Robustness and Pricing with Uncertain Growth

Marco Cagetti
University of Virginia

Lars Peter Hansen
University of Chicago

Thomas Sargent
Stanford University

Noah Williams
Princeton University

We study how decision-makers' concerns about robustness affect prices and quantities in a stochastic growth model. In the model economy, growth rates in technology are altered by infrequent large shocks and continuous small shocks. An investor observes movements in the technology level but cannot perfectly distinguish their sources. Instead the investor solves a signal extraction problem. We depart from most of the macroeconomics and finance literature by presuming that the investor treats the specification of technology evolution as an approximation. To promote a decision rule that is robust to model misspecification, an investor acts as if a malevolent player threatens to perturb the actual data-generating process relative to his approximating model. We study how a concern about robustness alters asset prices. We show that the dynamic evolution of the risk-return trade-off is dominated by movements in the growth-state probabilities and that the evolution of the dividend-price ratio is driven primarily by the capital-technology ratio.

This article shows how decision-makers' concerns about model misspecification can affect prices and quantities in a dynamic economy. We use the familiar stochastic growth model of Brock and Mirman (1972) and Merton (1975) as a laboratory. Technology is specified as a continuous-time hidden Markov model (HMM), inducing investors to make inferences about the growth rate. They form their opinions about the growth rate from current and past observations of technology that are clouded by concurrently evolving small shocks modeled as Brownian motions. We show how investors' desire to make their decision rules robust to misspecification of the evo-

Comments from John Heaton and Grace Tsiang are gratefully acknowledged. Both L. P. Hansen and T. Sargent were supported by the National Science Foundation while completing this project. Address correspondence to Lars Peter Hansen, Department of Economics, University of Chicago, 1126 East 59th St., Chicago, IL 60637, or e-mail: l-hansen@uchicago.edu.

The Review of Financial Studies Special 2002 Vol. 15, No. 2, pp. 363–404
© 2002 The Society for Financial Studies

lution of technology alters security prices and the intertemporal resource allocation.

Following the control theory literature, we formulate the robust decision-making process as a two-player game. For a game-theoretic approach to robust decision making see Basar and Bernhard (1995). For the recursive specification used here see Cagetti et al. (2000). An investor has somehow constructed a model of the technology shock process, but suspects that model to be misspecified. The investor wants to maximize the expected value of a discounted utility function, but doubting his approximating model, is unsure about what probability distribution to use to form mathematical expectations. To make decisions that perform well under a variety of models, the decision maker imagines that a second malevolent agent will draw technology shocks from a model that is distorted relative to his approximating model. The malevolent agent minimizes the decision-maker's objective function by choosing a model from a large set surrounding the approximating model. To represent the idea that the decision maker views his model as a good approximation, we restrict the surrounding models to be close to the approximating model, where closeness is measured by a statistical discrimination criterion of the gap between a distorted model and the approximating model.[1]

The malevolent agent in the decision problem provides an operational way to promote robustness by systematically exploring the types of model misspecification to which a proposed decision rule is especially fragile.[2] As argued by Huber (1981) in his discussion of an optimal robust statistical procedure:

> [A]s we defined robustness to mean insensitivity with regard to small deviations from assumptions, any quantitative measure of robustness must somehow be concerned with the maximum degradation of performance possible for an ϵ-deviation from the assumptions. The *optimally robust* procedure minimizes this degradation and hence will be a minimax procedure of some kind.

In this article we model investor preferences using a penalty approach that lets investors explore deviations from an approximating model of the technology evolution.[3] Formally, we append a term to the discounted expected utility

[1] As is typical with rational expectations models, we do not explore formally how investors arrive at an approximating model, but we do presume that the remaining models that are entertained are difficult to distinguish from the approximating model given historical data. A rational expectations counterpart model, however, would remove model approximation error from consideration.

[2] Economists use a max-min formulation when they use Lagrange multipliers. Lagrange multipliers allow us to convert a constrained maximization problem into an unconstrained maximum-minimum problem. The constraint is imposed by supposing there exists a fictitious *malevolent* agent whose aim it is to punish the original decision maker when the constraint is violated. This device for imposing a constraint is algorithmically convenient. So is our two-agent formulation of robustness.

[3] Although we use ideas from standard robust control theory, we modify them because our approximating models are stochastic. Among the few examples of stochastic robust control models are James (1982), who

that penalizes departures from a reference or approximating model. Penalty methods are common in both the robust control theory and statistics literature. Strictly speaking they imply preferences that are distinct from those that limit exploration only to ϵ-deviations as suggested by Hurber (1981) in the above quote, but the two approaches are related. [for example, see Dupuis, James, and Petersen (1998) and Hansen et al. (2001)]. We use a penalty function that tolerates only perturbations from the approximating model that are difficult to detect statistically. This leads us to use a log-likelihood ratio-based penalty term called relative entropy, appropriately adjusted to account for the HMM structure.[4]

Concern about model misspecification makes investors more cautious and enlarges measured risk premia. It enhances the usual precautionary motive.[5]

Our model economy is a continuous-time stochastic growth economy. The technology shock process has a two-state hidden Markov (HMM) structure. Our quantitative estimates confirm the findings of Hamilton (1989) and others that a two-state HMM model for the post-World War II United States recovers states that measure short recessions and sustained booms. Because it can be difficult to distinguish these two states from the observed technology level, we investigate the consequences of concealing the growth state from the decision maker. Our specification of the technology shock process formally follows Wonham (1964), David (1997), and Veronesi (1999). The latter two articles study pricing in production economies with linear technologies and hidden dividend growth processes.[6] Hidden information gives us a tractable setting to study what happens when the quality of investors' information fluctuates over time. This mechanism alone can alter the time-series evolution of the market risk prices and dividend-price ratios, but it cannot produce large enough risk premia to be empirically plausible. It is for this reason that we turn to robustness as a means of changing asset price predictions.

We decentralize the robust version of the stochastic growth model by computing shadow prices from a robust resource allocation problem. The continuous-time specification lets us use representations of local prices to assemble asset prices for intervals of time. The local prices include both the instantaneous interest rate and the risk price of the Brownian motion increment. Our model contributes an additional component of the factor risk price

---

studies deterministic differential games of robust decision making as small noise limits, and Dupuis, James, and Petersen (1998). A stochastic approximating model is essential for our application to finance. In emphasizing this structure we follow Hansen, Sargent, and Tallarini (1999), Maenhout (1999), and Anderson, Hansen, and Sargent (2000).

[4] Examples of robust control of deterministic systems with partial observations are James, Baras, and Elliott (1994) and James and Baras (1996).

[5] The precautionary motive is already present in models without quadratic preferences.

[6] David (1997) studies a model in which production is linear in the capital stocks with technology shocks that have hidden growth rates. Veronesi (1999) studies a permanent income model with a riskless linear technology. Dividends are modeled as an additional consumption endowment. Hidden information was introduced into asset pricing models by Detemple (1986), who considers a production economy with Gaussian unobserved variables.

that is attributable to a concern about model misspecification. This component also occurs in the continuous-time articles of Anderson, Hansen, and Sargent (2000) and Chen and Epstein (2001), and as an approximation in the discrete-time analysis of Hansen, Sargent, and Tallarini (1999).

Quantitatively, our two-state HMM specification has the following implications for intertemporal prices and allocations. The investor's signal extraction problem makes one state variable be the probability that the current growth rate is low. Another state variable is the ratio of capital to technology. We show that:

- The robust motive for precautionary savings increases the capital stock. This motive can be offset by making investors discount the future more.
- A concern about model misspecification adds a quantitatively important component to the risk-return trade-off as measured by financial econometricians. The component of risk prices due to robustness is particularly sensitive to growth-state probabilities and is largest when, under the approximating model, investors are most unsure of the hidden state.
- A concern about robustness causes price-dividend ratios to drop closer to the level observed in postwar data. In our model economy, these ratios are particularly sensitive to movements in capital-technology ratio. They respond very little to changes in the growth-state probabilities. The actual time-series trajectories for price-earnings ratios differ substantially from those implied by the model.

The first finding extends a result of Hansen, Sargent, and Tallarini (1999) to a nonlinear economy. Without prior information about the subjective discount factor, decision-makers' concerns about robustness can't be detected from macroeconomic quantities alone. The other two findings have important counterparts in a corresponding HMM rational expectation economy in which investors only care about risk. In such economies, market risk prices respond primarily to changes in growth-state probabilities, while dividend-price ratios are driven primarily by movements in the capital-technology ratio.

The article is structured as follows. Section 1 presents the economic environment. Section 2 describes the information structure and the signal extraction problem. Section 3 describes model distortions and measures of model misspecification. Section 4 presents differential equations for value functions that characterize equilibria of the hidden information games. Section 5 discusses implications for time series of capital stocks. Section 6 uses the link between statistical detection and robustness to restrict the degree of robustness in the asset calculations. Section 7 shows how risk-return trade-offs change over time. Section 8 displays the implied dividend-price ratios.

1. The Economy

We use a continuous-time formulation of a Brock and Mirman (1972) economy with production, capital accumulation, and stochastic productivity growth. There are two types of technology shocks: Brownian motion increments, and infrequent changes in the drifts of the Brownian motion modeled as a jump process. Investors observe productivity levels but the drift is hidden. The technology process is thus a special case of an HMM, confronting investors with a signal extraction problem. Current and past data must be used to make inferences about technological growth.

We use this model to study the precautionary motive for savings induced by a concern about robustness; the evolution of the measured market price of "risk"; and the evolution of price-dividend ratios.

1.1 Previous literature

The quantitative component of our investigation is designed to show how robustness alters the implications of the simple growth model familiar to economists. In the absence of robustness, the empirical implications of this model for consumption and investment are defective [e.g., see Watson (1993)] and the implied return to capital shows very little variation relative, for instance, to value-weighted returns on equity [e.g., see Cochrane (1991) and Rouwenhorst (1995)]. The absence of return variability is even more stark in the continuous-time embedding of this model. As noted by Merton (1975), the return to capital becomes locally riskless.

One remedy is to make capital locally risky. While this will enhance return variability, it may also result in excessive volatility in aggregate quantities. In addition, we might follow Boldrin, Christiano, and Fisher (2001) and others by introducing additional technological frictions and temporal nonseparabilities in preferences. Instead of mixing robustness with these other ways to complicate the short-run dynamics, we study the role of robust decision making in a simpler framework.

When looking at the asset pricing implications, we will be less ambitious than Boldrin, Christiano, and Fisher (2001) and Hansen and Singleton (1983),[7] and will study only the local or instantaneous risk-return relation and the time-series behavior for dividend-price ratios. Even the risk-return relation looks puzzling for a model without robust decision makers because the implied market price is too small to be plausible from the vantage point of aggregate models [Hansen and Jagannathan (1991) and Cochrane and Hansen (1992)]. As in Hansen, Sargent, and Tallarini (1999), Maenhout (1999), Anderson, Hansen, and Sargent (2000), and Chen and Epstein (2001) we explore the effects of a concern about model uncertainty on the measured risk premium in security market returns. We add to this literature by looking at the time-series variation both of risk prices and of model uncertainty prices. We show how disguising mean growth rates from investors can alter the time-series properties of the risk premia.

[7] We look at only a subset of the restrictions that they studied.

To study dividend-price ratios we will eventually posit an exogenous dividend claim that is distinct from the marginal product of capital. In this we imitate David and Veronesi (1999) and Veronesi (2000) except that we have an additional state variable (capital) and also explore implications for robustness.

1.2 Technology

We assume a Cobb–Douglas production function

$$f(K, L) = K^{\alpha}(YL)^{1-\alpha},$$

where K is the capital stock, L is the labor supply, and Y is the labor-augmenting technology parameter. For simplicity, we fix the total labor supply L at 1. Y evolves exogenously according to the continuous-time process

$$dy_t = s_t \cdot \hat{\kappa}\, dt + \sigma_y\, dB_t, \tag{1}$$

where B is a standard Brownian motion, $y = \log Y$, and s evolves according to a finite-state Markov chain. It can assume n possible values, $U_1, U_2, \ldots, U_n$, where U_j is a vector containing 1 in position j and zero everywhere else. $\hat{\kappa}$ is an n-dimensional vector that contains all possible values of the mean growth rate of the technology shock; $s_j \cdot \hat{\kappa}$ is therefore the growth rate in state j. The model of the technology shock can be viewed as a continuous-time embedding of the regime-shift models of Baum and Petrie (1966), Sclove (1983) and Hamilton (1989).[8]

Let δ be the depreciation rate of capital. The evolution equation for capital is given by

$$dK_t = \left[(Y_t)^{1-\alpha}(K_t)^{\alpha} - C_t - \delta K_t\right] dt, \tag{2}$$

where C_t is the instantaneous consumption flow. By construction, capital is locally predictable.

The technological process has a unit root in logarithms and is therefore nonstationary. As we show later, the ratio of capital to effective labor, $k_t = K_t/Y_t$, and that of consumption to effective labor, $c_t = C_t/Y_t$, are stationary. We will therefore represent the problem in terms of the variables k_t and y_t.

[8] An important qualification is that volatility is independent of the Markov chain state. Thus we are ruling a process with high volatility in low growth or recession states. Baum and Petrie (1966) and Bonomo and Garcia (1996) allow for the state s_t to alter volatility and find it to be empirically plausible for postwar output data and century-long consumption data. We preclude this dependence in order that our state s_t remain difficult to detect from high-frequency data. Volatility changes are revealed by continuous data records.

Applying Ito's lemma, we get

$$dk_t = \mu_k(c, k, s)\, dt + \sigma_k(k)\, dB_t, \tag{3}$$

where the drift of Equation (3) is

$$\mu_k(c, k, s) \equiv k^\alpha - c - \left[s \cdot \hat{\kappa} + \delta - \frac{(\sigma_y)^2}{2} \right] k$$

and the local standard deviation is

$$\sigma_k(k) = -\sigma_y k.$$

1.3 Evolution of technology growth states

The finite-state Markov chain for s has an intensity matrix

$$A = N(Q - I),$$

where N is a diagonal matrix of jump intensities, each of which dictates the jump frequency conditioned on the current state. We let η_i denote the jump intensity for state i. The matrix Q is a transition matrix. Each row specifies the probability distribution of the jump location conditioned on a jump taking place. We normalize the transition matrix Q so that its $\{i, i\}$ entry is zero. That is, conditioned on a jump from state i taking place, there is no chance that the state will remain the same.[9] The element $\{i, j\}$ of A will be denoted by a_{ij}, and $a_{i,i} = -\sum_{j, j \neq i} a_{ij}$.

The transition probabilities over any interval of time can be constructed from the intensity matrix A via the exponential formula

$$T_\tau = \exp(\tau A), \tag{4}$$

and the intensity matrix can be deduced from the transition matrices by computing the right derivative of T_τ at $\tau = 0$.

2. The Hidden Information Problem

We consider models in which the mean growth rate $s_t \cdot \hat{\kappa}$ is hidden to investors. Thus they must solve a signal extraction problem by using past levels of technology shock increments to forecast mean growth rates. Before solving the stochastic growth model under the alternative games, we display the solution to the signal extraction problem. A *separation* property of recursive prediction and control in our resource allocation games allows us first to solve the

[9] There are other *normalizations* that might be adopted. For instance, we could make the jump intensity constant across states, provided that conditioned on a jump taking place, there is a positive probability of remaining in the same state. The constant intensity specification is sometimes used because it simplifies the characterization of the stationary distribution.

signal extraction problem using the approximating model and then to use the filtering equations as an input into the solution of the games. The details of the decision problem that justify separation are described in Cagetti et al. (2000).

2.1 General formulation

Since the state variable s is not observed, the decision maker has to infer information about the current state of the system by using the current and past observations of y. This hidden state model is due to Wonham (1964), and is described in Liptser and Shiryayev (1977) and Elliott, Aggoun, and Moore (1995). It has been used in asset-pricing models by David (1997), David and Veronesi (1999), and Veronesi (1999, 2000).

The expected value of the drift of y, $s_t \cdot \hat{\kappa}$, given the current information is $\hat{\kappa}_t = \hat{\kappa} \cdot \hat{p}_t$. The n-dimensional vector $\hat{p}_t$ contains the probabilities of being in each of the states, given the information set at time t $\{\mathcal{Y}_t : t \geq 0\}$. These conditional probabilities evolve according to the stochastic differential equation:

$$\begin{aligned} d\hat{p}_t &= A'\hat{p}_t\,dt + \sigma_{\hat{p}}(\hat{p}_t)\,d\widehat{B}_t \\ \sigma_{\hat{p}}(\hat{p}) &= \frac{1}{\sigma_y}\widehat{P}(I - \mathbf{1}_n\hat{p}')\hat{\kappa}, \end{aligned} \tag{5}$$

where $\widehat{P}$ is a matrix with the elements of $\hat{p}$ on the diagonal. The normalized innovation process $d\widehat{B}_t$ containing the new information used to generate $\mathcal{Y}_t$ is

$$d\widehat{B}_t = \frac{1}{\sigma_y}(dy_t - \hat{\kappa} \cdot s_t\,dt) = dB_t + \frac{\hat{\kappa} \cdot (s_t - \hat{p}_t)}{\sigma_y}\,dt. \tag{6}$$

The evolution of the technology shock under the innovation process $\widehat{B}$ is

$$dy_t = \hat{\kappa} \cdot \hat{p}_t\,dt + \sigma_y\,d\widehat{B}_t \tag{7}$$

and the evolution of k is

$$dk_t = \hat{\mu}(c_t, k_t, \hat{p}_t)\,dt + \sigma_k(k)\,d\widehat{B}_t, \tag{8}$$

where

$$\hat{\mu}(c, k, \hat{p}) = k^\alpha - c - \left[\hat{\kappa} \cdot \hat{p} + \delta - \frac{(\sigma_y)^2}{2}\right]k.$$

2.2 The two-state case

We consider in particular the case in which the state s_t can assume two values, corresponding to a positive growth rate $\hat{\kappa}_1$ in expansions and a negative growth rate $\hat{\kappa}_2$ in recessions. The technology shock process is shown in the top panel of Figure 1, and the bottom panel plots the estimated probabilities

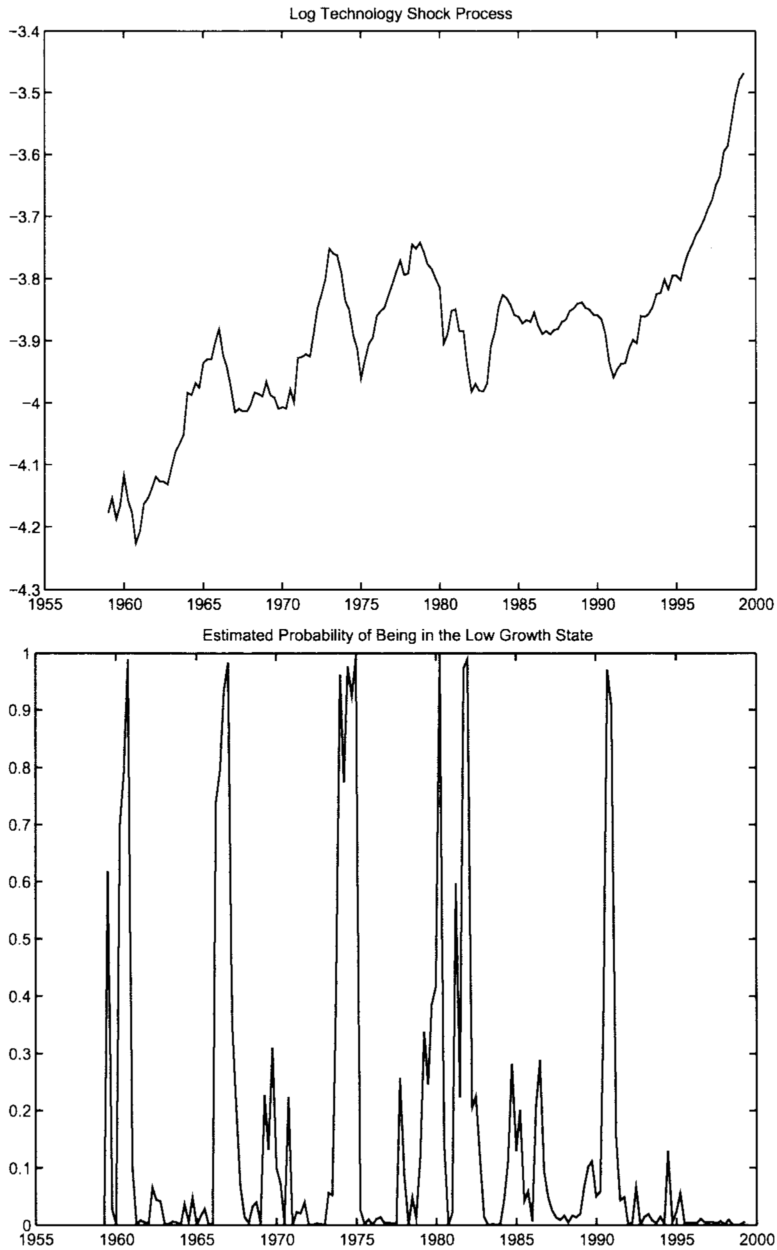

Figure 1
The top panel displays the logarithm of the technology level (the cumulation of the Solow residual) and the bottom panel displays the probability of being in the low-growth state.

Table 1
HMM model parameters

Parameter	Description	Quarterly value
$\hat{\kappa}_1$	High growth	0.0114
$\hat{\kappa}_2$	Low growth	−0.0290
σ_y	Technology shock standard deviation	0.0192
$1/a_{12}$	Mean duration of the high-growth state	13.58
$1/a_{21}$	Mean duration of the low-growth state	2.84

The parameter values were obtained by estimating the HMM using time-series data on the Solow residual.

of being in the low-growth state computed using our baseline estimates. Under our approximating model, the post–World War II experience was one characterized by sharp recessions and extended expansions. Although we use this specification as a model of the technology process, it mirrors that used by Hamilton (1989) in his study of output growth.

In constructing Figure 1, we used data on the cumulative Solow residual y from Citibase, following the construction of Stock and Watson (1999). These residuals are scaled so that they can be interpreted as labor-augmenting technology. These data are quarterly from 1959:Q1 to 1999:Q2 and are constructed from output [gross domestic product (GDP) less farm, housing, and government], capital (interpolation of annual values of fixed nonresidential capital stock using quarterly investment), and labor (hours of employees on nonagricultural payrolls). As in Stock and Watson, we then construct the technology shock process using a labor's share value of 0.65.

To construct probabilities for a two-growth state model, we estimated the HMM using an EM algorithm to compute maximum-likelihood estimates applied to discrete time data, as described by Hamilton (1990). We used the estimated transition probabilities over quarterly intervals to set the intensity matrix A as in Equation (4). Since there are only two states, $1/a_{12}$ is the average length of expansions and $1/a_{21}$ is that of recessions. The intensities and growth rates are reported in Table 1. The time series of probabilities of Figure 1 were computed using both a discrete-time filter and a continuous-time approximation to the filter applied to discrete-time data. Cagetti et al. (2000) give comparisons between the two methods and show that the discretization bias is small. As the table shows, under the approximating model the recessions are short lived and the expansions are persistent.[10]

3. Model Distortions and Approximation Measures

In Sections 1 and 2 we described the technological process and derived the filtering equations for the approximating model, ignoring robustness. In

[10] While the parameters of the HHM model can assume values that accommodate the longer-term productivity slowdowns that are often used to describe the 1970s and early 1980s, the maximum-likelihood estimates instead feature a recession-expansion classification, as is consistent with the related empirical literature in macroeconomics.

this section we study how a concern for robustness affects the approximating model. We present two different ways to incorporate robustness, called game I and game II. While the main intuition behind these two representations is similar, each of them analyzes different ways in which the approximating model can be perturbed. Then, in Section 3.5, as a benchmark we also describe the full information robustness case, where the growth state s is perfectly observed (and thus there is no filtering problem), but the decision maker still entertains the possibility that the approximating model describing the technological process is misspecified.

3.1 An overview

Here we give a brief overview of how to enforce robust decision making. As explained in the introduction, the decision maker is uncertain about the exact behavior of the state variables of the model. He starts from an approximating model. One possible interpretation for the approximating model is that it represents the estimates from the time series of observable variables (y in our case) obtained using some relatively simple model (such as the linear evolution equation for y with a two-state jump in the drift). However, the decision maker recognizes that the assumed model is just an approximation, and that the true behavior of y may have more complicated dynamics both in terms of the functional form and the length of the history of observations entering it. He thus considers many other possible models in addition to the approximating one. One way to represent all these other models is to add perturbations to the approximating model and to allow the perturbations to feed back on the history of the state, thereby modifying the evolution equation of the state variables. Sections 3.3–3.5 describe in detail how to add these perturbations. As mentioned before, the type of admissible perturbations is different between game I and game II.

How does the decision maker treat all these possible models? As a vehicle to explore the directions in which a candidate decision rule is most fragile, the decision maker considers the model that delivers the worst utility. We are not saying that he believes that the true model is the worst one, but that by planning a best response against it, he can devise a rule that is robust under a set of models surrounding his approximating model.

We formalize this conservative decision-making process by introducing a second fictitious player. This player has the ability to choose the distortion to the approximating model and plays against the original decision maker. This player thus chooses the model distortion in order to minimize the utility of the first player. The objective for this optimization problem contains a penalty function used to limit the potency of the minimization. We will have more to say about this in Section 6. The original player then chooses his consumption and investment in order to maximize his utility, given the distortion introduced by the fictitious player. The robust decision problem becomes a max-min problem, where min refers to the fictitious player trying to minimize utility, and max refers to the decision maker in the economy, trying to maximize his utility. The exact problem is described in Sections 3.3–3.5.

So far we have not talked about the set of all possible models, that is, the set of possible distortions that can be chosen by the fictitious player. It is of course necessary to restrict such a set. We thus consider only models that are close to the approximating one, and ones that would be difficult to distinguish statistically given the observed time series of past data. We do not want to consider models that are wildly different from the approximating one and that would be easily rejected. The aim of our exercise is to show that even considering very small distortions, that is, distortions very difficult to detect statistically, robustness generates quantitatively relevant effects on prices; even assuming that the approximating model is a very good approximation, small deviations can generate quantitatively relevant effects. To restrict the set of possible models, we penalize the fictitious player for choosing models that are very different from the approximating one. As a measure of the distance between two models, we use conditional relative entropy, a discrepancy measure built from log-likelihood ratios between models. Thus when solving the minimum problem of the fictitious player, we add a penalty term to make sure that this player chooses models that generate low relative entropy, that is, are very close to the approximating one. Sections 3.3–3.5 describe the formula for relative entropy, while Section 6 describes how to interpret it and which values were chosen in our simulations.

Since the evolution of technology shocks are not under the influence of investors, there is a strong form of separation in the optimal resource allocation problem in the absence of model misspecification. As we have seen, we can solve the state prediction problem prior to and separate from the investment problem. Both of the two hidden Markov games we explore preserve features of this separation, but in different ways.

Like the resource allocation problem without misspecification, the first game constructs a Markov-state evolution equation for the hidden-state probability vector. Armed with this and other evolution equations, we solve a full information robustness game with an appropriately constructed state variable. The first game thus treats $\hat{p}$ as any other state variable. Thus the HMM problem is used merely to produce a more complicated evolution equation for the technology shock and treats $\hat{p}$ as a directly observable state variable.

The second game considers a more primitive starting point and views $\hat{p}$ differently than other observable state variables. The decision makers explore misspecification of the evolution of the unobserved state s and its relation to technology growth. We continue to partially separate the prediction problem (the construction of hidden-state probabilities using past data) from the control problem (the selection of investment and consumption profiles). We achieve this separation by considering a recursive, Markovian game that allows the date t decision makers to reconsider their use of past data when making inferences about probabilities of the hidden growth states. We introduce new decision makers at each date as a device to avoid having the growth-state predictions depend on past contributions to utility. These new players in effect remove commitments to backward-looking filtering rules. Probability assessments are made to be more conservative based on the objective of the decision makers from date t forward.[11]

3.2 Perturbations and relative entropy

As previously explained, we need to introduce distortions to the approximating model, and we need to measure the distance between the approximating model and the distorted one. Let us consider first a diffusion process (without jumps in the drift) generated by a Brownian motion B_t. A very general way to introduce perturbations is to change B_t into $B_t + \int_0^t h_s \, ds$, where h_s is some process adapted to the filtration generated by B. The process h does not have any specific parametric form, but it is disguised by the Brownian motion. By adding h, we are describing a large class of models without making any parametric assumption about the uncertainty structure. Since the h process is masked by a Brownian motion, it can be hard to detect given historical data.

To measure the discrepancy between the approximating ($h = 0$) and the distorted model, we use conditional relative entropy. Relative entropy is typically defined as the expected value of the log-likelihood ratio, that is, of the expected value of the log of the Radon–Nikodym derivative of the perturbed model with respect of the approximating one, where the expected value is computed using the density generated by the perturbed model. The relative entropy equals zero when the two models coincide (the Radon–Nykodim derivative is 1), and is positive when they do not. Anderson, Hansen, and Sargent (2000) and Hansen et al. (2001) give the exact formulas for the relative entropy, and we will say more about how to interpret the entropy measure in Section 6. The time derivative of the relative entropy, that is, the contribution of the current $h_t dt$ to the relative entropy, turns out to be simply $\frac{1}{2}|h|^2$ (so that the larger $|h|$, the larger the distance between two models).

The HMM structure can change the interesting class of perturbations and the associated relative entropy measure, as we will see in Section 3.4. We consider two HMM robustness games that differ in the type of perturbations that are entertained. Conveniently, both can be formulated as solutions to Hamilton–Jacobi–Bellman equations, as we will describe in Section 4.

3.3 Hidden information: representation I

We follow Anderson, Hansen, and Sargent (2000) in formulating our first robust resource allocation problem. The state vector for the hidden Markov

[11] This approach differs from others in the robust control theory, such as Basar and Bernhard (1995). They consider time 0 commitment games, in which both agents may commit to a certain set of decision rules. Therefore, in their framework, the backwards-looking filtering of data to estimate the current vector of growth-state probabilities depends on past contributions to utility.

game is $(k, \hat{p}, y)$, which evolves according to Equations (5) and (8). The stochastic evolution of these state vectors is governed by a single Brownian motion technology shock, $d\widehat{B}_t$. We disguise model misspecification within the Brownian motion shock. Thus, as explained in the previous section, we replace $d\widehat{B}_t$ with $d\widehat{B}_t + h_t dt$, where the process h_t is a process adapted to the filtration $\mathcal{Y}_t$, and we alter Equations (5) and (8) to be

$$\begin{aligned} dk_t &= [\hat{\mu}_k(c_t, k_t, \hat{p}_t) + \sigma_k(k)h_t]\, dt + \sigma_k(k)\, d\widehat{B}_t \\ d\hat{p}_t &= [A'\hat{p}_t + \hat{\sigma}_{\hat{p}}(\hat{p})h_t]\, dt + \hat{\sigma}_{\hat{p}}(\hat{p}_t)\, d\widehat{B}_t \\ dy_t &= [\hat{\kappa} \cdot \hat{p}_t + \sigma_y h_t]\, dt + \sigma_y\, d\widehat{B}_t. \end{aligned} \tag{9}$$

There are two control variables in the decision problem. The maximizing player selects c and the minimizing player chooses h. As noted by Fellner (1965), such probability slanting should not be interpreted as the *beliefs* of the decision maker defined independently of the context of the decision problem. The drift, hdt, appended to the Brownian motion contributes $\frac{(h_t)^2}{2}$ to the systematic part of the instantaneous log-likelihood ratio.

Anderson, Hansen, and Sargent (2000) motivated their analysis by treating the composite state vector as observable. Here the component $\hat{p}_t$ is observable, but it conveys no new information beyond current and past values of the technology process. Instead $\hat{p}_t$ is a variable constructed to make the decision problem Markovian. Because it is constructed by the agent as a function of the history of the observations, it seems to have a different status than the technology process y_t, which is constructed by nature. Our next representation of the hidden state recognizes this difference and thereby prepares the way for a richer class of model perturbations when we eventually compose our zero-sum two-player game to promote robust decisions.

3.4 Hidden information: representation II

We now consider a second representation of the hidden information that was developed in Cagetti et al. (2000) and is the nonlinear counterpart to one explored by Hansen, Sargent, and Wang (2002). Instead of hiding perturbations in the reduced information Brownian motion $d\widehat{B}_t$, we consider perturbations in the original Brownian motion dB_t, in the growth rate vector $\hat{\kappa}$ and in the evolution of the state s_t. We now suppose that there is a vector g_t with the same dimension as the number of states. That is, we assume that the malicious player can observe the underlying stochastic processes (that are not observed by the decision maker) and perturb them. This vector adds a drift to the full information Brownian motion, and the increment dB_t is replaced by $g_t \cdot s_t dt + dB_t$. This same vector, g_t, also alters the forward evolution of $\hat{p}_t$. The drift $A\hat{p}_t$ of the probability updating equation is modified to be

$$A'\hat{p}_t\, dt + \eta(g_t, \hat{p}_t)\, dt,$$

where

$$\eta(g,\hat{p}) = \begin{bmatrix} \hat{p}^1[\frac{1}{2}(g^1)^2 - \text{ent}(g,\hat{p})] \\ \cdot \\ \cdot \\ \cdot \\ \hat{p}^n[\frac{1}{2}(g^n)^2 - \text{ent}(g,\hat{p})] \end{bmatrix}$$

and

$$\text{ent}(g,\hat{p}) = \frac{1}{2}\sum_i (g^i)^2 \hat{p}^i.$$

Superscripts on $\hat{p}$ and g are used to denote entries of the respective vectors and subscripts are used to denote time. Note that the drift perturbation η satisfies $\mathbf{1}_n \cdot \eta(g,\hat{p}) = 0$. Thus the local mean of $\mathbf{1}_n \cdot \hat{p}_t$ is zero, as it should be since probabilities must add up to one. Whenever g^i is large relative to the other components of g, the resulting drift in $\frac{d\hat{p}^i}{\hat{p}^i}$ is increased through the term $\frac{1}{2}(g^i)^2 - \text{ent}(g,\hat{p})$

Altering the underlying growth states by replacing $\hat{\kappa}$ with $\hat{\kappa}+g$ changes the accuracy of the probability estimates. Highly dispersed growth states should be easier to detect. This shows up in our analysis by replacing $\hat{\sigma}_{\hat{p}}(\hat{p})$ with

$$\sigma_p(g,\hat{p}) = \frac{1}{\sigma_y}\widehat{P}(I - \mathbf{1}_n\hat{p}')(\hat{\kappa}+g)$$

in the stochastic contribution to the probability evolution. Thus, in solving these games, we use the forward evolution

$$\begin{aligned} dk_t &= [\hat{\mu}_k(c_t,k_t,\hat{p}_t) + \sigma_k(k)g_t \cdot \hat{p}_t]\,dt + \sigma_k(k)\,d\widehat{B}_t \\ dp_t &= [A'\hat{p}_t + \eta(g_t,\hat{p}_t) + \sigma_p(g,\hat{p})g_t \cdot \hat{p}_t]\,dt + \sigma_p(g,\hat{p})\,d\widehat{B}_t \\ dy_t &= (\hat{\kappa} + \sigma_y g_t)\cdot \hat{p}_t\,dt + \sigma_y\,d\widehat{B}_t, \end{aligned} \tag{10}$$

where the probability vector p_t is *initialized* at $p_t = \hat{p}_t$.

To formalize robustness, we require measures of model misspecification. If s were fully observed, appending a drift to Brownian motion implies a predicted likelihood increment of $\frac{(g_i)^2}{2}$. Since this measure conditions on the unknown state, we average it using the $\hat{p}$ probabilities to obtain the overall measure $\text{ent}(g,\hat{p})$. Averaging the full information likelihood in this way imitates the EM construction of maximum-likelihood estimates of HMMs, as in Dembo and Zeitouni (1986).

In summary, a change in the vector g does two things. It alters the vector $\hat{\kappa}$ of potential growth rates, and it simultaneously changes the probability evolution. This dual role for g emerges as the solution to a specification error minimization problem in which a richer class of perturbations is entertained. In particular, A, $\hat{\kappa}$, and $\hat{p}$ can be altered in distinct ways. The restrictions

Table 2
Distortions for two representations with hidden information

Game	dB distortion	$d\widehat{B}$ distortion	$A'\hat{p}$ distortion	$\hat{\kappa}$ distortion
Hidden I		$d\widehat{B}+h\,dt$		$\hat{\kappa}+\sigma_y h\mathbf{1}_n$
Hidden II	$dB+g\cdot s\,dt$	$d\widehat{B}+g\cdot\hat{p}\,dt$	$A'\hat{p}+\eta(g,\hat{p})$	$\hat{\kappa}+\sigma_y g$

imposed on the relation between distortions can be defended from a more primitive starting point. [See Cagetti et al. (2000) for a formal treatment.]

Table 2 summarizes the differences between the two representations of information under a potential distortion. Notice that the perturbations in representation I make no explicit reference to the evolution of the hidden state s. A drift distortion $h\,dt$ (independent of s) is appended to the Brownian motion increment for the Brownian motion associated with investors' information. In contrast, the counterpart to h, denoted by g, in representation II can depend on states and hence can capture changes in the hidden growth rates.

We will subsequently define two HMM games associated with the two representations of information. These games are designed to deliver forms of robustness by having a fictitious agent *choose* perturbations in a malevolent way. Thus the decision variable for this fictitious agent is h for HMM game I and g for HMM game II. We limit the malevolence by adding quadratic penalties to the objectives of decision makers scaled by θ. We use the statistical discrepancy measures described earlier as penalties. These penalties are $\frac{h^2}{2}$ for game I and $\sum_i \frac{(g_i)^2}{2}$ for game II. The resulting two-player games, described later in Equation (12), will be Markov, which gives computationally tractable alternatives to the Markov decision problem without robustness.

3.5 A full information benchmark

For comparisons we will also consider a model in which s_t is directly observed. Like hidden information game I, this game is formulated as in Anderson, Hansen, and Sargent (2000), except with a different vector of state variables. The actual state s replaces vector $\hat{p}$ of state probabilities. We use $\hat{p}$ to form a time series for s by assigning s to the low-growth state when $\hat{p}$ is greater than one-half and to the high-growth state when $\hat{p}$ is less than one-half.

4. Value Functions

We have described representations of information for an HMM model and another full information benchmark. We now use these representations of information to solve resource allocation problems that incorporate a concern about robustness. These model specification games are all Markov and have a single value function. In this section we report the partial differential equations for these functions, which are known as Hamilton–Jacobi–Bellman (HJB) equations. To study these games we will solve these differential equations numerically as described in Appendix C.

4.1 Preferences and discounting

It is convenient to scale consumption and capital by the technology level. This will eventually allow us to derive HJB equations that do not depend on y, but only on $(k, \hat{p})$, thus simplifying our numerical analysis. This scaling, however, has the effect of introducing stochastic discounting into the preferences. We will use this same discounting to evaluate model misspecification.

Let $U(C)$ be the instantaneous flow of utility, which we parameterize using a constant elasticity of substitution. The time t contribution to the discounted power utility function is

$$\begin{aligned} \exp(-\rho t)U(C_t) &= \exp(-\rho t)(C_t)^{1-\gamma}/(1-\gamma) \\ &= \exp[(1-\gamma)y_t - \rho t]\ (c_t)^{1-\gamma}/(1-\gamma) \\ &= Y_t^* \frac{(c_t)^{1-\gamma}}{1-\gamma}, \end{aligned}$$

where ρ is the subjective rate of discount and

$$Y_t^* \equiv \exp[(1-\gamma)y_t - \rho t].$$

We now view Y_t^* as a stochastic discount factor and $c = \frac{C}{Y}$ as a decision variable. Notice that

$$dY_t^* = Y_t^* \left((1-\gamma)\, dy_t + \left[\frac{(1-\gamma)^2(\sigma_y)^2}{2} - \rho \right] dt \right). \tag{11}$$

In the robustness games, Y_t^* is used to discount instantaneous utilities and discrepancy measures.

4.2 Hamilton–Jacobi–Bellman equations

Let x be a composite-state vector that includes $(k, \hat{p}, Y^*)$. The stochastic evolution [Equations (9) and (11)] for game I can be written as

$$dx_t = \mu_x^1(c_t, h_t, x_t)dt + \sigma_x^1(x_t)d\widehat{B}_t,$$

where h_t is a scalar perturbation and σ_x^1 is a column vector. This column vector depends on x, but not on h. As explained in Appendix B, the HJB equation for game I is

$$\begin{aligned} 0 = \max_c \min_h Y^* \left[\frac{c^{1-\gamma}}{1-\gamma} + \theta \frac{h^2}{2} \right] + \mu_x^1(c, h, x)' \cdot \nabla W^1(x) \\ + \frac{1}{2} [\sigma_x^1(x)]' \frac{\partial^2 W^1(x)}{\partial x \partial x'} [\sigma_x^1(x)], \end{aligned} \tag{12}$$

where ∇ is the gradient. The HJB equation is composed of two parts, one representing the contribution of the current reward to the value function, the other representing the expected change in the value function W^1. The first part is

$$Y^*\left[\frac{c^{1-\gamma}}{1-\gamma}+\theta\frac{h^2}{2}\right],$$

which contains the current utility of the decision maker, plus, as explained in previous section, a penalty for the malicious player. As discussed, the penalty is given by the entropy measure of the discrepancy between the approximating and perturbed models. Note that the entropy term $\frac{h^2}{2}$ is multiplied by a (positive) parameter θ that controls the degree of robustness that is sought. When θ is ∞, the penalty is also ∞, and therefore no perturbations are allowed. Therefore $\theta=+\infty$ gives the usual resource allocation problem without robustness. Smaller θ result in smaller penalties and hence accommodate a larger exploration of alternative models. The calibration of θ in our simulation is described in Section 6.

For game II we replace Equation (9) with Equation (10). We write the stochastic evolution as

$$dx_t=\mu_x^2(c_t,g_t,x_t)\,dt+\sigma_x^2(g_t,x_t)\,d\widehat{B}_t,$$

where g_t is an n-dimensional perturbation and $\sigma_x^2(g_t,x_t)$ is a column vector. In contrast to game I, this volatility vector depends on the perturbation g. The HJB equation is of the same form as game I, but with a different stochastic evolution:

$$0=\max_c\min_g Y^*\left[\frac{c^{1-\gamma}}{1-\gamma}+\frac{\theta}{2}\sum_i(g^i)^2\hat{p}^i\right]+\mu_x^2(c,g,x)\cdot\nabla W^2(x)$$
$$+\frac{1}{2}[\sigma_x^2(g,x)]'\frac{\partial^2W^2(x)}{\partial x\partial x'}[\sigma_x^2(g,x)]. \tag{13}$$

Under game II, the local mean of the value function W^2 is the negative of

$$Y^*\left[\frac{c^{1-\gamma}}{1-\gamma}+\frac{\theta}{2}\sum_i(g^i)^2\hat{p}^i\right]$$

appropriately optimized. We again nest a decision problem without concern for robustness by setting the tuning parameter θ to infinity.

4.3 Computations

We compute the value functions for these games numerically. Given the presumed structure of shocks, we thought it best to avoid using the linearization techniques commonly employed in macroeconomics. Conveniently the value functions are linear in the stochastic discount factor. That is, they satisfy

$W(k, \hat{p}, Y^*) = Y^* V(k, \hat{p})$, and as a consequence we can focus our computations on determining V. This scaling property follows because differential equations [Equations (12) and (13)] are both linear in Y^*.

For both games, consumption satisfies

$$c^* = \left[\frac{\partial V}{\partial k}(k, \hat{p}) \right]^{-\frac{1}{\gamma}}. \tag{14}$$

Moreover, Equation (12) is quadratic in the scalar h and Equation (13) is quadratic in the vector g. Thus for given value functions, the control laws for c, g, and h are easy to compute. The solution algorithm in Appendix C exploits this simplicity.

We solve the complete information game in an entirely analogous fashion, except in this case we eliminate the dependence on $\hat{p}$ and carry along a vector of value functions (one for each growth state) that only depend on k.

5. An Illustrative Growth Economy

As we have seen, the HMM version of the stochastic growth model separates as follows. We can first solve a signal extraction problem and deduce the hidden state probabilities. We can then use these hidden state probabilities in conjunction with the directly observed technology shock process as a *multivariate* stochastic forcing process for a robust resource allocation problem. In this section we take the time paths for the technology process and the hidden state probabilities as inputs into our calculations. These are the exogenous forcing processes for our models. The robust resource allocation problems imply trajectories for capital, consumption, and investment. We now study these quantity implications to understand better the precautionary motive induced by robustness.

5.1 Implications for capital accumulation

To illustrate the impact of robustness, we compute the implied time series for k and compare them to actual data. To make this comparison we must fully parameterize preferences. Initially we set the power utility parameter $\gamma = 2$, subjective discount rate $\rho = .04$, and depreciation rate $\delta = .07$. We initialize the initial capital technology ratio to the corresponding level in the data in 1959:Q1 and then compute the solution for the various decision problems by using the trajectories of y and $\hat{p}$ shown in Figure 1.

Figure 2 shows the evolution of our endogenous state variable k, which can be interpreted as the capital-effective labor ratio. The figure reports the evolution of k for the data and for the hidden information decision problem without robustness ($\theta = \infty$). The time path for k implied by the $\theta = \infty$ model mimics the actual data, although the model generates a higher trajectory early on and a lower one later. The overall similarity to the data should come as

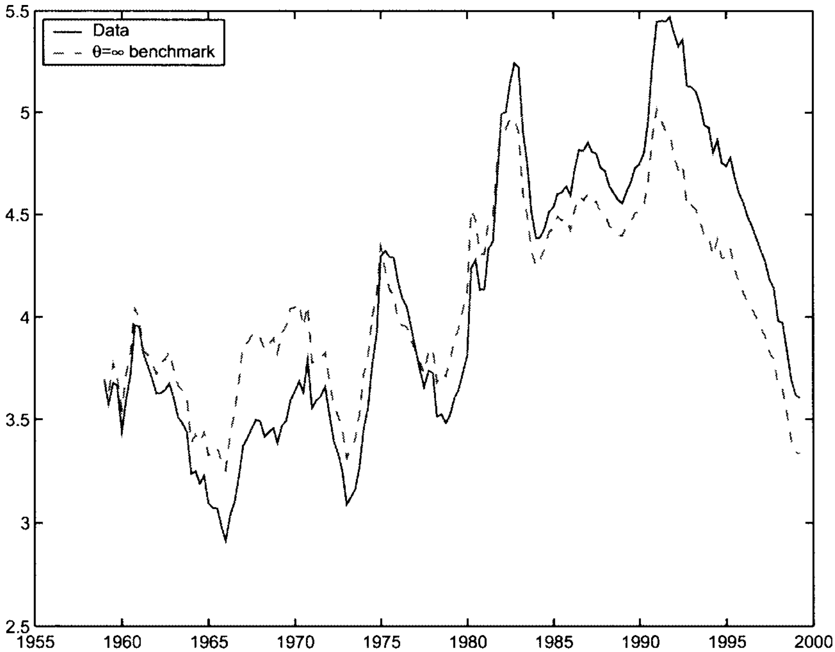

Figure 2
This figure displays the observed time series of the capital-technology ratio k and the ratio implied by an HMM benchmark model with $\theta = \infty$ (i.e., no preference for robustness).

no surprise since the parameter configurations were selected in part to match the growth features of the model.[12]

The effect of robustness on the evolution of k is depicted in Figure 3. The figure shows that the robustness games imply a higher level of capital than the corresponding nonrobust decision problems. After starting with the same initial capital, a decision maker endowed with a concern about model misspecification builds up the capital stock more quickly. Robustness acts as an additional precautionary motive for saving and generates a higher buffer stock of capital.

In the standard expected utility framework, precautionary savings are generated by the possibility of bad shocks coming from a given and prespecified probability distribution. These effects are known to be small when calibrated to macroeconomic measures of uncertainty.[13] In our robust setup, the decision maker also considers the potential misspecification of his approximating

[12] Moreover, the technology process was itself extracted from aggregate quantity data.

[13] See Aiyagari (1994) and Krusell and Smith (1998) among others for a discussion of this issue. Small precautionary savings are needed to justify the long-standing tradition of macroeconomic calibrations to steady-state relations.

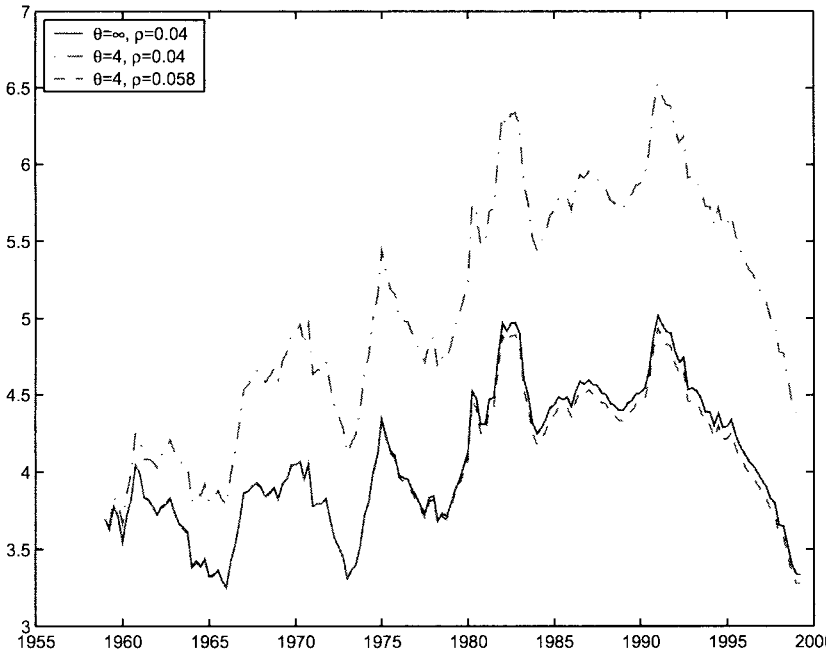

Figure 3
This figure displays the capital technology trajectories implied by different values of the subjective discount rate (ρ) and the robustness parameter (θ). Introducing robustness results in additional capital accumulation holding fixed the discount rate. This is seen by comparing the $(\theta, \rho) = (\infty, 0.04)$ trajectory to the $(\theta = 4, \rho = 0.04)$ trajectory. This increase in savings can be offset by simultaneously increasing the subjective rate of discount. This is seen by comparing the $(\theta = 4, \rho = 0.04)$ trajectory to the $(\theta = 4, \rho = 0.058)$ trajectory.

model. This induces an additional precautionary mechanism that may not be quantitatively small. The robust social planner will save more to build up a higher capital level. He fears that the future growth rates will be lower than those implied by the approximating model, and as a consequence will keep a larger buffer stock of capital. Figure 3 shows that this additional precautionary motive can be very important quantitatively.

Macroeconomic calibrations of parameters based on mean growth rates and average returns typically ignore the impact of precautionary savings.[14] When robustness is introduced, these calibrations must be modified. Decreasing the robustness parameter θ increases this robust precautionary motive and increases the average level of capital. A similar effect can, however, be

[14] Even when the macroeconomic model justifies a quantitatively small amount of precautionary savings, macroeconomic calibrators face the dilemma of which average return to use in pinning down parameters. As emphasized by Cochrane and Hansen (1992), the same macroeconomic model that is used to explain the evolution of capital is poorly suited for accounting for the observed risk premia. This is just a restatement of the so-called equity premium puzzle.

obtained by decreasing the discount rate ρ instead. A more patient decision maker will also want to hold more capital. Hansen, Sargent, and Tallarini (1999) and Hansen, Sargent, and Wang (2002) show that for a discrete-time linear quadratic permanent income model, the extra precautionary effect due to robustness can be fully offset by increasing the subjective discount rate. In our setup, with power utility and nonlinear state evolutions, this is not exactly true. However, Figure 3 shows that this result does hold as a remarkably good approximation. The figure plots the implied capital stock time series for alternative values of θ and ρ, and we see that by simultaneously lowering θ and increasing ρ we can preserve the quantity implications of the nonrobust ($\theta = \infty$) model. Therefore the same quantity data can be generated by various configurations of θ and ρ, and in our pricing calculations below we will look at those configurations that leave the quantity data unchanged, as in Hansen, Sargent, and Tallarini (1999).

5.2 Other quantity implications

The approximating model fails to capture some aspects of the data. In addition to abstracting from labor supply, the model is known to imply too much consumption volatility. The ratio of the standard deviation in consumption growth to that in output is approximately one in our model, but only one-half in the data. The excess consumption volatility implied by the model suggests that the risk premia are likely to be larger than in a model with more plausible consumption variability. However, as we will see in Section 7, they are still substantially lower than in the data, and this increased consumption variability cannot, per se, explain the observed risk premia. The aim of our exercises is to take a simple, widely used, and pedagogically valuable model and study how a concern about robustness alters what is typically viewed as risk premia by a financial econometrician. Adding various elements to the approximating model (such as a richer dynamic for the technological process, various types of adjustment costs and intertemporal complementarities in preferences) may generate more realistic consumption variabilities, but will also make the model more difficult to solve and arguably distract us from our primary interest. We have chosen to use a simpler model to emphasize how a concern about model misspecification can alter security market prices.

6. Tuning Robustness

The parameter θ is used to govern the extent of robustness in HMM games I and II, and in the full information counterpart. It is difficult to restrict the parameter θ a priori independent of the approximating model and of the rules of the game. In this section we describe a different approach that follows an idea in (Anderson, Hansen, and Sargent 2000). We explore families of robustness games indexed by θ. We do not presume that θ for one game has the same meaning as for another, and the θs used in our calculations will differ across games. Associated with each θ (and each game) is an implied alternative *worst-case* model used to enforce robustness in decision making. This alternative model is obtained as part of the solution to a two-player game. Moreover, there is a corresponding conditional relative entropy process that measures the discrepancy between the approximating model and the alternative worst-case model. A smaller value of θ reduces the penalty for exploring model misspecification. Typically this results in a larger entropy process associated with the worst-case model.

Comparing the entropy processes across games is more informative than comparing directly the values of θ. As we have noted previously, the discrepancy measures we use are information-based measures related closely to log-likelihood ratios. A larger entropy process implies that it is easier for a statistician to discriminate between the approximating model and the worst-case model given historical data. While rational expectations presume there is no discrepancy ($\theta = \infty$), we wish to allow for the possibility of model misspecification that is small in a statistical sense. We operationalize this to be small by choosing θ to limit the entropy process that comes from the Markov game solution.

Given a time series of y, we could compute the log-likelihood ratio to test one model (the approximating model $h = 0$) versus the alternative misspecified one ($h = h^*$). As explained in Section 3.2, the entropy term $\frac{(h^*)^2}{2}$ represents the conditional expectation of the instantaneous contribution to the log-likelihood ratio. Cumulating the process $\{\frac{(h_t^*)^2}{2}\}$ over an interval gives the predictable component to the log-likelihood over that same interval.[15] To simplify comparisons across games, we will hold fixed the cumulation over the observed sample. In our computations, we set this to .6, which is arbitrary but gives us a useful benchmark for making comparisons across games.

For HMM game II, matters are more complicated because of the hidden Markov state. Our measure of discrepancy is based on an averaging conditional entropy over the hidden states: $\frac{1}{2}\sum_i (g_i)^2 \hat{p}_i$. An analogous approach is used in likelihood estimation based on the EM algorithm. This algorithm is commonly used because of its computational convenience and its relation to likelihood ratios. Thus the discrepancy measure is distinct, but closely related to a conditional expectation of a log-likelihood ratio.

Anderson, Hansen, and Sargent (2000) and Hansen, Sargent, and Wang (2002) provide more sophisticated discussions linking statistical detection

[15] This reference to a predictable component is ambiguous without stating which model is used. It turns out that under either the approximating model or the worst-case model, this same expression is valid. Expectations of functions of this cumulation differ depending upon which model was used in the computation. The symmetry in the instantaneous contributions does not hold in a full information game with observable jump components.

to the choice of θ. Consider the following simplified problem of statistical discrimination. Use historical data to make a pairwise choice between an approximating model and a worst-case model. This can be formulated as a Bayesian decision problem. Two types of errors are possible depending upon which model is correct. The resulting detection error probabilities provide a formal way to quantify statistical discrimination as originally suggested by Chernoff (1952). Like Chernoff, Anderson, Hansen, and Sargent (2000) use bounds on detection error probabilities between competing models to help understand better any particular value of θ. For instance, a large value of θ is one for which the worst-case model is hard to detect statistically given the approximating model. Similar to the rational expectations intuition, the misspecified models a decision maker aims to be protected against are those that could not have been ascertained easily given historical data.

At least in the case of diffusions, conditional entropy is an important input into these bounds on detection error probabilities. Instead of using bounds, Hansen, Sargent, and Wang (2002) compute detection error probabilities via simulation and allow for hidden Markov states. Computations such as these are more refined ways to use the Bayesian theory of statistical discrimination to assist in the choice of θ. Here we use a less ambitious approach of fixing the sample accumulation of the relative entropy process. When relative entropy is constant, however, we may compute easily a Chernoff-type bound on the detection error probabilities in terms of the entropy accumulation. For an entropy accumulation of .6, the upper bound on the detection error probability is .43.[16] Thus if forced to decide between the approximating model and the worst-case model, a statistician should expect mistakes to occur. Unfortunately, as we will see, that assumption of constant relative entropy process is counterfactual for our models. Thus our use of .6 for the entropy accumulation is only a rough way to get comparability across games and is not meant as a way to give a definitive choice of the parameter θ. In fact, we suspect that even considerably smaller values of θ might be reasonable for our model economies. Moreover, there are other potential approaches for tuning the robustness parameter θ. While detection error analysis eliminates candidate models that should be easy to uncover from historical data, it considers only a very highly stylized model selection problem. The utility consequences of being robust and the costs of active learning arguably should also come into play.[17]

[16] The probability of making a classification error is bounded above by $\frac{1}{2}\exp(-\frac{1}{4}ce)$, where ce is the accumulation of the conditional entropy.

[17] We envision that investors confront a much more complicated problem in statistical discrimination, one that would entertain a wide array of models and would use the decision problem to weight the errors in misclassifying models. Implicitly we are presuming that when confronted with such a problem, the investors decide that model discrimination is too difficult and focus instead on making their investment decisions robust.

Table 3
Parameter settings for θ and ρ

Game	θ	ρ
Hidden I	4.0	.058
Hidden II	6.5	.046
Full information	4.4	.055
HMM benchmark	∞	.040

This table reports the values of the robustness parameter θ and the subjective discount rate ρ that are used in our subsequent calculations. These configurations leave the time path for the capital-technology ratio virtually the same as that for the $\theta = \infty$ economy with hidden information. The sample sum of the conditional entropies are .60 for HMM game I, .60 for HMM game II, and a slightly larger .61 for the full information benchmark.

Based on detection error considerations, we employed the θ and ρ pairs reported in Table 3 in our subsequent calculations. These choices make the cumulation of relative entropy remain about the same across games. We report the time series of entropy measures for both HMM games in Figure 4. There is considerable variability in these measures, with more variability in the entropy measures for HMM game II than for game I. The entropies for both HMM games are particularly large when the investors are unsure which growth state they are in as measured by $\hat{p}$.

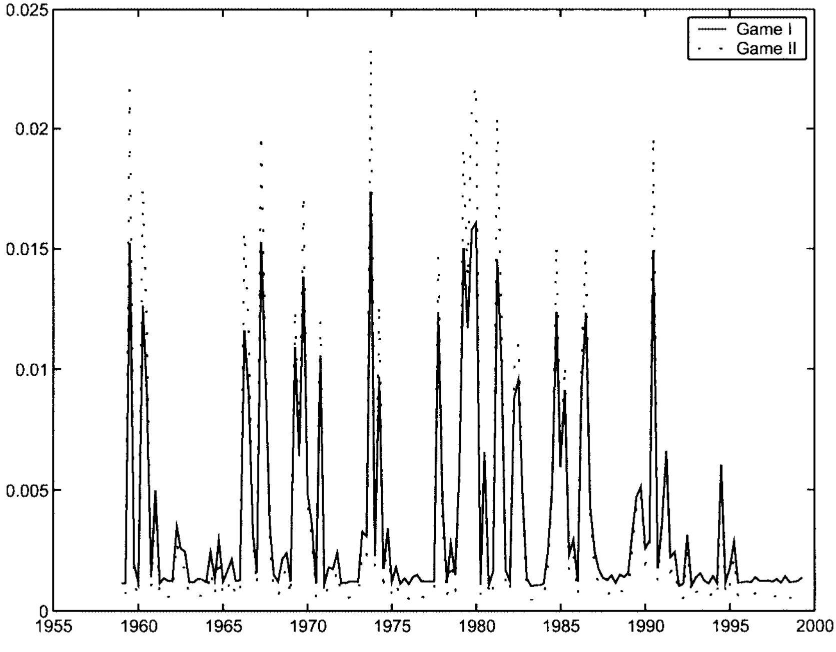

Figure 4
The top panel gives the time series for the conditional relative entropies for the two HMM games. For HMM game I we report $\frac{(h^*)^2}{2}$ and for HMM Game II we report $\frac{\sum_i (g_i^*)^2 \hat{p}_i}{2}$.

7. Local Risk and Uncertainty Prices

One common way of studying the dynamic implications of a model is to report the implied impulse response functions. Since our model is explicitly nonlinear, impulse response functions based on linear approximations seem ill suited to characterize the implications of these models. Instead we report plots of prices as functions of the state variables in the model and the time series implied by the historically observed technology levels.

The HMM formulations imply a particular Markov evolution for the technology shock process where the Markov chain state probabilities become an additional state variable. As we showed in Figure 1, a time series of these state probabilities can be constructed from the observed data on the technology level. For both HMM games we solve numerically for the law of motion for the capital stock. Using this solution and a given time series for the state probabilities, we recursively generate a time series for the implied capital stock. Thus from a given time-series trajectory for the technology shock process, we can compute time series for the state probabilities and the capital stock implied by the model. These time series will be used in some of the calculations that follow.

As we noted above, the effect of robustness on the capital evolution can be offset by increasing the rate of discount. In the calculations in this section we experiment with different levels of the robustness parameter θ, varying ρ to maintain the quantity implications. Given that the capital stock trajectory remains essentially the same across games, the instantaneous risk-free rate measured by the marginal product of capital also remains the same. The shadow prices of the Brownian motion increments, however, will differ.

We study the (*local*) pricing of $d\widehat{B}$ for our two decision models in the presence of hidden information. Consider the local relation between the instantaneous return on a security μ_r, its factor loading on the Brownian increment σ_r, and the risk-free rate μ_f:

$$\mu_r - \mu_f = \sigma_r \lambda.$$

Then λ is the factor risk price and $|\lambda|$ is the absolute slope of the mean standard deviation frontier. In our model so far, there is a single Brownian motion factor $d\widehat{B}$. Anderson, Hansen, and Sargent (2000) and Chen and Epstein (2001) show that the factor price of the Brownian increments $d\widehat{B}$ can be decomposed into two prices: the usual price for risk and a price of model uncertainty.[18] Thus the factor price λ is the sum

$$\lambda = \lambda_m + \lambda_u.$$

[18] Anderson, Hansen, and Sargent (2000) and Chen and Epstein (2001) differ in the way the beliefs that dictate prices are deduced. The local prices for game I can be viewed as a special case of those in Anderson, Hansen, and Sargent (2000). As we have seen, however, the beliefs that support the local game II solution exploit more details of the HMM structure. Thus the game II solution and prices are not special cases of those reported in Anderson, Hansen, and Sargent (2000).

The price of risk λ_m is obtained by applying a formula from Breeden's (1979) analysis of a consumption-based asset pricing model. This risk price is given by (minus) the weighting coefficient on $\widehat{B}_t$ in the evolution for the process of the log marginal utility of consumption:

$$\begin{aligned}\pi(k,\hat{p},y) &= \log(C^{-\gamma})\\ &= -\gamma\log c-\gamma y\\ &= \log V_k(k,\hat{p})-\gamma y,\end{aligned}$$

where we have used Equation (14) for the consumption-technology ratio, c. The coefficient on $d\widehat{B}$ is computed as the sum of partial derivatives with respect to y, $\hat{p}$, and k:

$$\lambda_m=\gamma\sigma_y+\sigma_y k\frac{V_{kk}}{V_k}-(1-\hat{p})\hat{p}\frac{\kappa_1-\kappa_2}{\sigma_y}\frac{V_{k\hat{p}}}{V_k}.$$

In addition to the usual component λ_m that depends on the marginal utility of consumption, there is a second component, related to the worst-case model, that emerges in the solution to the two-player game. It is the worst-case model that is used in place of the approximating one for computing shadow prices. Under this worst-case model, S, the price of the asset, evolves according to

$$\frac{dS}{S}=\mu_r\,dt+\sigma_r(d\widehat{B}+h^*\,dt)$$

in game I (and with $\hat{p}\cdot g^*$ instead of h^* in game II). That is, it is as if the drift of the process were $(\mu_r+\sigma_r h^*)\,dt$ instead of $\mu_r\,dt$. This augmented drift can be used in Breeden (1979), which introduces an additional component to pricing vis-à-vis the approximating model. Making this drift adjustment, we get $\lambda_u=-h^*$ for game I and $\lambda_u=-\hat{p}\cdot g^*$ for game II. (Note that h^* and $\hat{p}\cdot g^*$ will affect the expression for the risk-free rate μ_f, since μ_f is derived from the drift of the marginal utility of consumption.) Anderson, Hansen, and Sargent (2000) formalize these arguments and provide a more complete derivation of the factor price decomposition.

Figure 5 gives the price functions for model uncertainty and for risk. It shows how these prices vary with the probability of being in the low state holding the capital-technology ratio at its median level. These price functions are highly nonlinear, with peak effects occurring near probability one-half. Peak effects are associated with having little information about which growth-state regime is in place. As mentioned above, the reported risk prices are overstated, because consumption volatility in our models is about double that in the data. Nevertheless, the uncertainty prices dominate those of the risk prices and display more sensitivity to the probabilities.

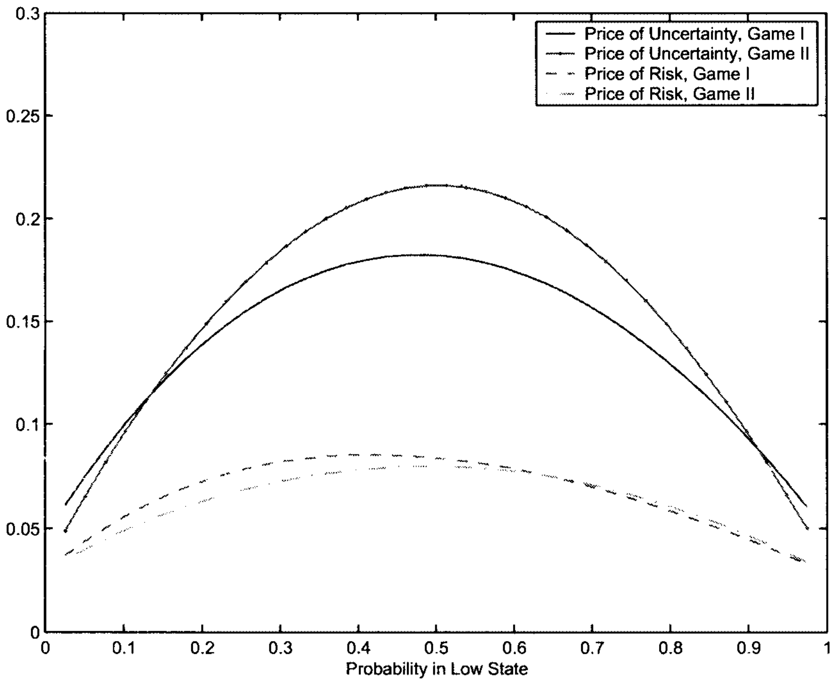

Figure 5
This figure shows the prices of risk and uncertainty as functions of the probability of being in the low-growth state. These functions are computed holding k fixed at the sample median.

Figure 6 shows the prices of model uncertainty as functions of k (the capital-technology ratio) for the two hidden information games and for two values of the probability of being in the low-growth state. The dependence of this price on k is mostly linear and relatively flat in comparison to the dependence on the probability. As a consequence, the implied time series for the uncertainty prices are dominated by movements in the state probabilities. The implied time-series trajectories are reported in Figure 7. These uncertainty prices display cyclical fluctuations, with the peak effects occurring at the beginning and end of recessions. These effects are associated with points in time in which the state probabilities are each about one-half. The prices are lower in game I than game II for probabilities near one-half, but higher for probabilities close to zero or one (see also Figure 5). Thus the model produces substantial cyclical fluctuations in what financial econometricians might mistakenly call risk premia. The high market prices of uncertainty occur not because of confidence in low growth but rather because of ambiguity about which growth regime is currently in play.

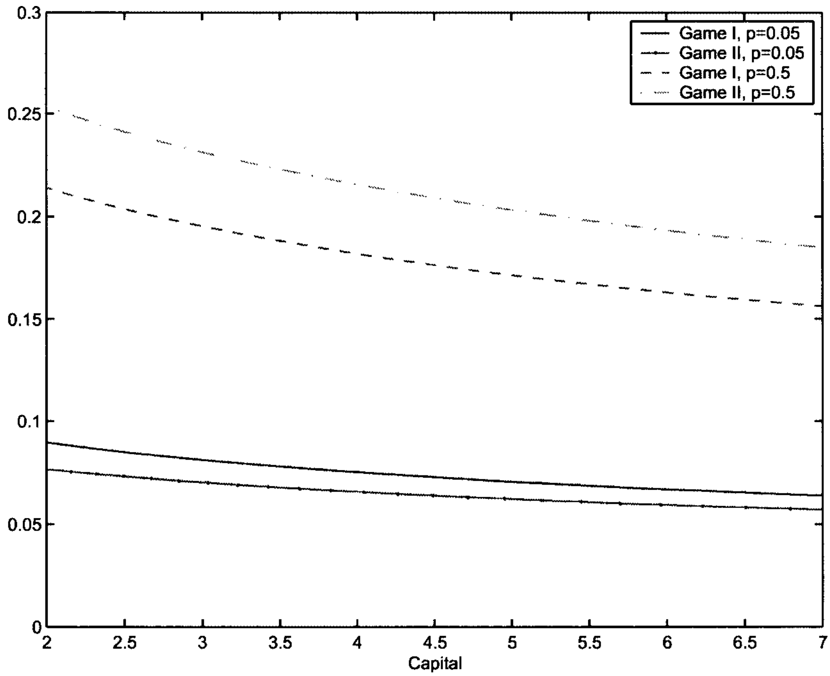

Figure 6
This figure shows the market price of uncertainty as a function of k for games I and II and for two different values of the probability of being in a low-growth state.

8. Price-Earnings Ratios

To study price-earnings ratios, we apply the HMM simultaneously to the technology shock process and an earnings process. Growth rates in technology and earnings respond to the same two-state hidden Markov chain. For the purposes of pricing, we treat the earnings process as a stream of *dividends*: claims to consumption to be priced. These dividends are not an extra source of consumption, but merely a specification of an intertemporal payoff stream to be priced. Production takes place as before, but the growth rate in earnings and technology both respond to the same two-state Markov chain. The two Brownian motions that disguise this state are independent. Investors use data on earnings and technology to make inferences about the common hidden growth state. This formulation closely follows David and Veronesi (1999) except that the state variable k comes into play in our analysis and our aim is to study how a concern about robustness changes the prices. Its discrete-time, full information counterpart has been used extensively in the asset pricing literature [for example, see Cecchetti, Lam, and Mark (1993) and Bonomo and Garcia (1996)].

The time series for technology and earnings are plotted in Figure 8. We used data on reported earnings on the Standard & Poors Stock Price Index

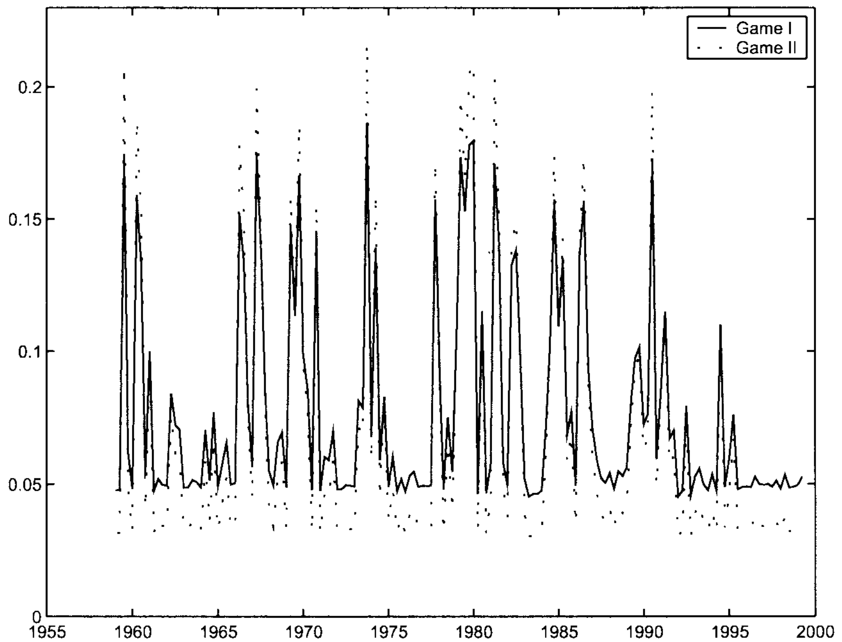

Figure 7
This figure shows the time series for the market price of uncertainty for games I and II. The market price of uncertainty is measured as $-h^*$ for game I and as $-g^* \cdot \hat{p}$ for game II.

divided by the price deflator for fixed investment. The earnings data were obtained from DRI.

8.1 Earnings evolution

In our robust resource allocation problems we used the following model for log technology and log earnings:

$$\begin{bmatrix} dy_t \\ de_t \end{bmatrix} = \begin{bmatrix} \hat{\kappa}_{11} & \hat{\kappa}_{12} \\ \hat{\kappa}_{21} & \hat{\kappa}_{22} \end{bmatrix} s_t dt + \begin{bmatrix} \sigma_y & 0 \\ 0 & \sigma_e \end{bmatrix} dB_t,$$

where dB_t is now a two-dimensional Brownian motion. We used the parameter values reported in Table 4. These parameters were obtained by estimating a discrete-time counterpart HMM using an EM algorithm. While we used a two-state specification to compute prices, we were actually compelled to fit a more complicated model to capture the seasonality in the earnings data.[19] The

[19] An independent, two-state seasonal Markov chain was introduced in the estimation. This seasonal chain only altered the growth rates for earnings. According to our estimates, in one seasonal state growth is reduced by .0836, and in the other it is enhanced by this same amount. The quarterly transition from reduced growth seasonal state to high seasonal state is essentially one, while the transition in the other direction is .986. We deliberately avoid the common practice of using a quarterly series of yearly averages of earnings. While the averaging removes seasonality, it should also change how the signal extraction problem is posed and solved.

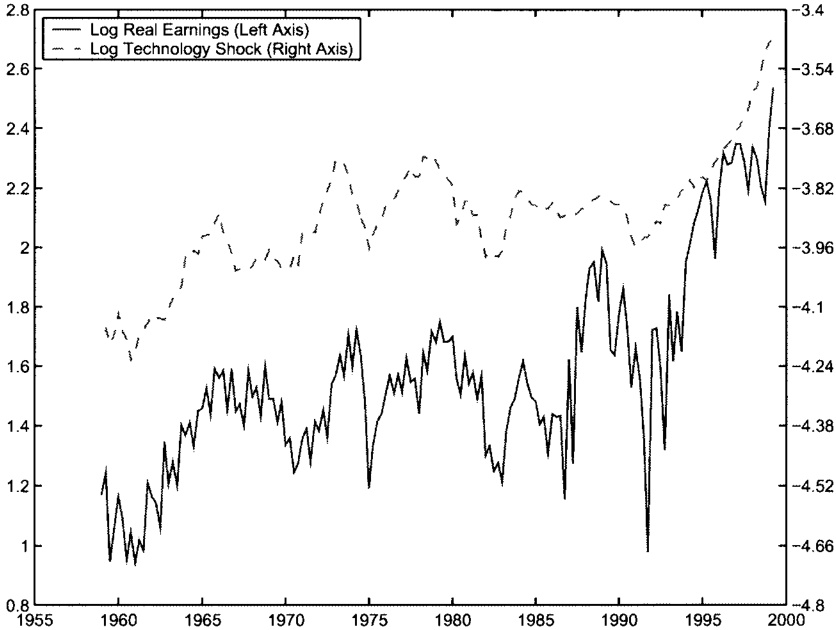

Figure 8
In this figure we report the time series for the cumulative Solow residual and S&P earnings.

estimates of the parameters that govern the technology evolution were very close to those obtained when we estimated the HMM using only the technology data. In our calculations we used the same parameters for the technology evolution for both models. (Compare Table 1 to Table 4.) Since the standard deviation for the earnings process is substantially higher than that for technology, the implied state probability estimates for the technology-earnings model are very close to those reported in Figure 1. The technology process is the primary source of information about the hidden growth state.

Table 4
Parameters for the technology-earnings model

Parameter	Description	Quarterly value
$\hat{\kappa}_{11}$	High technology growth	0.0114
$\hat{\kappa}_{21}$	High earnings growth	0.0205
$\hat{\kappa}_{12}$	Low technology growth	−0.0290
$\hat{\kappa}_{22}$	Low earnings growth	−0.0612
σ_y	Technology shock standard deviation	0.0192
σ_e	Earnings shock standard deviation	0.1291
$1/a_{12}$	Mean duration of the high-growth state	13.58
$1/a_{21}$	Mean duration of the low-growth state	2.84

These parameters were estimated using time-series data on the Solow residual and earnings on the S&P 500 stock index. Since the estimated evolution parameters for technology in the technology-earnings model were essentially the same as those estimated without earnings, we used the same values for the common parameters in our calculations.

8.2 Price calculation

To compute the price function for the HMM games, we use the familiar implication that the local mean of the marginal utility scaled price should be minus the marginal utility scaled dividend. The solutions to the respective robustness games provide us with the formula of the local mean of the marginal utility scaled dividends.

To formalize this idea, we form a state vector z that contains $(k, \hat{p}, Y^*, e)$. Let μ_z^1 denote the drift implied by the Markov solution to HMM game I and Σ_z^1 the corresponding diffusion matrix for the bivariate Brownian motion $d\widehat{B}_t$ associated with the investors' information set. The drift vector μ_z^1 includes terms which reflect how the worst-case perturbations h^* affect the state variables and hence adjusts for model uncertainty. This is consistent with our incremental pricing results above. Let Π^1 denote the marginal utility scaled pricing function that maps the composite state z into the marginal utility scaled price. This function satisfies the partial differential equation

$$\mu_z^1 \cdot \nabla \Pi^1 + \frac{1}{2} \text{trace} \left(\Sigma_z^1 \frac{\partial^2 \Pi^1}{\partial z \partial z'} \right) = -\exp(e)(C^1)^{-\gamma} Y^*, \tag{15}$$

where C^1 is the consumption function implied by the Markov solution to game I. The left-hand side of this equation is the local mean of the marginal utility scaled price and the right-hand side is the negative of the marginal utility scaled dividend (measured by earnings). There is an analogous pricing equation for HMM game II.

For computational purposes, it is convenient to transform the partial differential equation [Equation (15)]. The price-dividend ratio in this economy can be shown to depend only on k and $\hat{p}$ and not on Y^* and e.[20] We exploit this reduced dependence in solving for the equilibrium price of an earnings claim. The figures that follow report the price-dividend ratio as a function of k and $\hat{p}$ without the marginal utility scaling.

8.3 Results and figures

We solved for price-earnings functions for the HMM robustness games. For comparison we also computed the full information prices and the $\theta = \infty$ HMM prices. The former prices presume the growth state is known to investors and the latter prices abstract from robustness.

Figure 9 shows how price-earnings ratios vary with the probability of being in the low-growth state. These functions decrease with that probability and are essentially linear. When investors are confident they are in the low-growth state, the price is lower. The HMM robustness games imply a lower price-earnings ratio than the HMM benchmark without robustness. A concern about

[20] The state variables Y^* and e can be eliminated by dividing both sides by $\exp(e)Y^*$ and solving for $\frac{\Pi^1}{Y^* \exp(e)}$ instead of for Π^1.

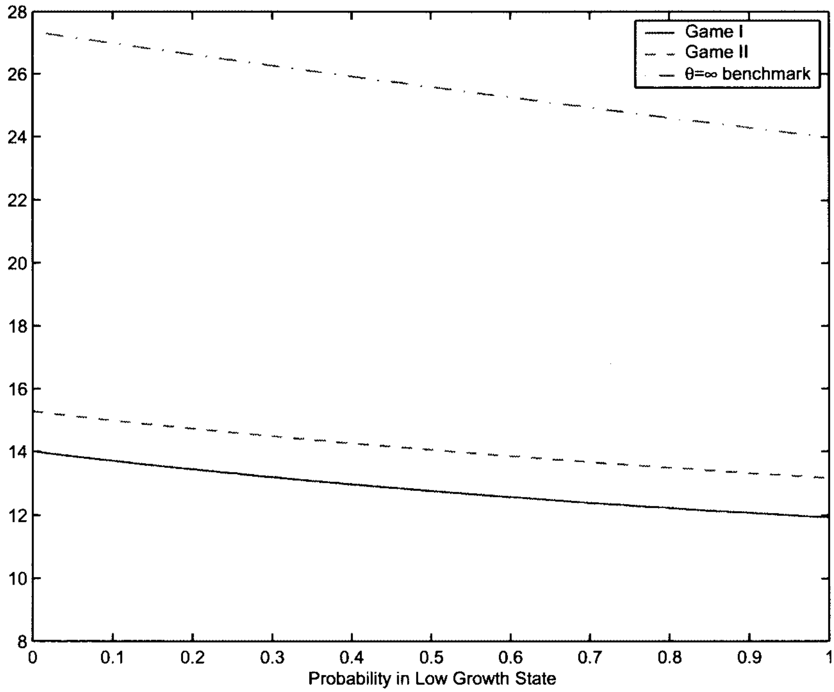

Figure 9
This figure displays the price-earnings ratio as a function of the probability of being in the low state for HMM games I and II. For comparison, it also displays the price-earnings ratio for the HMM economy that abstracts from robustness ($\theta = \infty$). The functions are plotted by fixing capital at its median value.

model misspecification makes the security less attractive to investors. The corresponding prices for the full information robustness game are 16.43 for the high state and 14.25 for the low state.

Figure 10 shows how the price-earnings ratios vary with k (the capital-technology ratio). The relation is increasing for all of the robustness games. When capital is high relative to the technology, the equity asset is more valuable to investors. Price-earnings ratios are predicted to be more responsive to historical movements in the capital-technology ratio than to movements in growth-state probabilities. While both k and $\hat{p}$ respond to common Brownian motion shocks, the capital-technology channel is more potent than the signal extraction channel as a source of fluctuation for price-earnings ratios. This implication for the pricing of the infinite earnings stream is in contrast to the implied behavior of local prices. Recall that the local factor prices of the Brownian increments are very responsive to changes in growth-state probabilities, but in a nonlinear way.

The median price-earnings ratio in the data for our sample is 16.46. As seen in Figures 9 and 10, the robust decision games bring this level more

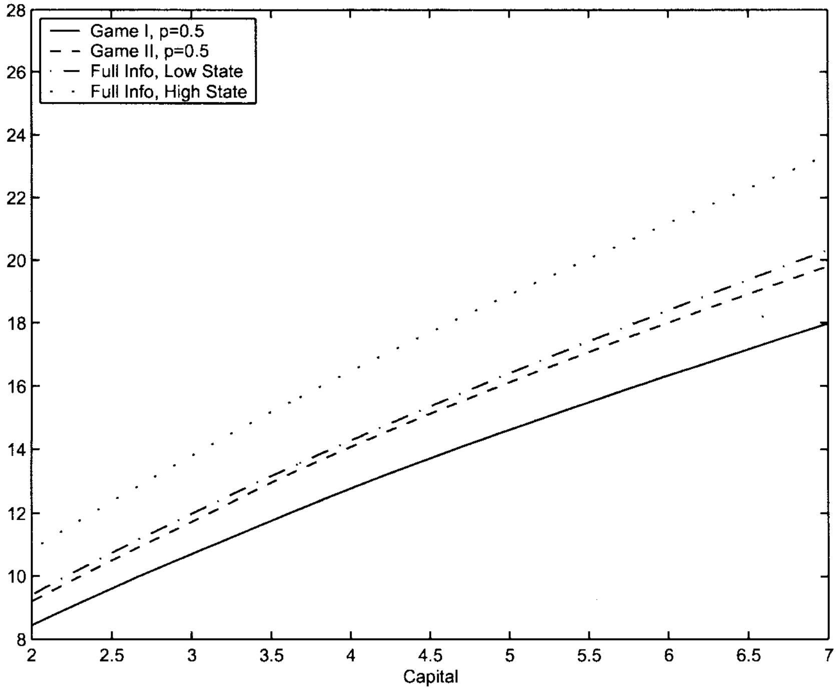

Figure 10
This figure displays price-earnings ratios as functions of the capital-technology ratio k for HMM games I and II. For comparison, it also reports the corresponding functions for the full information robustness game. The growth-state probabilities $\hat{p}$ are fixed at the designated values.

in line with these historical data. In fact, the hidden information games with the chosen values of θ have median price-earnings ratios that are too low. In contrast, the median price-earnings ratio for the $\theta = \infty$ benchmark is much too large.

While robustness delivers an empirically plausible downward shift in the price-earnings ratio, the model-based time-series trajectories do not track well some of the movements in the actual price-earnings ratios. The model-based, price-earnings trajectories are dominated by changes in the capital-technology ratio k. The capital-technology ratio in our model is stationary, with increases in recession states followed by decreases in high-growth states. Thus the model predicts that the long boom evident in our data from 1992 onward should result in a falling capital-technology ratio and therefore a corresponding drop in the price-earnings ratio. This predicted decline is displayed in Figure 11, which plots the price-earnings time series implied by the full information and HMM robustness games. During this period the capital-technology ratio fell in the data as in the model. However, during much of this time span the stock market was booming and the price-earnings ratios

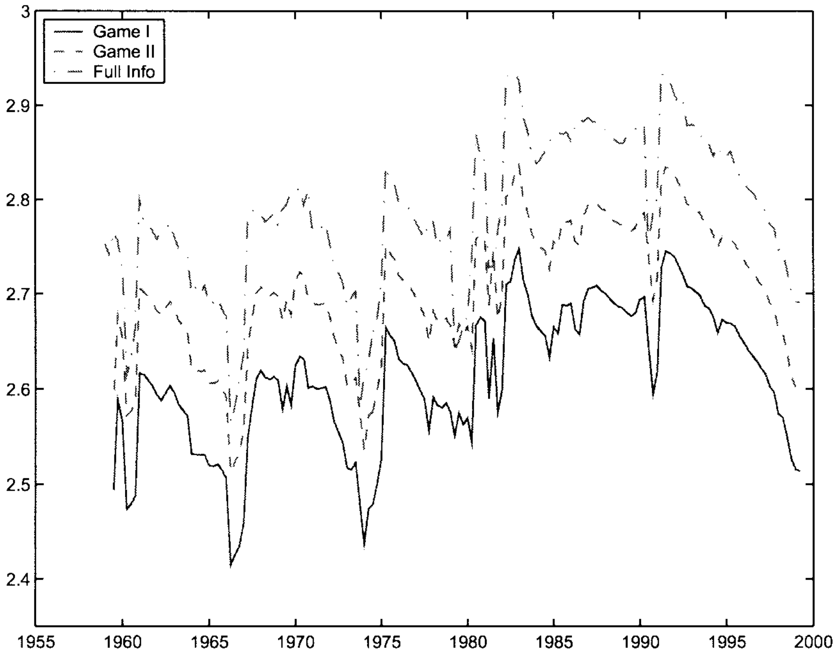

Figure 11
This figure plots the time series for the logarithm of the price-earnings ratio implied by models of robust decision making.

in the data increased dramatically. Clearly other factors outside of the model would be necessary to explain this dramatic runup in asset prices. Moreover, actual price-earnings ratios are much more variable than those implied by our models.[21] Nearly all of the variability in the price-earnings ratio in our model reflects changes in the capital-technology ratio. The models' price-earnings ratios inherit the smoothness in the capital-technology trajectory. Perhaps an approximating model with lower-frequency movements in technology growth rates would enhance the role of growth-state probabilities in price-earning ratios. HMMs of business cycle growth, however, are apparently ill equipped to explain price-earnings ratios even when robustness is added.

9. Conclusion

This article explores how a representative decision-maker's concern about model misspecification embeds itself in the time-series evolution for quantities and prices. We considered an environment in which there are abrupt

[21] The volatility in the logarithm of the price-earnings ratio in the data is four to five times that of our models.

movements in the time series and decision makers use historical data to update their beliefs about these movements. Abrupt movements are formal changes in the growth rates of technology and earnings in a stochastic growth model. We find the following:

- Factor "risk prices" encode model uncertainty premia in the model of robust decision making. These factor prices are largest when, under the approximating model, investors are most unsure of the hidden state. Thus there is an intriguing interaction between ambiguity, as reflected by state probabilities, and model uncertainty, as encoded in the local factor prices.
- While a concern about robustness brings the overall level of model-based price-earnings ratios closer to that in postwar data, the time series trajectories implied by the model do not track well the postwar data.

We nest an HMM rational expectations model inside our models of investor robustness by setting $\theta = \infty$. Thus some of conclusions carry over to this rational expectations model. Factor risk prices are predicted to be largest when investors are most unsure of the hidden state. Even abstracting from the overall level, the rational expectations model-based price-earnings ratio does not track well that observed in postwar data. The HMM information structure we used apparently does little to confront the well-known excess volatility in security prices.

We have used a stochastic growth model common in macroeconomics and finance. This stylized model has been an important benchmark for macroeconomics, and so presents a good laboratory for understanding how a concern about misspecification can alter implications. On the other hand, the stochastic growth model has known empirical deficiencies. Allowing for robustness repairs only some of these shortcomings. Richer transient dynamics and possibly multiple sectors and consumers are needed to produce models with better empirical underpinnings. In particular, richer learning dynamics may result in quantitatively important asymmetries in uncertainty premia in expansions and recessions. For example, Chalkley and Lee (1998) and Veldkamp (2000) build models in which business cycle asymmetries are linked to learning about an unobserved state. The asymmetry is driven by the interaction between risk aversion and private information in the Chalkley and Lee (1998) model. A concern about model misspecification could well imitate an enhancement of risk aversion in this environment and amplify the asymmetry between recessions and booms. Veldkamp shows how investors' decisions to initiate and terminate projects create more public information about the hidden state in the economy in good times than bad times. Introducing this feature into a model like ours could make uncertainty premia vary more between low-growth and high-growth regimes.

Appendix A: Generators

To represent the robustness games, we will make use of the generator of Markov processes. It is therefore useful to define it here briefly.[22] In our applications the Markov state contains either (k, s, Y^*) in the full information benchmark or $(k, \hat{p}, Y^*)$ in the hidden information case.

A Markov process can be specified in terms of the transition probabilities. Associated with each transition interval is a conditional expectation operator. The family of such operators is a Feller semigroup. The time derivative of the semigroup at the zero interval gives the generator of the semigroup. The generator captures the local evolution of the process. Since the family (semigroup) of conditional expectation operators can be built from the generator, we may model a Markov process by specifying its generator.

A.1 Full information

In the full information case, the state space is $\mathbb{X} = \mathbb{R}^+ \times \mathbb{S} \times \mathbb{R}^+$, where $\mathbb{R}$ is the real line, $\mathbb{R}^+ = (0, +\infty)$, and $\mathbb{S}$ is the collection of coordinate vectors in $\mathbb{R}^n$. Consider a function f mapping this state space into the real line $\mathbb{R}$. This function can be thought of equivalently as a $\phi : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^n$, where each coordinate function of ϕ is matched to a state s. If we begin with a function $f(k, s, Y^*)$ we may form

$$\phi(k, Y^*) = \overrightarrow{f(k, s, Y^*)}.$$

The $\vec{}$ operation stacks $f(k, s, Y^*)$ for the n values of s.

To construct a Feller semigroup of conditional expectation operators we consider the space of functions that are continuous on the one-point compactification of the state space $\mathbb{X}$. The collection of the restrictions of these functions to $\mathbb{X}$ is denoted by $\widehat{C}$. The conditional expectation operator for interval τ can be represented as

$$\mathcal{T}_\tau f(k, s, Y^*) = E\big[f(k_\tau, s_\tau, Y^*_\tau) | k_0 = k, s_0 = s, Y^* = Y^*_0\big]$$

for any $\tau \geq 0$ and for any f in $\widehat{C}$. The generator of the semigroup is the time derivative computed using the metric induced by the *sup* norm in (k, Y^*):

$$\mathcal{G} f = \lim_{\tau \downarrow 0} \frac{\mathcal{T}_\tau f - f}{\tau}$$

for each of the n values of s.

The generator is composed of two parts, one relative to the diffusion process for k and Y^*, and one relative to the jump process s. The generator of the bivariate process for k and Y^* (holding s fixed) can thus be depicted in terms of a drift μ and σ. μ and σ are column vectors containing, respectively, the drifts and the diffusion coefficients of k and Y^*. The exact forms of μ and σ depend on whether we are considering the nonrobust or the robust decision problems, since the evolution of these processes can be perturbed in the robustness games.

At least on the space C^2_K of twice continuously differentiable (in k and Y^*) functions with compact support in the interior of $\mathbb{R}^+ \times \mathbb{R}^+$, the generator is given by a second-order differential operator:

$$\mathcal{D} f = \mu \cdot \nabla f + \frac{1}{2} \sigma' \Delta f \sigma,$$

where ∇f is the gradient and $\triangle f$ the Hessian matrix (with respect to k and Y^*).

[22] For a rigorous treatment, see Ethier and Kurz (1986).

The generator for the jump process s can be depicted in terms of the matrix A as

$$\mathcal{A}f(s) = s'A\overrightarrow{f(s)}.$$

The composite generator under complete information is formed by adding the two components,

$$\mathcal{G}f = \mathcal{D}f + \mathcal{A}f.$$

A.2 Partial information

In the case of partial information, the state variables are $x = (k, \hat{p}, Y^*)$ and the state space is $\mathbb{X} = \mathbb{R}^+ \times \mathbb{P} \times \mathbb{R}^+$, where $\mathbb{P}$ is an $n-1$-dimensional set of $\mathbb{R}^n$ appropriate for a vector of nondegenerate probabilities. The semigroup is again defined using a one-point compactification. Since x is now a diffusion, the resulting generator, denoted $\widehat{\mathcal{G}}$, is

$$\widehat{\mathcal{G}}f = \hat{\mu} \cdot \nabla f + \frac{1}{2}\hat{\sigma}'\Delta f\,\hat{\sigma},$$

where ∇f and Δf are the gradient and the Hessian with respect to the composite state vector $(k, \hat{p}, Y^*)$, and $\hat{\mu}$ and $\hat{\sigma}$ are the column vectors containing the drift and the diffusion coefficients. The exact expressions for $\hat{\mu}$ and $\hat{\sigma}$, depend on the robustness game, and were denoted in the article by μ_x^i and σ_x^i, $i = 1, 2$ depending on the game.

The domain of the Feller semigroup is sufficiently rich for the purposes of constructing Markov processes, but it is too confining for the purposes of control theory, since the value functions for our resource allocation games are unbounded. The domain of the generator of a semigroup may be extended to unbounded functions f by finding functions g such that

$$M_t = f(x_t) - f(x_0) - \int_0^t g(x_u)\,du, \tag{16}$$

$x \in \mathbb{X}$, is well defined and a local martingale. In this case we use the notation $\mathcal{G}f$ or $\widehat{\mathcal{G}}f$ to denote the function g used in this construction. $\mathcal{G}$ and $\widehat{\mathcal{G}}$ are called the *extended* generators of the corresponding Markov processes [see Ethier and Kurz (1986) and Davis (1993)].

Appendix B: HJB Equations

In the case of full information and no robustness, using the generator notation previously introduced, the resulting HJB equation is

$$\max_c Y^*U(c) + \mathcal{G}W(k, s, Y^*) = 0.$$

For the nonrobust, hidden information case, the equation is

$$\max_c\ Y^*U(c) + \widehat{\mathcal{G}}W(k, \hat{p}, Y^*) = 0.$$

In the robustness games, we introduce a minimizing player and an entropy penalty. The corresponding equation for game II is

$$\max_c \min_h Y^*\left[U(c) + \theta\frac{h^2}{2}\right] + \widehat{\mathcal{G}}W^1(x) = 0.$$

By substituting the formula for $\widehat{\mathcal{G}}$, we get Equation (12).

For game II, we have

$$\max_c \min_g Y^* \left[U(c) + \frac{\theta}{2} \sum_i (g^i)^2 \hat{p}^i \right] + \widehat{\mathscr{G}} W^2(x) = 0,$$

which results in Equation (13).

In all the cases, the value functions are linear in Y, $W(x) = Y^* V(k, \hat{p})$. This can be verified by substituting this guess in the HJB equations, working out the algebra, and noticing that Y^* can be factored out of the resulting equations. Numerically, therefore, one need to solve only for V, which depends on fewer state variables than W.

Given V, we can compute the decision rules for c and k, and the prices of risk and uncertainty, λ_m and λ_u, using the formulas presented in the text. Note that while λ_m only depends on the derivatives of V, the distortions g and h also depend on the level of V; therefore it is important to compute precisely both V and its derivatives.

Appendix C: The Algorithm to Compute the Value Functions

The HJB equations are second-order partial differential equations in V. To solve for V, we adapt the algorithm described by Candler (1999). The idea is to fix the decision rules, solve the resulting linear, second-order partial differential equations, update the decision rules with the new solution for the value functions, and iterate until convergence.

The algorithm is similar for all the games considered, so we will explain it here in general terms. Let z denote the arguments of V (k and s for the full information case, k and $\hat{p}$ for the hidden information ones), and i the control variables (c for the nonrobust full information case, c and h for game I, c and g for game II). Since we consider the case of only two possible values for the mean rate of growth of the technology, there is only one probability in z, and $\hat{p}_L = 1 - \hat{p}_H$. Let $\text{HJB}(V, i)$ be the differential equations described above, keeping the decision rule i fixed. Following Candler (1999), we explicitly introduce time and solve

$$\frac{\partial V(z, t)}{\partial t} + \text{HJB}(V(z, t), i) = 0.$$

This corresponds to the backwards iteration often used to solve the dynamic programming problem. We start at time T with some guess $V(z, T)$ and then solve the problem backwards one time period.

The algorithm therefore consists of the following:

1. Start with a guess $V(z, T)$.
2. Given $V(T, z)$, compute the optimal decision rules i_T.
3. Given i_T, solve the second-order partial differential equation backwards one time interval Δt, obtaining the new value function $V(z, T - \Delta t)$,

$$\frac{\partial V(z, T - \Delta t)}{\partial t} + \text{HJB}(V(z, T - \Delta t), i_T) = 0.$$

   The equation is linear and we use an implicit, upwind, finite difference method.
4. Given $V(z, T - \Delta t)$, compute the optimal decision rules $i_{T-\Delta t}$ and iterate until the distance between $V(z, t)$ and $V(z, t - \Delta t)$ is small.

As mentioned in the previous section, we can use the numerical solution for V to compute the decision rules and the prices of risk. The derivatives of V necessary to evaluate these variables are computed numerically using central differences. The numerical solution for V is also used as an input for the pricing equation [Equation (15)], which is then solved using a finite difference method.

References

Aiyagari, S. R., 1994, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics,* 109, 659–684.

Anderson, E., L. P. Hansen, and T. Sargent, 2000, "Robustness, Detection and the Price of Risk," mimeo.

Basar, T., and P. Bernhard, 1995, *H_∞-Optimal Control and Related Minimax Design Problems.* Boston: Birkhauser.

Baum, L. E., and T. Petrie, 1966, "Statistical Inference for Probabilistic Functions of a Finite State," *Annals of Mathematical Statistics,* 37, 1554–1563.

Boldrin, M., L. Christiano, and J. Fisher, 2001, "Habit Persistence, Asset Returns and the Business Cycle," *American Economic Review*, 91, 149–166.

Bonomo, M., and R. Garcia, 1996, "Consumption and Equilibrium Pricing: An Empirical Assessment," *Journal of Empirical Finance*, 3, 239–265.

Breeden, D., 1979, "An Intertemporal Asset Pricing Model With Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

Brock, W. A., and L. J. Mirman, 1972, "Optimal Economic Growth and Uncertainty: The Discounted Case," *Journal of Economic Theory*, 4, 479–513.

Cagetti, M., L. P. Hansen, T. J. Sargent, and N. Williams, 2000, "Robust Recursive Prediction and Control," mimeo.

Candler, G. V., 1999, "Finite Difference Methods for Continuous Time Dynamic Programming," in R. Marimon and A. Scott (eds.), *Computational Methods for the Study of Dynamic Economies*, Oxford University Press, Oxford, pp. 172–194.

Cecchetti, S. G., P. Lam, and N. C. Mark, 1993, "The Equity Premium and the Risk-Free Rate," *Journal of Monetary Economics*, 31, 21–45.

Chalkley, M., and I. H. Lee, 1998, "Learning and Asymmetric Business Cycles," *Review of Economic Dynamics*, 1, 623–645.

Chen, Z., and L. Epstein, 2001, "Ambiguity, Risk and Asset Returns in Continuous Time," forthcoming in *Econometrica*.

Chernoff, H., 1952, "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations," *Annals of Mathemtatical Statistics*, 23, 493–507.

Cochrane, J., and L. P. Hansen, 1992, "Asset Pricing Explorations for Macroeconomics," in O. J. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*, MIT Press, Cambridge, MA, pp. 115–165.

Cochrane, J. M., 1991, "Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations," *Journal of Finance*, 46, 209–237.

David, A., 1997, "Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility," *Journal of Financial and Quantitative Analysis*, 32, 457–462.

David, A., and P. Veronesi, 1999, "Option Prices with Uncertain Fundamentals: Theory and Evidence on the Dynamics of Implied Volatilities and Over-underreaction in the Options Market," mimeo.

Davis, M. H., 1993, *Markov Models and Optimization*, Chapman & Hall, London.

Dembo, A., and O. Zeitouni, 1986, "Parameter Estimation of Partially Observed Continuous Time Stochastic Processes via the EM Algorithm," *Stochastic Processes and Their Applications*, 23, 91–113.

Detemple, J., 1986, "Asset Pricing in a Production Economy with Incomplete information," *Journal of Finance*, 41, 383–390.

Dupuis, P., M. R. James, and I. Petersen, 1998, "Robust Properties of Risk Sensitive Control," LCDS 98-15, Brown University.

Elliott, R. J., L. Aggoun, and J. B. Moore, 1995, *Hidden Markov Models. Estimation and Control*, Springer-Verlag, New York.

Ethier, S. N., and T. G. Kurz, 1986, *Markov Processes. Characterization and Convergence*, Wiley, New York.

Fellner, W., 1965, *Probability and Profit: A Study of Economic Behavior Along Bayesian Lines.* Richard D. Irwin, Homewood, Ill.

Hamilton, J. D., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Hamilton, J. D., 1990, "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics*, 45, 39–70.

Hansen, L. P., and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

Hansen, L. P., T. Sargent, and T. Tallarini, 1999, "Robust Permanent Income and Pricing," *Review of Economic Studies*, 66, 873–907.

Hansen, L. P., T. J. Sargent, G. A. Turmuhambetova, and N. Williams, 2001, "Robustness and Uncertainty Aversion," mimeo.

Hansen, L. P., T. J. Sargent, and N. E. Wang, 2002, "Robust Permanent Income and Pricing with Filtering," forthcoming in *Macroeconomic Dynamics*.

Hansen, L. P., and K. J. Singleton, 1983, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, 249–265.

Huber, P. J., 1981, *Robust Statistics*, Wiley, New York.

James, M. R., 1992, "Asymptotic Analysis of Nonlinear Stochastic Risk Sensitive Control and Differential Games," *Mathematics of Control, Signals, and Systems*, 5, 401–417.

James, M. R., and J. Baras, 1996, "Partially Observed Differential Games, Infinite Dimensional Hamilton-Jacobi-Isaacs Equations, and Nonlinear H_∞ Control. *SIAM Journal on Control and Optimization*, 34, 1342–1364.

James, M. R., J. Baras, and R. J. Elliott, 1994, "Risk-Sensitive Control and Dynamic Games for Partially Observed Discrete-Time Nonlinear Systems," *IEEE Transactions on Automatic Control*, 39, 780–791.

Krusell, P., and A. Smith, Jr., 1998, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Liptser, R. S., and A. N. Shiryayev, 1977, *Statistics of Random Processes*, Springer-Verlag, New York.

Maenhout, P. J., 1999, "Robust Portfolio Rules and Asset Pricing," mimeo.

Merton, R. C., 1975, "An Asymptotic Theory of Growth Under Uncertainty," *Review of Economic Studies*, 42, 375–393.

Rouwenhorst, K. G., 1995, "Asset Pricing Implications of Equilibrium Business Cycle Models," in T. F. Cooley (ed.), *Frontiers of Business Cycle Theory*, Princeton University Press, Princeton, N.J., PP. 294–330.

Sclove, S. L., 1983, "Time-Series Segmentation: A Model and a Method," *Information Sciences*, 29, 7–25.

Stock, J., and M. Watson, 1999, "Business Cycle Fluctuations in US Macroeconomic time series," in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Amsterdam, Elsevier.

Veldkamp, L. L., 2000, "Slow Boom, Big Crash," manuscript.

Veronesi, P., 1999, "Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model," *Review of Financial Studies*, 12, 976–1007.

Veronesi, P., 2000, "How Does Information Quality Affect Stock Returns?" *Journal of Finance*, 55, 807–837.

Watson, M. W., 1993, "Measures of Fit for Calibrated Models," *Journal of Political Economy*, 101, 1011–1041.

Wonham, W. J., 1964, "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering," *SIAM Journal on Control*, 2, 347–368.