-state REE. Further, S is strictly increasing, differentiable with $S'(\hat{r}) < 1$, and maps the positive reals into a bounded interval. It follows that no REE exists other than the steady state: If $r_1 > \hat{r}$ initialized a REE, then returns would follow the upward-sloping path determined by $r_t = S(r_{t+1})$, which eventually would generate returns too large to be compatible with market clearing. Similarly, $r_1 < \hat{r}$ would initialize a downward-sloping path of returns that would eventually violate market clearing.[4]

3. CALCULATION FRAMEWORK

3.1. Overview

Before giving a detailed presentation of the particular calculation algorithms that we consider, it is helpful first to present a general overview of our expectation-formation framework. The basic idea is that agents possess a recursive numerical algorithm that they use to calculate expectations in each period. From the information that they hold at the start of a period, the agents determine a starting point, or initial input, for calculation in the period. Agents then perform iterations of the algorithm, each of which takes the output of the preceding iteration as input. After some number of iterations, agents cease calculating and use the output of the final iteration as the expectation on which their economic decisions are based.

More formally, let β be an *expectation parameter* that represents possible expectations agents may hold. Let $\mathcal{B}$ give the set of all β, and let $\mathcal{B}^k$ denote the set of all k-tuples of elements of $\mathcal{B}$ for given k. Let h_t denote an agents' *information set* at period t, which comprises (1) the history of realized returns up to period

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

$t-1$; (2) expectations held by the agent in preceding periods; and (3) information about future economic structure that the agent possesses at period t, such as preannouncements of future government policy. Let $\mathcal{H}_t$ give the set of h_t. The expectation-formation framework is composed of three ingredients:

1. A *calculation algorithm*, $A : \mathcal{H}_t \times \mathcal{B} \to \mathcal{B}$, which recursively determines outputs $\beta^k = A(h_t, \beta^{k-1})$ from an initial input β^0.
2. An *input function*, $I : \mathcal{H}_t \to \mathcal{B}$, which maps information sets to initial inputs $\beta^0 = I(h_t)$.
3. A CDR, $K : \mathcal{H}_t \times \cup_{k=0}^{\infty} \mathcal{B}^{k+1} \to$ {Continue, Stop}, which indicates whether to continue calculating or to stop following the kth iteration, given h_t and the ensuing calculation trajectory $\beta^0, \ldots, \beta^k$.

Calculation ceases at the first value of k such that $K(h_t, \beta^0, \ldots, \beta^k) = \text{Stop}$. Let this value of k be denoted by k_t^C; thus, k_t^C gives the number of calculations that are performed in period t, given the information set h_t. Let $\beta_t^C = \beta^{k_t^C}$ denote the final output of the algorithm in period t. In the context of the structural model of Section 2, β_t^C determines an expected return $r_{t+1}^e(\beta_t^C)$ and a realized price $q_t(r_{t+1}^e(\beta_t^C))$. Further, for an information set h_{t+1} that is a continuation of h_t, the expectations β_t^C and β_{t+1}^C calculated in periods t and $t+1$ determine the realized rate of return $r_{t+1} = T(r_{t+1}^e(\beta_t^C), r_{t+2}^e(\beta_{t+1}^C))$.[5]

3.2. Calculation Algorithms

We now introduce some particular specifications of calculation algorithms that can be used to implement our framework. These algorithms have been designed for calculation of REE in dynamic macroeconomic applications, making them well-suited for expectation calculation by our agents. Each of the algorithms we consider is based upon the T-mapping described in the previous section, i.e., upon the true structural model of the economy. This assumption could be relaxed to allow agents to have misspecified models of the economy, but endowing them with a correct structural model is a natural benchmark specification for our calculation framework.

3.2.1. Extended path (EP) algorithm. The EP algorithm of Fair and Taylor (1983) calculates expectations by iterating a structural model of the economy out to a given forecast horizon. Expectations in period t are parameterized by $\beta_t = (r_{t,1}, r_{t,2}, \ldots, r_{t,H+1})$, where H gives the forecast horizon of the algorithm, and $r_{t,s}$ denotes the expectation of r_{t+s} held at period t. The EP algorithm fixes $r_{t,H+1}$ at its initial input level $r_{t,H+1}^0$, and on each iteration it adjusts the other rates $r_{t,s}$, $s \le t$ toward the realized rates that would be generated by the expectation β_t^{k-1}. Thus, iterations are performed as follows:

$$\begin{aligned} r_{t,H+1}^k &= r_{t,H+1}^{k-1}, \\ r_{t,s}^k &= \lambda T\left(r_{t,s}^{k-1}, r_{t,s+1}^{k-1}\right) + (1-\lambda) r_{t,s}^{k-1} \end{aligned} \tag{4}$$

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

for $s = 1, \ldots, H$. The damping factor $\lambda \in (0, 1)$ is applied to counteract numerical problems in calculating equilibria, such as explosive oscillation.

3.2.2. Modified extended path (MEP) algorithm. The EP algorithm can be modified by making use of the fact that $r^k_{t,H+1}$ remains fixed, while the values $r^{k-1}_{t,s+1}$ are recursively fed backward into revisions of earlier forecasted rates. This makes it possible to use newly calculated values $r^k_{t,s+1}$, rather than the values $r^{k-1}_{t,s+1}$ from the previous iteration, in revising the earlier rates. In particular, the MEP algorithm operates as follows:

$$\begin{aligned} r^k_{t,H+1} &= r^{k-1}_{t,H+1}, \\ r^k_{t,s} &= \lambda T\left(r^{k-1}_{t,s}, r^k_{t,s+1}\right) + (1-\lambda) r^{k-1}_{t,s}. \end{aligned} \tag{5}$$

Observe that the only difference between EP and MEP is that $r^{k-1}_{t,s+1}$ in equation (4) is replaced by $r^k_{t,s+1}$ in equation (5). In essence, EP iterates the entire string $(r_{t,1}, \ldots, r_{t,H+1})$, whereas MEP revises the string element-by-element via backward recursion. An advantage of MEP is that it makes use of information that is calculated within the current iteration.

3.2.3. Parameterized expectations (PE) algorithm. The PE algorithm of Marcet (1991) specifies a function $J : \mathcal{H}_t \times \mathcal{B} \to R^{H+1}$ that, for a given information set h_t, maps the expectation parameter β to a string $(r_{t,1}, \ldots, r_{t,H+1})$ giving the forecasted rates of return. The key feature of the PE algorithm is that the set of possible expectation strings. $(r_{t,1}, \ldots, r_{t,H+1})$, which may be very large, is reduced to a more managable subset that is parameterized by β. Each iteration of the PE algorithm consists of two steps.

Step 1. Simulated data $\hat{r}^k_s, s = 1, \ldots, H$ are generated by combining the forecasted returns $J(h_t, \beta^{k-1}_t) = (r^{k-1}_{t,1}, \ldots, r^{k-1}_{t,H+1})$ with the T-mapping:

$$\hat{r}^k_s = T\left(r^{k-1}_{t,s}, r^{k-1}_{t,s+1}\right). \tag{6}$$

Step 2. The model $\hat{r}^k_s = J_s(h_t, \beta), s = 1, \ldots, H$ is used to estimate $\hat{\beta}^k_t$, e.g., by nonlinear least squares, and the expectation parameter is adjusted toward $\hat{\beta}^k_t$:

$$\beta^k_t = \lambda \hat{\beta}^k_t + (1-\lambda)\beta^{k-1}_t.$$

For expositional ease, we have thus far focused on a stationary structural environment; in particular, the parameters y_1, G, and d have been taken to be invariant across periods. It is immediate, however, to extend the preceding calculation algorithms to incorporate nonstationarity. For example, if first-period endowment incomes are given by a sequence $y_{1,t}, t = 1, 2, \ldots$, then we may let $T_t(r^e_{t+1}, r^e_{t+2})$ give the T-mapping that corresponds to the pair $y_{1,t}, y_{1,t+1}$.[6] The preceding algorithms then may be modified by using $T_{t+s}(r^e_{t+1}, r^e_{t+2})$ in (4), (5), and (6). In this

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

way, the algorithms can incorporate agents' beliefs about future changes in the environment that are contained in h_t.

3.3. Input Functions

The input function generates an agent's initial guess of the REE path that is obtained from past observed rates of return and the agent's own past calculations. As such, the input function represents the passive adaptation component of our expectation-formation framework. Suitable specifications of the input functions may be drawn from the variety of learning schemes that have been studied in the recent literature.[7] The following gives a representative example of a recursive least-squares learning rule.

3.3.1. Autoregressive forecasting rule. Let $(r^0_{t,1}, \ldots, r^0_{t,H+1})$ be determined recursively by

$$r^0_{t,s+1} = \alpha_{0,t} + \alpha_{1,t} r^0_{t,s} \tag{7}$$

for $s = -1, 0, 1, \ldots, H$, where under our timing assumptions, $r^0_{t,-1} = r_{t-1}$ gives agents' most recent observation of the realized rate of return as of period t. The parameters $\alpha_t = [\alpha_{0,t}, \alpha_{1,t}]'$ are estimated using recursive least squares. Putting $x_{t-1} = [1, r_{t-1}]'$, we have

$$\alpha_t = \alpha_{t-1} + \gamma_t Z_t^{-1} x_{t-1}(r_{t-1} - x'_{t-1}\alpha_{t-1}), \tag{8}$$

$$Z_t = Z_{t-1} + \gamma_t(x_{t-1}x'_{t-1} - Z_{t-1}) \tag{9}$$

for given initial values α_0 and Z_0, where $\gamma_t, t = 1, 2, \ldots$ is a prespecified gain sequence (recursive least-squares sets $\gamma_t = 1/t$). The string $(r^0_{t,1}, \ldots, r^0_{t,H+1})$ gives β^0_t directly for the EP and MEP algorithms, whereas for the PE algorithm the initial value β^0_t is obtained by fitting $r^0_{t,s} = J_s(h_t, \beta)$, $s = 1, \ldots, H+1$.

3.3.2. Prior calculations. Alternatively, an agent may rely solely on his calculations from preceding periods to provide initial inputs for calculations in the current period. Let $(r^C_{t-1,1}, \ldots, r^C_{t-1,H+1})$ give the final outputs from period $t-1$ calculations. Period-t inputs are given by $r^0_{t,s} = r^C_{t-1,s+1}$ for $s = 1, \ldots, H$, whereas $r^0_{t,H+1}$ is determined using some additional relationship. Input functions also may combine past calculations with past observations, e.g., initial inputs might be an average of the autoregressive forecasts and the past calculations.

3.4. CDR

The CDR determines the number of iterations to be performed as a function of the information set and the realized calculation trajectory. One simple specification of the CDR, which mimics the practice of applied macroeconomists, would set a tolerance $\delta > 0$ and stop as soon as $|\beta^k_t - \beta^{k-1}_t| < \delta$. We are interested, however, in obtaining a CDR that reflects a balancing of the costs and benefits of calculation;

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

for example, an agent with a lower cost of computation presumably would set a smaller value of δ.

We consider a CDR wherein agents perform a *diagnostic calculation* to estimate the gain in utility due to reduced forecast errors, and then compare this estimate to the costs of calculation in order to evaluate the net benefit of continued calculation. Let $U(r^e_{t+1}, r_{t+1})$ denote the agent's utility when the expected return is r^e_{t+1} and the realized return is r_{t+1}:

$$U\left(r^e_{t+1}, r_{t+1}\right) = V\left(c_1\left(r^e_{t+1}\right)\right) + W\left(\left(y_1 - c_1\left(r^e_{t+1}\right)\right)r_{t+1}\right).$$

Let v_k denote the utility cost of the kth iteration for $k = 1, 2, \ldots, n$, where n gives the maximum number of iterations that are possible in a period. Beginning with an initial input β^0_t and $k = 1$, the calculation decision is determined as follows:

Step 1. Calculate β^k_t.

Step 2. Calculate a *target expectation* $r^*_{t+1} = T(r^k_{t,1}, r^k_{t,2})$. Stop calculating if either $k = n$ or

$$U\left(r^k_{t,1}, r^*_{t+1}\right) - v_k < U\left(r^{k-1}_{t,1}, r^*_{t+1}\right). \tag{10}$$

Otherwise, continue to the $k + 1$st iteration.

Here, step 1 represents the usual operation of the calculation algorithm, whereas step 2 gives the diagnostic calculation that assesses the performance of the algorithm by estimating the net benefit of the preceding iteration, relative to a target return computed from the T-mapping. The target return indicates roughly where the calculation trajectory is headed, allowing agents to assess the marginal impact of the latest iteration on forecast accuracy. Note that the CDR requires at least one iteration of the algorithm to be performed in each period. This initial *test calculation* forces at least one pass through the algorithm in every period, and it thereby ensures that anticipations of future structural changes are incorporated into the estimated net benefits of calculation.

This CDR is *adaptive*, in the sense that it relies only on information that is either in the agent's information set or obtained by the agent's own calculations. The CDR is also *sequential*, because a decision as to whether to stop calculating is made after each iteration.[8] Thus, we refer to it as the *adaptive sequential CDR*. Importantly, the adaptive sequential CDR does not necessarily generate a calculation decision that is *fully optimal* from the agent's point of view, in contrast to the optimal CDR considered by Evans and Ramey (1992, 1995). In the present case, calculation decisions may be suboptimal either because the criterion (10) serves to stop calculation prior to carrying out the optimal number of iterations, or because the target return r^*_{t+1} does not give the actual rate of return that would be realized given the calculation decisions of the other agents. The advantage of the adaptive sequential CDR, however, is that it gives a complete description of how agents form expectations, unlike the optimal CDR, which endows agents with correct conjectures of the benefits of calculating without modeling how those conjectures are obtained.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

4. ANALYTICAL RESULTS

4.1. Convergence

In this section, we investigate whether the algorithms of Section 3.2 are well behaved, in the sense that, for the stationary environment, the calculation trajectories converge to the steady-state REE from any initial inputs. Convergence may be problematical, however, in that the outputs of the algorithm are potentially unbounded: As expected rates of return r^e_{t+1} approach zero, the corresponding prices $F(r^e_{t+1})$ also approach zero, so that from (3) we have $T(r^e_{t+1}, r^e_{t+2}) \to \infty$. To address this difficulty, we impose a *projection facility* that replaces rates $r^k_{t,s} > \bar{r}$, either inputs or outputs, with the rate $r^k_{t,s} = \bar{r}$, for some given $\bar{r} > \hat{r}$. We may take $\bar{r}$ to be as large as desired.[9] Further, we need to ensure that the algorithms are sufficiently damped; for the propositions that follow, we assume that $\lambda \in (0, \bar{\lambda})$, for a given level $\bar{\lambda} > 0$ that is derived in the Appendix.

The following proposition establishes convergence of EP and MEP for arbitrary initial inputs.

PROPOSITION 1.

(i) *For any β^0_t, infinite iterations of either EP or MEP yield a limit β^∞_t that satisfies $r^\infty_{t,s} = S(r^\infty_{t,s+1})$, $s = 1, \ldots, H$; and*

(ii) *For any given neighborhood of $\hat{r}$, there exists $\underline{H}$ such that, for any β^0_t and $H \geq \underline{H}$, $r^\infty_{t,1}$ lies in the neighborhood.*

Part (i) of the proposition states that EP and MEP converge to a string $(r^\infty_{t,1}, \ldots, r^\infty_{t,H+1})$ that gives a perfect-foresight solution for periods t to $t + H$. Because we have $r^\infty_{t,H+1} = r^0_{t,H+1}$ by definition, the limiting expectations define a perfect-foresight path leading to $r^e_{t+H+1} = r^0_{t,H+1}$. This convergence result utilizes the projection facility and damping assumption to generate calculation trajectories that are *monotonic* for each element of the string. Monotonicity ensures that limits of (4) and (5) exist, and clearly these limits must give perfect-foresight solutions. Part (ii) indicates that, in the stationary case, sufficiently long perfect-foresight paths must have initial elements r^e_{t+1} that are very close to the steady-state REE, as a consequence of the fact that the backward dynamics of the perfect-foresight solution converge to $\hat{r}$.

To illustrate how the PE algorithm can be implemented in the stationary environment, let us take β simply to be a nonnegative scalar giving the agents' expectation of the steady-state rate of return. It then suffices to put $J_s(h_t, \beta) = \beta$, $s = 1, \ldots, H + 1$ (i.e., for any information set, J generates a constant string of length $H + 1$) and to specify any unbiased estimator for step 2. We refer to this specification as the *steady-state PE* algorithm. By analogy to Proposition 1, we have the following proposition.

PROPOSITION 2. *For any β^0_t, infinite iterations of steady-state PE yield a limit β^∞_t that satisfies $\beta^\infty_t = \hat{r}$.*

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

Here, $H = 1$ suffices to generate convergence to the steady-state REE; larger H serves only to generate redundant data.[10]

4.2. Calculation of Steady-State REE

Under the adaptive sequential CDR, the intensity of agents' calculation is related to calculation costs, where the latter are reflected by the parameters n and v_k. In this section, we consider calculation behavior under low calculation costs in the stationary environment. Here, the input functions are assumed to generate constant strings $r^0_{t,s} = r^0$, $s = 1, \ldots, H + 1$; this reflects the notion that agents are seeking to calculate the steady-state REE. If calculation costs are sufficiently low, then the adaptive sequential CDR leads agents to calculate expectations that are arbitrarily close to the steady-state REE, as the following proposition demonstrates.

PROPOSITION 3. *For any given neighborhood of $\hat{r}$ and every H sufficiently large, there exist $\underline{n}$ and $\overline{v}$ such that if $n \geq \underline{n}$ and $v_k < \overline{v}$, $k = 1, \ldots, n$, then calculation ceases under either EP or MEP only if $r^k_{t,1}$ lies in the neighborhood, uniformly over β^0_t.*

Proposition 3 expresses the sense in which expectations under our calculation framework approximate full rationality when calculation costs are low. High levels of n and low levels of v_k mean that agents are able to perform a great many iterations in a period, at a low cost per iteration. As shown in Proposition 1(ii), intense calculation generates expectations that approximate the steady-state REE if the forecast horizon H is sufficiently long. Thus, agents obtain approximately rational expectations when they can look far ahead and carry out many iterations at low cost. The proof of Proposition 3 exploits the monotonicity of convergence under our assumptions, which ensures that agents perceive increases in their utility as they continue to calculate.[11]

We have a parallel result for steady-state PE; as above, it suffices to set $H = 1$.

PROPOSITION 4. *For any given neighborhood of $\hat{r}$, there exist $\underline{n}$ and $\overline{v}$ such that, for $n \geq \underline{n}$ and $v_k < \overline{v}$, $k = 1, \ldots, n$, calculation ceases under steady-state PE only if $r^k_{t,1}$ lies in the neighborhood, uniformly over β^0_t.*[12]

4.3. Nonstationary Environments

We conclude our theoretical analysis by establishing some properties of the adaptive sequential CDR that hold in nonstationary environments. Input functions now are allowed to generate arbitrary strings (subject to the projection facility). The following proposition considers how calculation costs affect the intensity of calculation.

PROPOSITION 5. *If v_k is reduced for every k, then for any given h_t and β^0_t, calculation in period t ceases under any algorithm after a weakly greater number of iterations.*

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

This proposition, which is a trivial implication of the stopping rule (10), shows that the adaptive sequential CDR delivers the correct comparative statics with respect to calculation costs.

We next consider how the results pertaining to EP and MEP for stationary environments extend to nonstationary settings. Proposition 1(ii) continues to hold in the nonstationary case, given that the parameters y_1, G, and d are restricted to a compact set. Further, if the environment is stationary for all periods following t', for some $t' > t$, then the perfect-foresight path emanating backward from the limiting steady state determines a unique REE; as an immediate corollary to Proposition 1(ii), the limiting expectations $r_{t,1}^{\infty}$ can be made arbitrarily close to the corresponding REE values by taking H sufficiently large.[13]

Finally, we derive the following special property of MEP.

PROPOSITION 6. *If $v_k = 0, k = 1, \ldots, n$, then for any given h_t and β_t^0, calculation does not cease under MEP until $k = n$.*

This full-calculation result emerges from the fact that, under MEP, monotonic convergence of $r_{t,1}^k$ ensures that perceived utility must rise with every iteration. As long as the cost per iteration is zero, the adaptive sequential CDR leads agents to make as many calculations as possible.[14] Given that the economy eventually achieves a steady state, it follows that the adaptive sequential CDR yields an arbitrarily close approximation of the REE when H and n are sufficiently large and the values v_k, $k = 1, \ldots, n$ are sufficiently small. We do not obtain a corresponding result for EP because the needed monotonic convergence cannot be guaranteed for arbitrary nonstationary environments.[15]

5. NUMERICAL EXAMPLE

In this section, we analyze a numerical example in which agents are confronted by an anticipated future change of a structural variable. Let the utility function be given by

$$V(c_1) + W(c_2) = \frac{c_1^{1-\sigma}}{1-\sigma} + \frac{\phi c_2^{1-\sigma}}{1-\sigma},$$

where we take $\sigma = \phi = 0.5$. The economy begins in a steady-state REE with $y_1 = 30$ and $G = d = 1$, in which $\hat{r} \cong 1.15$. At the beginning of period 1, the government announces an income tax such that y_1 will be permanently reduced to 20 beginning in period 11, giving a new steady-state REE of $\hat{r} \cong 1.215$.[16] Solution paths for expected rates of return r_{t+1}^e are shown in Figure 1. The announcement produces a unique REE that is depicted by the solid line in Figure 1: The large capital loss in period 10 is fully anticipated, and it leads to a steady decline in returns leading up to period 10. In contrast, expectations under the adaptive rule (11), along with the correct expected returns corresponding to adaptive expectations, are given by the dashed lines. Under adaptive expectations, agents take no account whatsoever of the policy announcement, and they react only following period 11, when the policy change begins to show up in the observed data.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

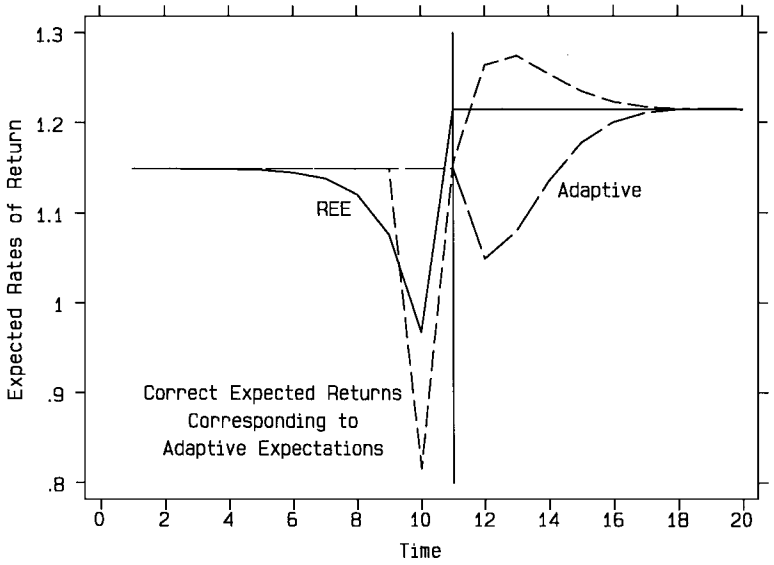


FIGURE 1. Rational and adaptive expectations solutions.

The calculation framework is specified as follows: Let the algorithm be given by MEP, with $H = 5$ and $\lambda = 0.1$. Here, we invoke an alternative projection facility that replaces any values $r^k_{t,s} < 0.1$ with 0.1. For the input function, we use the following simple adaptive rule:

$$r^0_{t,s} = 0.3r_{t-1} + 0.7r^0_{t-1,1} \tag{11}$$

for $s = 1, \ldots, H + 1$. Observe that (11) relies only on past observed rates of return and generates a constant string as the initial input. The adaptive sequential CDR has $n = 60$ and $v_k = v, k = 1, \ldots, n$, i.e., the cost per iteration is constant. We consider the effect of calculation costs on the solution paths by varying the level of v.

Under our calculation framework, the announcement enters agents' information sets h_t for $t \geq 1$, and calculation algorithm incorporates the corresponding future path of $y_{1,t}$. Figure 2 depicts paths of the number of calculations k^C_t generated by the CDR for various levels of v. For sufficiently low calculation costs, intensive calculation commences prior to period 11, as agents seek to adjust their expectations *in anticipation of* the structural change. Further, as calculation costs fall, agents react sooner as well as more intensively. Figure 3 reports the corresponding paths of calculated expected rates of return $r^C_{t,1}$, together with the REE and adaptive paths from Figure 1. As calculation costs fall, expectations shift smoothly from adaptive to REE levels. Note that the shift occurs both *before* and after the policy change.[17]

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

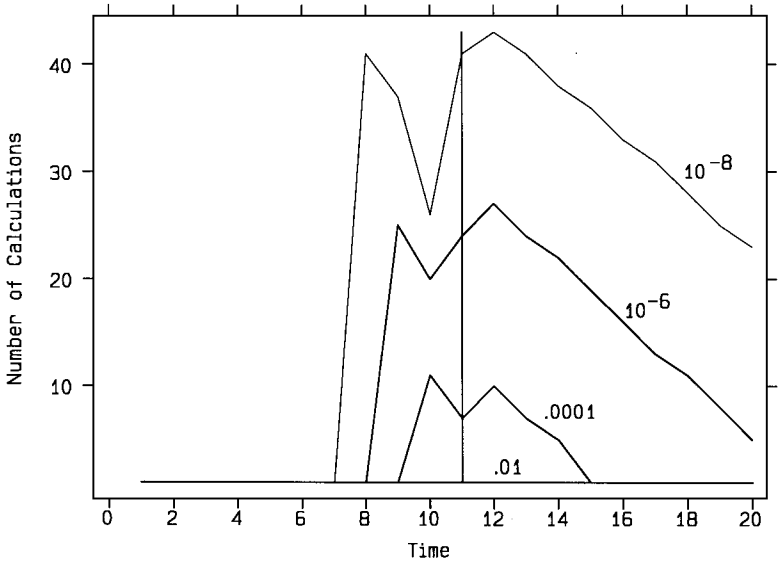


FIGURE 2. Number of calculations per period.

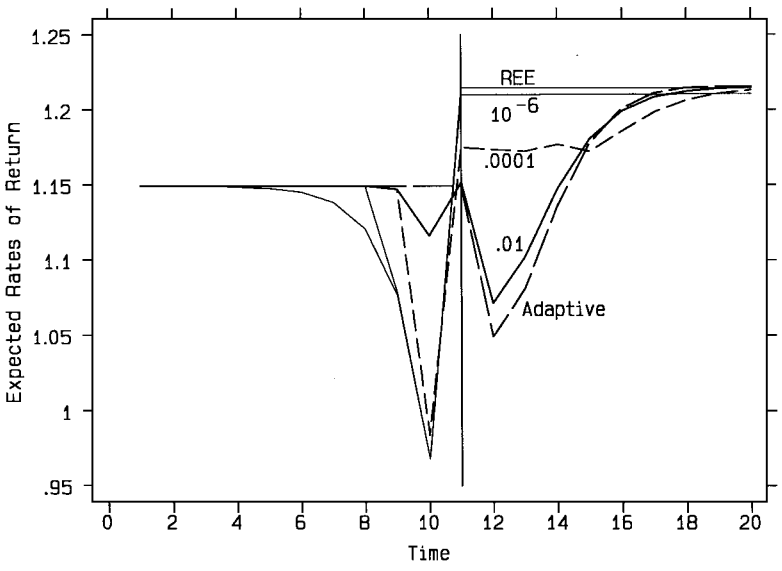


FIGURE 3. Paths of expected returns.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

6. APPLICATIONS

6.1. Asset-Pricing Bubble

The structural model of Section 2 exhibits bubble-like dynamics, in the sense that, for rates of return above the steady-state REE, the perfect-foresight dynamics given by the map $r_{t+1} = S(r_{t+2})$ generate temporarily self-fulfilling upward-sloping paths of returns (and analogously downward-sloping paths for rates of return below the steady state). Temporarily explosive or bubble-like paths may emerge under adaptive expectations, given forecasting rules that allow for the possibility of such phenomena.[18] In this section, we present an example of our calculation framework in which an adaptive component of the input function generates a self-sustaining upward bubble that leads eventually to a crash. As calculation costs fall, agents become better able to anticipate the crash; correspondingly, the bubble becomes smaller, and it breaks sooner.

Consider the example of the preceding section, with $y_1 = 30$ for all periods and the calculation algorithm and CDR unchanged. The input function is given by a combination of the autoregressive forecast rule developed in Section 3.3 and past calculations. Let the forecasted rates determined by (7)–(9) be denoted by $r^A_{t,s}$, $s = 1, \ldots, H+1$, and let the inputs be determined by

$$r^0_{t,s} = 0.2r^A_{t,s} + 0.8r^C_{t-1,s+1}$$

for $s = 1, \ldots, H+1$, where we take $r^C_{t-1,H+2} = r^C_{t-1,H+1}$. The autoregressive rule is initialized by $\alpha_{0,0} \cong -1.15$, $\alpha_{1,0} = 2$ with Z_0 equal to the 2×2 identity matrix and a fixed gain of $\gamma_t = 0.02$. These initial parameter values are consistent with a steady-state REE at $\hat{r} \cong 1.15$, whereas the high value of $\alpha_{1,0}$ ensures that agents are prone to forecast a bubble should higher rates appear in the data. Past calculations are initialized so that agents begin by conjecturing a slight bubble, with $r^C_{0,1} = \hat{r}$ and

$$r^C_{0,s+1} = r^C_{0,s} + 0.001\left(r^C_{0,s}\right)^2$$

for $s = 1, \ldots, H$. A key feature of this input function is that inputs are not restricted to constant strings, but rather can take the form of upward- or downward-sloping bubbles.

Solution paths for expected rates of return are shown in Figure 4. Under adaptive expectations, returns are driven steadily upward by feedback from the adaptive forecasting rule. The bubble breaks in period 15, and a sharp crash ensues.[19] When the calculation algorithm is utilized, there is an initial wave of calculation that counteracts the upward path of the bubble, followed by a second wave that counteracts the crash. These waves occur sooner for lower levels of v. This pattern is seen in Figure 4, where the bubble is smaller and peaks sooner, and the crash is less severe for lower v. At $v = 10^{-7}$, the bubble is barely distinguishable from the steady-state REE.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

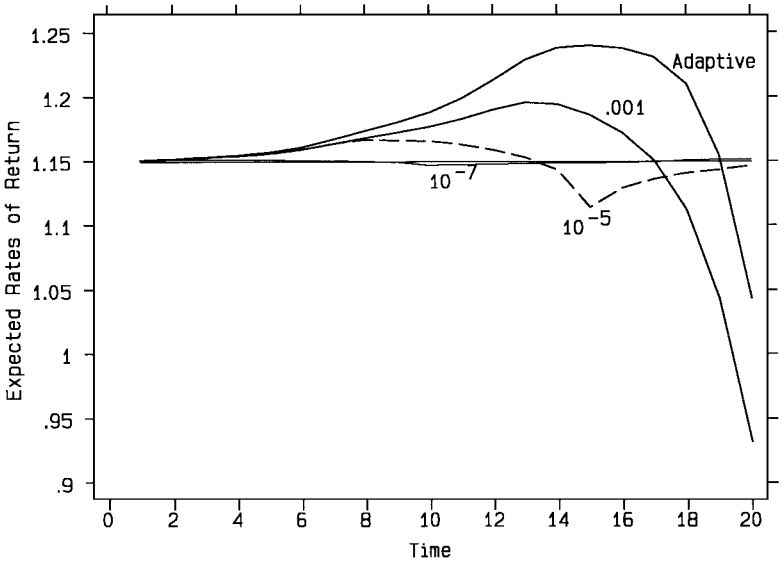


FIGURE 4. Expected return paths for bubble example.

6.2. Deficit Finance and Hyperinflation

To illustrate the relevance of our calculation framework for macroeconomic policy analysis, we modify the structural model of Section 2 to incorporate government spending financed by borrowing. Let G_t denote the stock of government debt in period t. For convenience, we assume that the debt is entirely composed of zero-coupon bonds, so that $d = 0$; thus, G_t can be interpreted as the stock of money. The goods price of money then is given by q_t, and $r_t = q_t/q_{t-1}$ gives the rate of return on holding money, which is the inverse of the inflation rate. Let $b > 0$ denote the amount of government borrowing within a given period. The government budget constraint is given by

$$q_t G_t = b + q_t G_{t-1}.$$

As in Section 2, q_t is determined by market clearing in the asset market, i.e., $q_t G_t = y_1 - c_1(r^e_{t+1})$. The level of real money demand can be written

$$F\left(r^e_{t+1}\right) = y_1 - c_1\left(r^e_{t+1}\right).$$

From the government budget constraint, we obtain $r_t = [F(r^e_{t+1}) - b]/F(r^e_t)$, provided $F(r^e_{t+1}) > b$.[20]

The new feature of the deficit finance model is the possibility that, for large b and small r^e_{t+1}, there does not exist a nonnegative price that clears the asset market and satisfies the government budget constraint. In this case, we assume that the government defaults and that its currency becomes worthless, i.e., q_t falls to zero in the current and all subsequent periods. We refer to this event as *currency*

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

collapse. Taking into account this possibility, we arrive at the following mapping that determines realized rates of return:

$$r_{t+1} = T\left(r^e_{t+1}, r^e_{t+2}\right) \equiv \max\left\{0, \left(F\left(r^e_{t+2}\right) - b\right)/F\left(r^e_{t+1}\right)\right\}, \qquad \textbf{(12)}$$

provided $r_t > 0$, and $r_{t+1} = q_{t+1} = 0$ if $r_t = 0$. Associated with (12) is a perfect-foresight solution $r_{t+1} = S(r_{t+2})$ that gives $r_{t+1} > 0$ if $r_{t+2} > F^{-1}(b)$, and $r_{t+1} = 0$ otherwise.

Multiple monetary steady states, corresponding to multiple strictly positive fixed points of the S map, will exist if b is not too large. All such steady states must satisfy $\hat{r} < 1$, representing positive inflation; further, as long as $c_1'' \geq 0$, there will be, at most, two of them. Let $\hat{r}$ denote the higher of the two steady-state rates of return associated with the lower rate of inflation. The collapse steady state exists for all b. The calculation framework can be applied directly to this case, and the results of Section 4 hold for a suitable choice of the projection facility.[21]

We now consider a policy experiment in which the government temporarily runs an unsustainable deficit. The economy begins in an initial monetary steady state with $b = 1$, which corresponds to $\hat{r} = 0.8$ under the parameterization of Section 5.[22] In period 1, the government announces that the deficit will be increased to $b = 2$ for 10 periods, reverting to $b = 1$ in period 11. The deficit level $b = 2$ is too high to sustain a monetary steady state, and, in this example, the unique REE following the announcement is immediate collapse.

Results are reported in Figure 5. Under adaptive expectations, expected rates of return begin to adjust downward in the direction of collapse, but adjustment

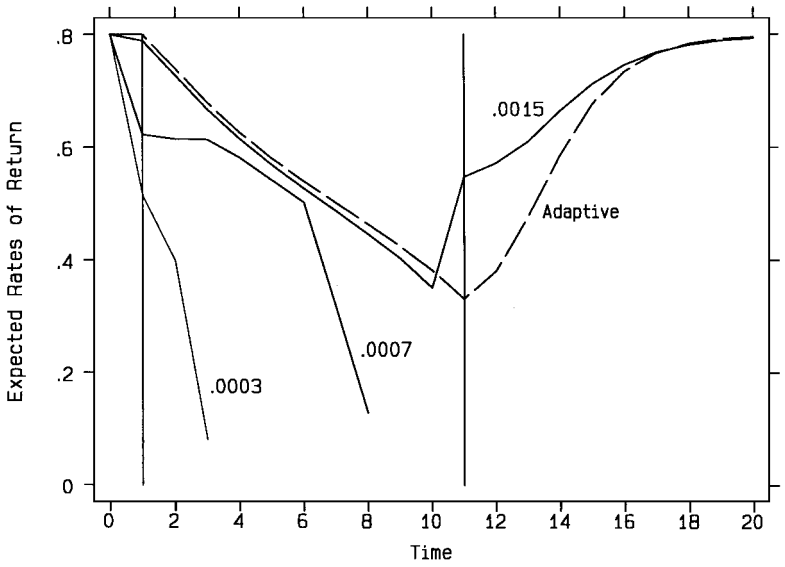


FIGURE 5. Expected return paths for hyperinflation example.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

proceeds slowly enough to prevent collapse from occuring prior to period 11; the policy reversal then leads to restoration of the monetary steady state. A similar outcome arises for calculation costs of $v = 0.0015$. When calculation costs are lowered to 0.0007, however, more rapid adjustment leads to collapse in period 8, and collapse occurs even sooner, in period 3, for calculation costs of 0.0003. This example dramatically illustrates how long-run policy effects, as well as the importance of timely policy responses, can hinge on the cognitive abilities of the agents.

7. CONCLUSIONS

Building on our previous two papers, this paper develops a framework wherein agents form expectations by means of a recursive numerical algorithm. Calculation in each period is initialized by input functions that draw on both observed data and calculations from prior periods. The intensity of calculation, as reflected by the number of iterations of the algorithm in a given period, is determined by a calculation decision rule that utilizes diagnostic calculations to assess the performance of the algorithm in reducing forecast errors.

Our approach has the major advantage of making it possible for expectations to adjust rapidly, should agents perceive sufficient benefits from doing so, as well as to take account of future structural changes that agents anticipate. In essence, we have provided an approach to modeling active cognition that takes passive adaptation as a base, and builds on it a capacity to calculate, in the form of a numerical algorithm. Nevertheless, the framework in this paper remains fully adaptive, in the sense that expectations are derived exclusively on the basis of information that agents possess at the start of a period. Importantly, the framework makes no use of equilibrium notions whereby agents obtain correct beliefs through some unmodeled process.

We present numerical examples that illustrate the performance of our calculation algorithm in applications. A full range of outcomes between adaptive and rational expectations is obtained as calculation costs decline. Notably, we provide a natural resolution of the tension that arises between adaptive and rational approaches to modeling phenomena such as degree of anticipation of future events and expectational bubbles; our framework gives a synthetic treatment of such phenomena, and resolves the tension by reference to the level of calculation costs. Further, we have tried to provide a straightforward and easily implemented setup that can be applied readily in a variety of settings. The framework would appear to be especially useful in the context of monetary models, where calculation costs would serve as a natural source of nominal rigidities.[23] Because the detailed implications of our analysis depend critically on the size of calculation costs, we hope that future research will attempt to estimate these costs, e.g., through empirical studies of the calculation horizons of different economic agents.

It would be natural to apply our framework to the analysis of overlapping-generations models with multiperiod lives. Many important policy issues, such as the reform of social security, involve tax and expenditure changes that extend decades into the future. Auerbach and Kotlikoff (1987) and Laitner (1990), for

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

example, study these kinds of issues in the context of perfect-foresight solutions of an overlapping-generations model with 55-period lifetimes. In settings such as these, solutions that assume nonnegligible calculation costs and limited forecast horizons seem highly salient, both for their realism and potential importance for policy.

One limitation of the framework developed in this paper is that it does not explicitly take account of stochastic elements in the economic structure. In the case of the EP algorithm and its modified version, noise is easily handled simply by treating expectations as certainty equivalents; indeed, this is the approach taken by Fair and Taylor in applying their algorithm to stochastic environments. In practice, however, parameterized expectations appear to give better results in the presence of noise.[24] We plan to investigate more detailed specifications of the parameterized expectations algorithm in the context of a stochastic structural model. A key issue is that the parameterization must exhibit a degree of time-dependence in order to capture anticipations of future events.

A deeper limitation of our framework is that agents' inferential processes are not sufficiently rich to allow them to take full account of the behavior of other agents; in particular, agents cannot give any recognition to the fact that other agents are engaging in expectation calculation, just as they themselves are. In other words, our agents are adaptive robots that cannot think about how the calculations of their fellow robots ought to affect their expectations. This limitation is the price we pay for moving away from the optimal CDR of our earlier work, which posited that agents take *full* account of the calculation behavior of other agents. We believe that a fruitful middle ground can be developed between the extremes of zero and full mutual recognition of rationality. A natural formulation would have agents use their algorithms to calculate the expectations of *other* agents; such a formulation could be applied to analyze problems having a game-theoretic structure, such as optimal monetary policy.[25]

Deeper still, our framework imposes a number of very sharp modeling assumptions in the interests of coherence and tractability. Breaking expectation formation into an adaptation stage, in which data are incorporated costlessly into an input function, followed by a calculation stage that employs a single prespecified algorithm, is only one possible specification from among a much richer set of possibilities. More generally, expectation formation could entail active efforts to obtain data from the environment, interspersed with data processing using various calculation algorithms. Further, agents may not have at hand an accurate structural model of the economy for use in calculation, as we have assumed, but instead may face the problem of estimating a structural model on top of the problem of calculating expectations. Our adaptive sequential CDR also gives a very simple version of a diagnostic decision procedure that, in principle, could be much more involved and potentially could give better decisions on average. However the framework is elaborated, our main point remains that expectation formation is most usefully regarded as emerging from agents' attempts to balance benefits and costs, just as with any other economic activity.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

NOTES

1. Papers that exemplify the passive adaptation approach include Bray (1982), Evans and Honkapohja (1994, 1995), Lucas (1986), Marcet and Sargent (1989a, b), and Woodford (1990). See Sargent (1993) for an extensive survey and a discussion of applications to macroeconomics.

2. Although this approach to modeling calculation decisions has the advantage of linking calculation decisions to the benefits and costs of calculation, it can be criticized on the same grounds as the rational-expectations hypothesis itself, i.e., it does not explain how agents acquire their beliefs about the forecast errors obtaining from their calculation decisions.

3. The algorithms are restricted by imposing a projection facility that rules out extremely large values of expected rates of return, along with suitably heavy damping of the expectation revision on each iteration. We extend the approximation result to a large class of nonstationary environments for one of the versions of the Fair-Taylor algorithm.

4. This type of argument for ruling out rational bubbles was used by Diba and Grossman (1988).

5. To simplify the framework, we have specified that agents cannot take actions to expand their information sets, e.g., by purchasing new information. Expanding the framework to incorporate active observation together with calculation is conceptually straightforward, however. Further discussion of issues in specifying the calculation framework is given in the Conclusions.

6. Let $c_{1,t}(r^e_{t+1})$ denote the solution to (1) for $y_1 = y_{1,t}$, and let $F_t(r^e_{t+1})$ denote the price function determined by (2) with $y_{1,t}$ and $c_{1,t}(r^e_{t+1})$. Then, $T_t(r^e_{t+1}, r^e_{t+2})$ is given by (3) with $F_t(r^e_{t+1})$ and $F_{t+1}(r^e_{t+2})$. We will consider nonstationary paths of y_1 in some of the numerical examples of Section 5.

7. Sargent (1993) surveys a broad class of adaptive learning algorithms drawn from the econometrics and artificial intelligence literatures that may serve as input functions.

8. As a simple alternative to the adaptive sequential CDR, agents may perform multiple iterations in step 1 before making the diagnostic calculation, thus giving a *nonsequential* CDR. Broader issues in specifying the CDR are discussed in the Conclusions.

9. We also replace initial inputs $r^0_{t,s} < 0$, which may emerge from an adaptive forecasting rule, with $r^0_{t,s} = 0$. In the numerical simulations of Sections 5 and 6, we use an alternative projection facility that replaces rates $r^0_{t,s} < \underline{r}$ with $r^0_{t,s} = \underline{r}$, for given $\underline{r} \in (0, \hat{r})$. Our theoretical results continue to hold under the latter specification.

10. Richer specifications of J become necessary when stochastic elements are introduced into the environment. See the Conclusions for further discussion of the use of PE in our framework.

11. Proposition 3 establishes upper bounds on calculation costs for given H that lead agents to approximate the REE. If H is increased, the bounds $\underline{n}$ and $\overline{v}$ may have to be tightened, because a larger number of iterations may be needed to bring $r^k_{t,1}$ into the neighborhood of $\hat{r}$.

12. We omit the proof of Proposition 4, because it is very close to that of Proposition 3.

13. Here we exploit Fair and Taylor's basic insight that REE can be calculated by positing the presence of a limiting steady state.

14. Proposition 6 provides a rationale for the use of H and n as an alternative measure of calculation costs when MEP is used, where the CDR would specify simply that agents perform n iterations per period at a forecast horizon of H.

15. The properties of PE for nonstationary environments depend on additional details of the parameterization J, which we do not take up in this paper.

16. We assume that household utility is additively separable in the utility of private consumption and the utility of government consumption.

17. Similar findings emerge if agents are confronted by a surprise permanent change in y_1: if y_1 shifts unexpectedly to 20 at the beginning of period 1, then expectations adjust more rapidly toward the new steady state as v falls. For very small v, we obtain an essentially fully rational outcome, in which expectations jump at once to the REE.

18. For example, see Brock and Hommes (1996).

19. In later periods, this initial crash is followed by a very sharp spike, eliciting an abrupt downward shift in $\alpha_{1,t}$. Rates of return then follow a cyclical path that converges slowly to the steady-state REE.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

20. This constitutes an overlapping-generations version of the Cagan-Fisher hyperinflation model; see Evans and Ramey (1995) for further discussion and references.

21. It is straightforward to verify that results of Section 4 hold when the projection facility restricts inputs and outputs to $[\underline{r}, \bar{r}]$, where $\underline{r}$ is chosen between the two monetary steady-state returns. In particular, the $\hat{r}$ steady state is stable under our algorithms, and agents will choose to calculate it when calculation costs are sufficiently small, for the indicated projection facility. Further, our propositions can be modified to establish stability of the collapse steady state for an appropriately chosen projection facility, whereas the lower of the two monetary steady states is unstable.

22. In this example, we lower the value of λ from 0.1 to 0.05, whereas the other parameter values are maintained. Here the projection facility simply replaces values $r < 0.1$ with 0.1 and because $0.1 < F^{-1}(1)$, it is possible for agents to calculate either of the two stable steady states.

23. Evans and Ramey (1992) illustrates how nominal rigidities arising from calculation costs give rise to new Keynesian phenomena, such as long-run nonneutrality and hysteresis in monetary policy.

24. See Taylor and Uhlig (1990) and the papers that follow.

25. This approach to applying our framework in game-theoretic settings is discussed further by Evans and Ramey (1988). See Selten (1991) for an alternative approach having some related features.

REFERENCES

Auerbach, A.J. & L.J. Kotlikoff (1987) *Dynamic Fiscal Policy*. Cambridge, England: Cambridge University Press.

Bray, M.M. (1982) Learning, estimation and stability of rational expectations. *Journal of Economic Theory* 26, 318–339.

Brock, W.A. & C.H. Hommes (1996) Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model. Mimeo, University of Wisconsin and Tinbergen Institute.

Conlisk, J. (1988) Optimization cost. *Journal of Economic Organization and Behavior* 9, 213–228.

Conlisk, J. (1996) Why bounded rationality? *Journal of Economic Literature* 34, 669–700.

de Palma, A., G.M. Myers & Y.Y. Papageorgiou (1994) Rational choice under an imperfect ability to choose. *American Economic Review* 84, 419–440.

Diba, B.T. & H.I. Grossman (1988) The theory of rational bubbles in stock prices. *Economic Journal* 98, 746–754.

Evans, G.W. & S. Honkapohja (1994) Learning, convergence, and stability with multiple rational expectations equilibria. *European Economic Review* 38, 1071–1098.

Evans, G.W. & S. Honkapohja (1995) Local convergence of recursive learning to steady states and cycles in stochastic nonlinear models. *Econometrica* 63, 195–206.

Evans, G.W. & G. Ramey (1988) Calculation Equilibria. Institute for Mathematical Studies in the Social Sciences technical report 530, Stanford University.

Evans, G.W. & G. Ramey (1992) Expectation calculation and macroeconomic dynamics. *American Economic Review* 82, 207–224.

Evans, G.W. & G. Ramey (1995) Expectation calculation, hyperinflation and currency collapse. In H. Dixon & N. Rankin (eds.), *The New Macroeconomics*: *Imperfect Markets and Policy Effectiveness*, pp. 307–336. Cambridge, England: Cambridge University Press.

Fair, R.C. & J.B. Taylor (1983) Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models. *Econometrica* 51, 1169–1185.

Laitner, J. (1990) Tax changes and phase diagrams for an overlapping generations model. *Journal of Political Economy* 98, 193–220.

Lucas, R.E., Jr. (1986) Adaptive behavior and economic theory. *Journal of Business* 59, S401–S426.

Marcet, A. (1991) Solving Non-Linear Models by Parameterizing Expectations. Mimeo, Carnegie Mellon University.

Marcet, A. & T.J. Sargent (1989a) Convergence of least squares learning mechanisms in self referential linear stochastic models. *Journal of Economic Theory* 48, 337–368.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

Marcet, A. & T.J. Sargent (1989b) Convergence of least squares learning in environments with hidden state variables and private information. *Journal of Political Economy* 97, 1306–1322.
Sargent, T.J. (1993) *Bounded Rationality in Macroeconomics*. Oxford, England: Clarendon Press.
Selten, R. (1991) Anticipatory learning in two-person games. In R. Selton (ed.), *Game Equilibrium Models, Volume 1*: *Evolution and Game Dynamics*. Berlin: Springer-Verlag.
Taylor, J.B. & H. Uhlig (1990) Solving nonlinear stochastic growth models: A comparison of alternative solution methods. *Journal of Business and Economic Statistics* 8, 1–17.
Woodford, M. (1990) Learning to believe in sunspots. *Econometrica* 58, 277–307.

APPENDIX

The subsequent proofs will make use of the following three lemmas.

LEMMA A.1.

$$\operatorname{sign}\left[r^e_{t+1} - S\left(r^e_{t+2}\right)\right] = \operatorname{sign}\left[S\left(r^e_{t+2}\right) - T\left(r^e_{t+1}, r^e_{t+2}\right)\right].$$

Proof. $r^e_{t+1} > S\left(r^e_{t+2}\right)$ implies that

$$T\left(r^e_{t+1}, r^e_{t+2}\right) = \frac{d + F\left(r^e_{t+2}\right)}{F\left(r^e_{t+1}\right)} < \frac{d + F\left(r^e_{t+2}\right)}{F\left(S\left(r^e_{t+2}\right)\right)} = S\left(r^e_{t+2}\right).$$

Clearly, the reverse implication holds as well. A similar implication holds for $r^e_{t+1} > S(r^e_{t+2})$. ■

LEMMA A.2. *For any $\bar{r} > \hat{r}$, there exists $\bar{\lambda} > 0$ such that, for all $r^e_{t+1}, r^e_{t+2} \leq \bar{r}$, and $\lambda < \bar{\lambda}$, we have*

$$\operatorname{sign}\left[\lambda T\left(r^e_{t+1}, r^e_{t+2}\right) + (1-\lambda) r^e_{t+1} - S\left(r^e_{t+2}\right)\right] = \operatorname{sign}\left[r^e_{t+1} - S\left(r^e_{t+2}\right)\right].$$

Proof. Put $r = r^e_{t+1}, x = r^e_{t+2}$. For (r, x) such that $T(r, x) \neq r$, put

$$\lambda(r, x) = \frac{S(x) - r}{T(r, x) - r}. \tag{A.1}$$

Lemma A.1 implies that $\lambda(r, x) \in (0, 1)$. Let $\bar{\lambda}$ denote the infimum of $\lambda(r, x)$ over $(r, x) \in [0, \bar{r}]^2$, $T(r, x) \neq 0$, and suppose that $\bar{\lambda} = 0$. Then there exists a sequence $\{(r_n, x_n)\}_{n=1}^{\infty} \subset [0, \bar{r}]$ such that $\lambda(r_n, x_n) \to 0$, and further, we may take the sequence to have a limit (r', x'). We have

$$\frac{S(x_n) - r_n}{T(r_n, x_n) - r_n} = 1 - \frac{T(r_n, x_n) - T(S(r_n), x_n)}{T(r_n, x_n) - r_n} = 1 + T_1(\tilde{r}_n, x_n) \frac{S(x_n) - r_n}{T(r_n, x_n) - r_n}, \tag{A.2}$$

where T_1 denotes the derivative of T with respect to its first argument, and $\tilde{r}_n$ lies between r_n and $S(x_n)$. Next note that

$$T_1(\tilde{r}_n, x_n) = \frac{-(d + F(x_n))}{F(\tilde{r}_n)^2} F'(\tilde{r}_n) \to \frac{-(d + F(x'))}{F(r')^2} F'(r').$$

Let us show that $F(r') > 0$. First, if $T(r', x') \neq r'$, then $\lambda(r, x)$ is defined at (r', x'), and we then have $\lambda(r', x') = 0$ because $\lambda(r, x)$ is continuous wherever it is defined, which is

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

a contradiction. Thus, $T(r', x') = r'$, and from (3) we have that $F(r') = 0$ would imply $r' = \infty$, which contradicts $r' < \bar{r}$.

Thus, $T_1(\tilde{r}_n, x_n)$ approaches a limit $T_1(r', x')$. Combining (A.1) and (A.2) gives

$$\lambda(r_n, x_n) = \frac{1}{1 - T_1(\tilde{r}_n, x_n)} \to \frac{1}{1 - T_1(r', x')} > 0,$$

which contradicts $\lambda(r_n, x_n) \to 0$. This proves $\bar{\lambda} > 0$.

Suppose that $\lambda < \bar{\lambda}$ and $r > S(x)$. Then, Lemma A.1 gives $T(r, x) < r$, and we have

$$\lambda < \bar{\lambda} < \frac{S(x) - r}{T(r, x) - r}$$

from which the result follows. A similar argument applies for $r < S(x)$. ■

LEMMA A.3. *Suppose that* $\operatorname{sign}(r^0_{t,s} - S(r^0_{t,s+1}))$ *is the same for all* $s = 1, \ldots, H$, *and that* $\lambda < \bar{\lambda}$. *Then, under EP, for all* k *and* $s = 1, \ldots, H$,

$$\operatorname{sign}\left(r^k_{t,s} - S\left(r^{k-1}_{t,s+1}\right)\right) = \operatorname{sign}\left(r^{k-1}_{t,s} - r^k_{t,s}\right) = \operatorname{sign}\left(r^0_{t,s} - S\left(r^0_{t,s+1}\right)\right).$$

The same holds for MEP, except that $r^{k-1}_{t,s+1}$ is replaced by $r^k_{t,s+1}$.

Proof. First consider EP. Suppose $\operatorname{sign}(r^0_{t,s} - S(r^0_{t,s+1})) > 0$. Using Lemmas A.1 and A.2 we have

$$r^0_{t,s} > r^1_{t,s} > S\left(r^0_{t,s+1}\right)$$

for $s = 1, \ldots, H$, where the first inequality follows from $r^0_{t,s} > T(r^0_{t,s}, r^0_{t,s+1})$. Now suppose that for $k = 1, \ldots, k' - 1$ we have

$$r^{k-1}_{t,s} > r^k_{t,s} > S\left(r^{k-1}_{t,s+1}\right). \tag{A.3}$$

Then $r^{k'-1}_{t,s+1} \le r^{k'-2}_{t,s+1}$, $s = 1, \ldots, H$ and the induction hypothesis imply that

$$r^{k'-1}_{t,s} > S\left(r^{k'-2}_{t,s+1}\right) \ge S\left(r^{k'-1}_{t,s+1}\right)$$

for $s = 1, \ldots, H$. Using Lemmas A.1 and A.2, it follows that equation (A.3) holds for $k = k'$, and by induction for all k. A similar argument applies for $\operatorname{sign}(r^0_{t,s} - S(r^0_{t,s+1})) \le 0$.

Now consider MEP. Suppose $\operatorname{sign}(r^0_{t,s} - S(r^0_{t,s+1})) > 0$. We have

$$r^0_{t,H} > r^1_{t,H} > S\left(r^0_{t,H+1}\right) = S\left(r^1_{t,H+1}\right).$$

Suppose that for $s = s' + 1, \ldots, H$ we have

$$r^0_{t,s} > r^1_{t,s} > S\left(r^1_{t,s+1}\right). \tag{A.4}$$

Because $r^0_{t,s'} > S(r^0_{t,s'+1}) > S(r^1_{t,s'+1})$, we have $r^0_{t,s'} > T(r^0_{t,s'}, r^1_{t,s'+1})$, using Lemma A.1, and so,

$$r^0_{t,s'} > r^1_{t,s'} > S\left(r^1_{t,s'+1}\right)$$

using Lemmas A.1 and A.2. Thus, equation (A.4) holds for $s = 1, \ldots, H$.

Now, suppose that for $k = 1, \ldots, k' - 1$ we have

$$r^{k-1}_{t,s} > r^k_{t,s} > S\left(r^k_{t,s+1}\right) \tag{A.5}$$

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

for $s = 1, \ldots, H$. Because $r_{t,H}^{k'-1} > S(r_{t,H+1}^0)$ implies that $r_{t,H}^{k'-1} > T(r_{t,H}^{k'-1}, r_{t,H+1}^0)$, we have

$$r_{t,H}^{k'-1} > r_{t,H}^{k'} > S\left(r_{t,H+1}^0\right).$$

Suppose that for $s = s' + 1, \ldots, H$ we have

$$r_{t,s}^{k'-1} > r_{t,s}^{k'} > S\left(r_{t,s+1}^{k'}\right).$$

Then, $r_{t,s'}^{k'-1} > S(r_{t,s'+1}^{k'-1}) > S(r_{t,s'+1}^{k'})$, where the first inequality uses equation (A.5); this implies that $r_{t,s'}^{k'-1} > T(r_{t,s'}^{k'-1}, r_{t,s'+1}^{k'})$, and so,

$$r_{t,s'}^{k'-1} > r_{t,s'}^{k'} > S\left(r_{t,s'+1}^{k'}\right).$$

Thus, equation (A.5) holds for every k. A similar argument applies for $\text{sign}(r_{t,s}^0 - S(r_{t,s+1}^0)) \leq 0$. ■

Proof of Proposition 1. First consider EP. Suppose that $r_{t,H}^0 > S(r_{t,H+1}^0)$ for given β_t^0. From Lemma A.1 we have $r_{t,H}^0 > T(r_{t,H}^0, r_{t,H+1}^0)$, so that $r_{t,H}^1 < r_{t,H}^0$. From Lemma A.2, we have $r_{t,H}^1 > S(r_{t,H+1}^0)$. By induction, $r_{t,H}^{k-1} > S(r_{t,H+1}^{k-1}) = S(r_{t,H+1}^0)$ implies $r_{t,H}^{k-1} > r_{t,H}^k > S(r_{t,H+1}^0)$; thus, $r_{t,H}^\infty$ exists, and clearly $r_{t,H}^\infty = S(r_{t,H+1}^0)$. A similar argument applies for $r_{t,H}^0 < S(r_{t,H+1}^0)$, while $r_{r,H}^0 = S(r_{t,H+1}^0)$ implies $r_{t,H}^k = S(r_{t,H+1}^0)$ for all k.

Now suppose that $r_{t,s}^\infty$ exists for $s = s' + 1, \ldots, H$, and $r_{t,s}^\infty = S(r_{t,s+1}^\infty)$. Choose any $\delta > 0$, and fix K so that $k > K$ implies that

$$\left|S\left(r_{t,s'+1}^{k-1}\right) - S\left(r_{t,s'+1}^\infty\right)\right| < \delta.$$

Let N denote the δ-ball around $S(r_{t,s'+1}^\infty)$. If $k > K$ and $r_{t,s'+1}^{k-1} < S(r_{t,s'+1}^{k-1})$, then Lemma A.2 gives

$$r_{t,s'}^{k-1} < r_{t,s'}^k < S\left(r_{t,s'+1}^{k-1}\right).$$

Because $S(r_{t,s'+1}^{k-1}) \in N$, we have that $r_{t,s'}^{k-1} \in N$ implies that $r_{t,s'}^k \in N$, whereas $r_{t,s'}^{k-1} \notin N$ implies that $r_{t,s'}^k$ lies either below or within N. If $r_{t,s'}^k$ lies below N, then we have $r_{t,s'}^k < S(r_{t,s'+1}^k)$ because $S(r_{t,s'+1}^k) \in N$. Thus, for $k > K$ and $r_{t,s'+1}^{k-1} < S(r_{t,s'+1}^{k-1})$, the sequence $r_{t,s'}^{k-1}, r_{t,s'}^k, \ldots$ either lies in N or increases monotonically toward N. In the latter case, suppose that $r_{t,s'}^k \to r'$ for some r' below N. Because $S(r) < \bar{r}$ for all $r \in [0, \bar{r}]$ [which follows from $\bar{r} > \hat{r}$, which implies that $S(\bar{r}) < \bar{r}$], we have $r_{t,s'}^k < \bar{r}$ for all $k > K$. Thus, for each k we have

$$r_{t,s'}^k - r_{t,s'}^{k-1} = \lambda\left(T\left(r_{t,s'}^{k-1}, r_{t,s'+1}^{k-1}\right) - r_{t,s'}^{k-1}\right).$$

Convergence implies that $r_{t,s'}^k - r_{t,s'}^{k-1} \to 0$, and so,

$$T\left(r_{t,s'}^{k-1}, r_{t,s'+1}^{k-1}\right) - r_{t,s'}^{k-1} \to 0 = T\left(r', r_{t,s'+1}^\infty\right) - r',$$

which implies that $r' = S(r_{t,s'+1}^\infty)$. Thus, the sequence must lie in N. Symmetric arguments hold for $r_{t,s'}^{k-1} > S(r_{t,s'+1}^{k-1})$, and because δ is arbitrary, we have $r_{t,s'}^k \to S(r_{t,s'+1}^\infty)$. The same proof applies for MEP, except that $r_{t,s'+1}^{k-1}$ in the above expressions is replaced by $r_{t,s'+1}^k$, and $r_{t,s'+1}^k$ is replaced by $r_{t,s'+1}^{k+1}$. This proves part (i).

Given the stationary environment, part (i) implies that $r_{t,1}^\infty = S^H(r_{t,H+1}^0)$. It is easily checked that $S'(\hat{r}) < 1$; thus, because $S(r) \neq r$ for any $r \neq \hat{r}$, we have that, for any given $\epsilon > 0$, $S(r) - r$ is bounded above zero for $r \leq \hat{r} - \epsilon$, and bounded below zero for $r > \hat{r} + \epsilon$.

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

It follows that there is some $\underline{H}$ such that $S^H(r) \in (\hat{r}-\epsilon, \hat{r}+\epsilon)$ for $H \geq \underline{H}$, uniformly over $r \in [0, \bar{r}]$. This proves part (ii). ■

Proof of Proposition 2. The parameterization for steady-state PE gives $r_{t,s}^k = \beta_t^k$, $s = 1, \ldots, H+1$, and so, step 1 gives

$$\hat{r}_s^k = T\left(\beta_t^{k-1}, \beta_t^{k-1}\right).$$

Unbiasedness of the estimator implies that $\hat{\beta}_t^k = T(\beta_t^{k-1}, \beta_t^{k-1})$, so that step 2 yields

$$\beta_t^k = \lambda T\left(\beta_t^{k-1}, \beta_t^{k-1}\right) + (1-\lambda)\beta_t^{k-1}. \tag{A.6}$$

For $\beta_t^0 = \hat{r}$, equation (A.6) gives $\beta_t^k = \hat{r}$ for all k. Suppose that $\beta_t^0 > \hat{r}$. Then, we have $\beta_t^0 > S(\beta_t^0)$, and Lemma A.1 gives $\beta_t^0 > T(\beta_t^0, \beta_t^0)$. From equation (A.6) we then have $\beta_t^1 < \beta_t^0$, whereas using Lemma A.2 gives $\beta_t^1 > S(\beta_t^0) > S(\beta_t^1)$; the latter implies that $\beta_t^1 > \hat{r}$. Proceeding inductively, we have $\beta_t^{k-1} > \beta_t^k > \hat{r}$, and there exists a limit $\beta_t^\infty \geq \hat{r}$. Using equation (A.6) and continuity of the T-mapping, we have $\beta_t^\infty = T(\beta_t^\infty, \beta_t^\infty) = \hat{r}$. A similar argument applies for $\beta_t^0 < S(\beta_t^0)$. ■

Proof of Proposition 3. Let $r_{t,1}^k(r^0)$ and $r_{t,1}^\infty(r^0)$ denote the values of $r_{t,1}^k$ and $r_{t,1}^\infty$ when $r_{t,s}^0 = r^0$, $s = 1, \ldots, H+1$. These functions are continuous in r^0 [where continuity of $r_{t,1}^\infty(r^0)$ follows from $r_{t,1}^\infty(r^0) = S^H(r^0)$, as implied by Proposition 1]. Choose any $\delta > 0$; Proposition 1(ii) allows us to fix H so that, for all $r^0 \in [0, \hat{r}]$,

$$\left| r_{t,1}^\infty(r^0) - \hat{r} \right| < (\delta/2). \tag{A.7}$$

For each $r^0 \in [0, \bar{r}]$, choose $k(r^0)$ to satisfy

$$\left| r_{t,1}^{k(r^0)}(r^0) - r_{t,1}^\infty(r^0) \right| < (\delta/2).$$

Continuity of $r_{t,1}^{k(r^0)}(r)$ and $r_{t,1}^\infty(r)$ implies that we can find a neighborhood $N(r^0)$ of r^0 such that, for all $r \in N(r^0)$,

$$\left| r_{t,1}^{k(r^0)}(r) - r_{t,1}^\infty(r) \right| < (\delta/2).$$

Because the neighborhoods $N(r^0)$ form an open cover of $[0, \bar{r}]$, we can find a finite subcover $N(r_i^0)$, $i = 1, \ldots, I$. Put

$$\underline{n} = \max_i k\left(r_i^0\right).$$

For any $r^0 \in [0, \bar{r}]$, we have $r^0 \in N(r_i^0)$ for some i, whence

$$\left| r_{t,1}^{k(r_i^0)}(r^0) - r_{t,1}^\infty(r^0) \right| < (\delta/2). \tag{A.8}$$

If $r^0 > S(r^0)$, then using Lemma A.3 we have, for any $n \geq \underline{n}$,

$$r_{t,1}^{k(r_i^0)}(r^0) > r_{t,1}^n(r^0) > r_{t,1}^\infty(r^0). \tag{A.9}$$

For $r^0 < S(r^0)$, equation (A.9) holds with reversed inequalities, whereas $r^0 = S(r^0)$ implies that $r_{t,1}^k = r_{t,1}^\infty = r^0$ for all k. Combining equations (A.8) and (A.9), it follows that, for all r^0

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

and all $n \geq \underline{n}$,

$$\left| r_{t,1}^{n}(r^0) - r_{t,1}^{\infty}(r^0) \right| < (\delta/2). \tag{A.10}$$

Combining equations (A.7) and (A.10) gives, for all r^0 and all $n \geq \underline{n}$,

$$\left| r_{t,1}^{n}(r^0) - \hat{r} \right| < \delta. \tag{A.11}$$

Next, put $N = (\hat{r} - \delta, \hat{r} + \delta)$ and

$$\Delta = \inf_{r^0 \notin N, k \leq \underline{n}} \left| r_{t,1}^{k-1}(r^0) - S\left(r_{t,2}^{k-1}(r^0)\right) \right|.$$

Suppose $\Delta = 0$. Then, there exists a sequence $\{r_j^0, k_j\}_{j=1}^{\infty}$ with

$$\left| r_{t,1}^{k_j - 1}\left(r_j^0\right) - S\left(r_{t,2}^{k_j - 1}\left(r_j^0\right)\right) \right| \to 0.$$

Further, $r_j^0 \in [0, \bar{r}] \sim N$ and $k_j \leq \underline{n}$, and so the sequence has a limit point $(\tilde{r}^0, \tilde{k})$; further, we may take a subsequence with $r_j^0 \to \tilde{r}^0$ and $k_j = \tilde{k}$ for all j. Continuity of $r_{t,1}^{\tilde{k}-1}(r^0)$ and $r_{t,2}^{\tilde{k}-1}(r^0)$, as well as of $S(r)$, then gives

$$\left| r_{t,1}^{\tilde{k}-1}\left(\tilde{r}^0\right) - S\left(r_{t,2}^{\tilde{k}-1}\left(\tilde{r}^0\right)\right) \right| = 0.$$

According to Lemma A.3 this can hold only if $\tilde{r}^0 = S(\tilde{r}^0)$, which implies that $\tilde{r}^0 = \hat{r}$. We have a contradiction, because $\tilde{r}^0 \notin N$ is necessary. Thus, $\Delta > 0$ is necessary.

We next consider the utility function U.

$$\frac{\partial}{\partial r_{t,1}^k} U\left(r_{t,1}^k, r_{t+1}^*\right) = \left[V' - W' r_{t+1}^*\right] \frac{\partial c_1}{\partial r_{t,1}^k}.$$

The first-order condition that determines $c_1(r_{t+1}^e)$ implies that $V' - W' r_{t+1}^* = 0$ for $r_{t,1}^k = r_{t+1}^*$. Moreover, at $r_{t,1}^k = r_{t+1}^*$, we have

$$\frac{\partial}{\partial r_{t,1}^k}\left[V' - W' r_{t+1}^*\right] = \left[V'' + W''\left(r_{t+1}^*\right)^2\right] \frac{\partial c_1}{\partial r_{t,1}^k} > 0.$$

Thus, we have $V' - W' r_{t+1}^* < 0$ for $r_{t,1}^k < r_{t+1}^*$ and $V' - W' r_{t+1}^* > 0$ for $r_{t,1}^k > r_{t+1}^*$. It follows that

$$\operatorname{sign}\left[\frac{\partial}{\partial r_{t,1}^k} U\left(r_{t,1}^k, r_{t+1}^*\right)\right] = -\operatorname{sign}\left(r_{t,1}^k - r_{t+1}^*\right). \tag{A.12}$$

Using (A.12), we have

$$\min_{r_{t,1}^k \leq r_{t+1}^* - \Delta} \frac{\partial}{\partial r_{t,1}^k} U\left(r_{t,1}^k, r_{t+1}^*\right) = \nu^+ > 0,$$

$$\max_{r_{t,1}^k \geq r_{t+1}^* + \Delta} \frac{\partial}{\partial r_{t,1}^k} U\left(r_{t,1}^k, r_{t+1}^*\right) = \nu^- < 0.$$

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press

Now consider EP. For $r^0 \geq \hat{r} + \delta$, $k \leq \underline{n}$, we have

$$r_{t,1}^{k-1} - r_{t,1}^k = \lambda\left(r_{t,1}^{k-1} - T\left(r_{t,1}^{k-1}, r_{t,2}^{k-1}\right)\right) > \lambda\left(r_{t,1}^{k-1} - S\left(r_{t,2}^{k-1}\right)\right) \geq \lambda\Delta.$$

Moveover, Lemma A.3 gives $r_{t,1}^k > S(r_{t,2}^k)$, and so,

$$r_{t,1}^k - T\left(r_{t,1}^k, r_{t,2}^k\right) > r_{t,1}^k - S\left(r_{t,2}^k\right) \geq \Delta.$$

Thus,

$$\begin{aligned} U\left(r_{t,1}^k, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) - U\left(r_{t,1}^{k-1}, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) &= -\int_{r_{t,1}^k}^{r_{t,1}^{k-1}} \frac{\partial}{\partial r_{t,1}^k} U\left(r, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) dr \\ &\geq -\left(r_{t,1}^{k-1} - r_{t,1}^k\right)\nu^- \geq -\lambda\Delta\nu^- > 0. \end{aligned} \tag{A.13}$$

Similarly, for $r^0 \leq \hat{r} - \delta$, we have

$$r_{t,1}^{k-1} - r_{t,1}^k < -\lambda\Delta < 0,$$

$$r_{t,1}^k - T\left(r_{t,1}^k, r_{t,2}^k\right) < r_{t,1}^k - S\left(r_{t,2}^k\right) < -\Delta.$$

Thus,

$$U\left(r_{t,1}^k, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) - U\left(r_{t,1}^{k-1}, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) \geq \left(r_{t,1}^k - r_{t,1}^{k-1}\right)\nu^+ \geq \lambda\Delta\nu^+ > 0. \tag{A.14}$$

Put $\overline{v} = \min\{\lambda\Delta\nu^+, -\lambda\Delta\nu^-\}$. It follows that if $r^0 \notin N$, then $v_k \leq \overline{v}$ for $k = 1, \ldots, \underline{n}$ implies that at least $\underline{n}$ iterations will be performed, in which case (A.11) assures $r_{t,1}^k \in N$ once calculation ceases. For $r^0 \in N$, we have $r_{t,1}^k \in N$ for all k, using Lemma A.3.

For MEP, in the case $r^0 \geq \hat{r} + \delta$, $k \leq \underline{n}$ we have, using $r_{t,2}^k \leq r_{t,2}^{k-1}$,

$$r_{t,1}^{k-1} - r_{t,1}^k = \lambda\left(r_{t,1}^{k-1} - T\left(r_{t,1}^{k-1}, r_{t,2}^k\right)\right) > \lambda\left(r_{t,1}^{k-1} - S\left(r_{t,2}^k\right)\right) \geq \lambda\left(r_{t,1}^{k-1} - S\left(r_{t,2}^{k-1}\right)\right) > \lambda\Delta$$

Thus, we have (A.13) as before, and similarly (A.14) continues to hold for $r^0 \leq \hat{r} - \delta$, $k \leq \underline{n}$. Defining $\overline{v}$ as above gives the result for MEP. ■

Proof of Proposition 6. Suppose that after k iterations we have $r_{t,1}^{k-1} \geq S(r_{t,2}^k)$. Using Lemma A.2, we have

$$r_{t,1}^{k-1} \geq r_{t,1}^k \geq S\left(r_{t,2}^k\right).$$

Lemma A.1 gives

$$r_{t,1}^{k-1} \geq r_{t,1}^k \geq T\left(r_{t,1}^k, r_{t,2}^k\right). \tag{A.15}$$

Thus, we have

$$U\left(r_{t,1}^k, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) - U\left(r_{t,1}^{k-1}, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) = \int_{r_{t,1}^{k-1}}^{r_{t,1}^k} \frac{\partial}{\partial r_{t,1}^k} U\left(r, T\left(r_{t,1}^k, r_{t,2}^k\right)\right) dr \geq 0,$$

where the inequality follows from equations (A.12) and (A.15). Combining this with the hypothesis that $v_k = 0$ for $k \leq n$, it follows that equation (10) cannot hold for any $k < n$. A similar argument applies for $r_{t,1}^{k-1} \leq S(r_{t,2}^k)$. ■

https://doi.org/10.1017/S1365100598007020 Published online by Cambridge University Press