Robustness and Pricing with Uncertain Growth

Marco Cagetti, Lars Peter Hansen, Thomas Sargent and Noah Williams[1]

February 20, 2001

Abstract. We construct models of robust decision-making and pricing when contemporaneous big and small shocks hit a stochastic-growth economy. Large shocks are infrequent changes and small shocks are continuous movements in the technology process. Large shocks evolve as a Markov jump process. Small shocks are a Brownian motion. Decision makers treat models as approximations and fear misspecification. Robust decision-making is formalized as a two-player game. To promote robustness, investors imagine that a malevolent player perturbs a baseline model. We study two economies, each of which decentralizes a robust resource allocation problem with hidden growth rates. The economies differ in how the same base line model is viewed as an approximation. We compare prices to those for an economy in which the growth state is fully revealed. We study the time-series implications for the measured risk-return tradeoff and the price-dividend ratio.

1 Introduction

This paper shows how decision makers' concerns about the robustness of their decisions to 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), inspiring investors to make inferences about the growth rate. They gather their opinions about the growth rate from current and

[1]Lars Peter Hansen: Department of Economics, University of Chicago (l-hansen@uchicago.edu). Thomas Sargent: Department of Economics, Stanford University (sargent@leland.stanford.edu). Marco Cagetti: Department of Economics, University of Virginia (cacio@virginia.edu). Noah Williams: Department of Economics, University of Chicago (nwilliam@midway.uchicago.edu).
Comments from John Heaton and Grace Tsiang are gratefully acknowledged. Both Hansen and Sargent were supported by the National Science Foundation while completing this project.

past observations of technology, which are clouded by concurrently evolving small shocks that are Brownian motion movements. We show how the investors' desire to make their decision rules robust to misspecification of the evolution of technology alters the evolution of security prices and 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, Hansen, Sargent, and Williams (2000). An investor maximizes a discounted utility function, but fears that a baseline forecasting model is misspecified. A second player is introduced as a device to make the decisions perform well under a variety of possible model misspecifications. The second player minimizes the decision maker's objective function by distorting the baseline model. To restrict the distortions, the objective function of the second player penalizes a measure of the discrepancy between the baseline model and the perturbed model. We allow only perturbations that are small and hard to detect statistically. Formally, this leads us to a log-likelihood-based (relative entropy) penalty modified to account for the HMM structure. We refer to the fictitious second agent as the *malevolent* agent.

The second player can seem mysterious because he exists only in the decision maker's mind. The two-player, zero-sum game is the decision maker's way to induce robustness. While every decision maker is arguably a fiction in any economic model, our second malevolent one helps the decision maker explore the consequences of model misspecification. As argued by Huber (1981) in his discussion of an optimal robust statistical procedure:

> ... as 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.

Economists use a max-min formulation when they use Lagrange multipliers. Lagrange multipliers allow us to convert a constrained maximization problem into an unconstrained max-min 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.

Although we use ideas from standard robust control theory, we modify them because, as is common in economic theory, our baseline models are stochastic and the two-player games are formulated in a fully recursive way that allows for discounting.[2] A stochastic baseline model is essential for our application to finance. In emphasizing this structure, we follow Hansen, Sargent, and Tallarini (1999), Anderson, Hansen, and Sargent (2000) and Maenhout (1999).

We apply the games to a continuous-time version of a stochastic growth economy. A technology shock process serves as a stochastic forcing process, modeled as an HMM. This specification reflects the post war US experience, characterized by Hamilton (1989) and others, of short recessions and sustained booms. Because it can be difficult to distinguish these two movements from the observed technology level, we explore what happens when we make the growth state hidden from the decision-maker. Our specification of the technology shock process formally follows Wonham (1964), David (1997) and Veronesi (1999). The latter two papers study pricing in production economies with linear technologies and dividend growth processes that are hidden.[3] The introduction of hidden information alone does not enhance the market price of risk and the price-dividend ratios into ranges that accord to data. For this reason we turn to robustness as means of distorting the beliefs of the decision-makers.

In confronting an uncertain future, a robust decision-maker uses an endogenously-determined pessimistic assessment of the underlying probability model. This pessimism formally captures Fellner (1965)'s conjecture that the probabilities used in actual decision-making might deviate from those implied by a tractable model that is used to quantify risks. Fellner saw decision makers conservatively adjusting probabilities in context specific ways. In our model, investment decisions can appear to be based on a model in which the duration of low growth states is longer than those quantified by looking at historical time series data (as is done when imposing rational expectations). This motive augments the usual precautionary mechanism coming from the baseline stochastic model. As in Hansen, Sargent, and Tallarini (1999), the effects of a preference for robustness on saving-consumption profiles and therefore on equilibrium quantities can be offset (at least approximately) by a change in the subjective rate of discount. But also as in Hansen, Sargent, and Tallarini (1999), equilibrium prices are affected by such joint alterations in the discount and robustness parameters.

[2] Among the few examples of stochastic robust control models are James (1992), who studies deterministic differential games of robust decision-making as small noise limits, and Dupuis, James, and Petersen (1998). Examples of robust control of deterministic systems with partial observations are James, Baras, and Elliott (1994) and James and Baras (1996).

[3] In particular, 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.

We decentralize the robust version of the stochastic growth model by computing shadow prices from a robust resource allocation problem. Given the continuous-time nature of the model, we can build up asset prices for intervals of time from local or instantaneous prices. These local prices include both the instantaneous interest rate and the risk price of the Brownian motion increment. In our models, the factor risk price has an additional component attributable to a concern about model misspecification. This term also occurs in the continuous-time papers of Anderson, Hansen, and Sargent (2000) and Chen and Epstein (1999), and as an approximation in the discrete-time analysis of Hansen, Sargent, and Tallarini (1999). The hidden growth state puts an extra source of variation into the risk premia. Risk premia fluctuate in part because of changes in the beliefs about the mean rate of growth.

The paper is structured as follows. Section 2 presents the economic environment. Section 3 describes the information structure and the signal extraction problem. Section secdistort describes the model distortions and measures of model misspecification. Section 5 presents the value function differential equations that must be solved for the hidden information games. Section 6 discusses the implications for time series of capital stocks. Section 7 uses the link between statistical detection and robustness to restrict the degree of robustness in the asset calculations. Section 8 shows how risk-return tradeoffs change over time. Section 9 displays the implied dividend-price ratios.

2 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 a 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';
- the evolution of price-dividend ratios.

2.1 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 valued-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 (1999) 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 (1999) and Hansen and Singleton (1983),[4] 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), Anderson, Hansen, and Sargent (2000), Chen and Epstein (1999) and Maenhout (1999), 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.

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 Veronesi (2000) and David and Veronesi (1999) except that we have an additional state variable (capital) and also explore implications for robustness.

2.2 Technology

We assume a Cobb-Douglas production function

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

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

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, $U_1, U_2, ..., U_n$, 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).

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

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

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.

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 (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:

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

2.3 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.[5] 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 $\tau = 0$.

3 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 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 justifies separation are described in Cagetti, Hansen, Sargent, and Williams (2000).

3.1 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 the 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), Veronesi (1999), Veronesi (2000) and David and Veronesi (1999).

[5] 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.

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 linear stochastic differential equation:

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

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

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

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

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

and the evolution of k is:

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

where

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

3.2 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 of being in the low growth state computed using our baseline estimates. Under our baseline model, the post WWII 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 (GDP less farm, housing and government), capital (interpolation of annual values of fixed

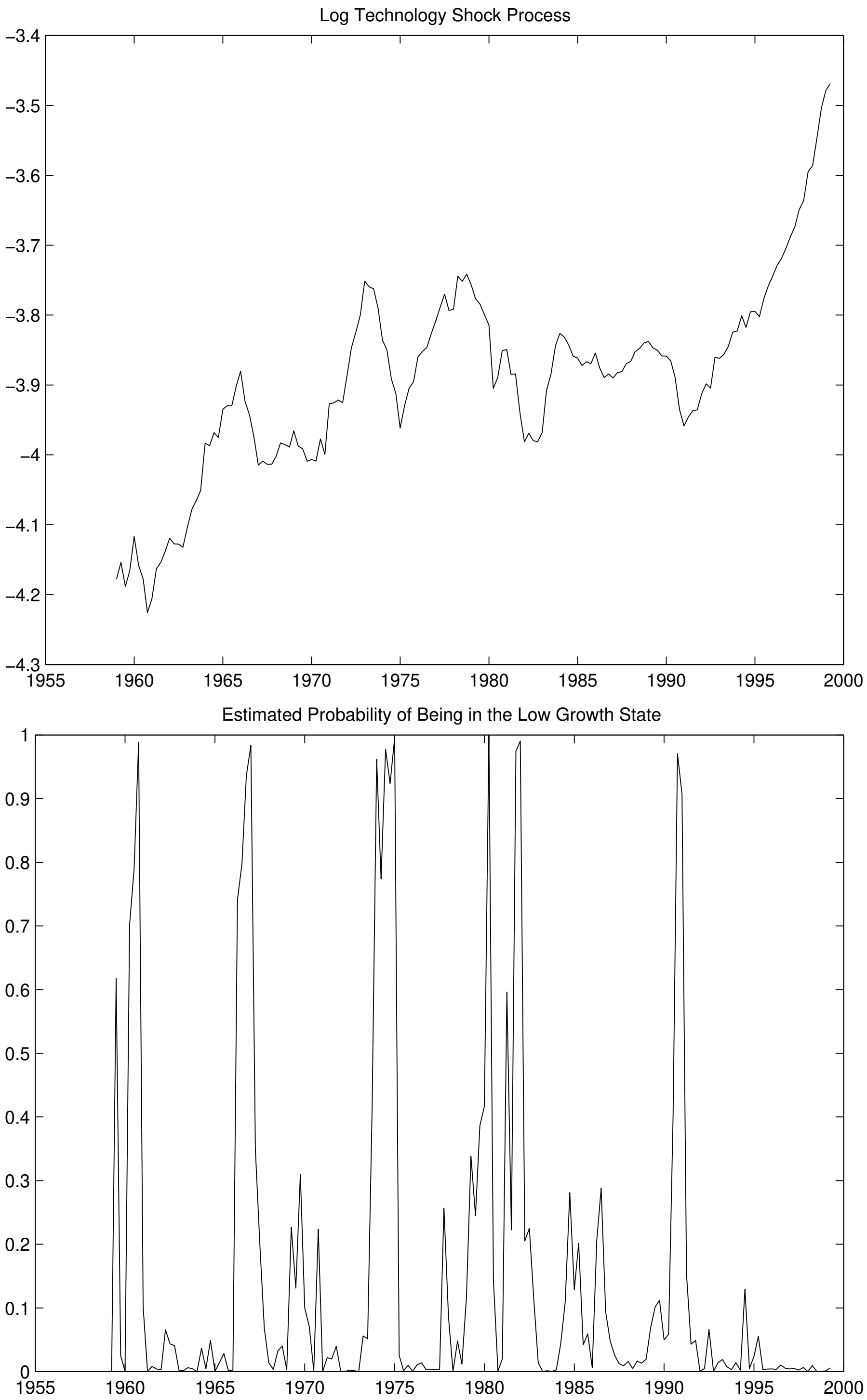

Figure 1: The top panel displays the logarithm of the technology level (the cumulation of the Solow residual) and the bottom panel displays the probability of being in the low growth state.

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

Table 1: This table contains the HMM model parameters. The parameter values were obtained by estimating the HMM model using time series data on the Solow residual.

non-residential capital stock using quarterly investment), and labor (hours of employees on non-agricultural 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 (4). Since there are only two states, $1/a_{12}$ is the average length of expansions, and $1/a_{21}$ 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, Hansen, Sargent, and Williams (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.[6]

4 Model Distortions and Approximation Measures

We describe two decision problems with hidden information and a full information benchmark. The hidden information problems share one common feature. The vector $\hat{p}_t$ constructed by updating probabilities with past data is used as an exogenous state vector for the decision problem. It is constructed using the filtering equation implied by the approximating model. Hence $\hat{p}_t$ becomes an integral part of the approximating model around which we consider perturbations.

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

Our formulation differs from others in the robust control theory, such as Basar and Bernhard (1995). In their framework, the backwards-looking filtering of data to estimate the current state depends on past utility rewards. In our formulation, the date t decision-makers use past data to make inferences without regard to past contributions to utility. Instead they are forward-looking in a recursive manner. The decision-makers in the current date care only about utility contributions in the current and future time periods. This perspective is familiar from derivations of Hamilton-Jacobi-Bellman equations for control problems and is at the heart of the recursive formulations of preferences. Robustness, however, adds a wrinkle because it requires a form of probability slanting that depends in part on the objective of the decision-maker. The date t perspective used here leads us to a criterion function for the decision maker and the malevolent agent different from the one typically used in the literature on robust control with hidden information. See Cagetti, Hansen, Sargent, and Williams (2000) for justification and elaboration of the approach taken here.[7] The outcome of our formulation is that $\hat{p}_t$ formed using the approximating model is a reference point in making robust adjustments to beliefs.

Our two hidden information games differ in the perturbations that are entertained. The specification of each game requires a statement of the perturbations and a measure of their discrepancy from the baseline model. For each game, we describe a limited array of structured perturbations. In each case this structure can be justified from a more primitive starting point.[8]

4.1 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 game is $(k, \hat{p}, y)$, which evolves according to (8) and (5). The stochastic evolution of these state vectors is governed by a single Brownian motion technology shock $d\hat{B}_t$. We disguise model misspecification within the Brownian motion shock. This makes the statistical discrimination between the approximating model and its competitors difficult. Thus we replace: $d\hat{B}_t$ by $d\hat{B}_t + h_t dt$ where the process h_t is a process adapted to the

[7]This date t perspective is best conceived as a dynamic game with different decision-makers at each date that have overlapping objectives. The standard robust control theory is premised on a form of commitment to an initial way of slanting probabilities that depends on the initial (date zero) objective.

[8]Thus our specification differs from the literature on structured uncertainty which imposes limits on the perturbations from the outset. Our structured perturbations arise as the outcome of an analysis of more general unstructured uncertainty. See Cagetti, Hansen, Sargent, and Williams (2000) for details.

filtration $\mathcal{Y}_t$, and we alter evolution equations (5) and (8) to be:

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

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.

We require a quantitative measure of model misspecification h. Following Anderson, Hansen, and Sargent (2000), we use a likelihood ratio or conditional entropy penalty. A drift hdt appended to a 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.

4.2 Hidden Information, Representation II

We now consider a second representation of the hidden information that was developed in Cagetti, Hansen, Sargent, and Williams (2000) and that is the nonlinear counterpart to one explored by Hansen, Sargent, and Wang (2000). Instead of hiding perturbations in the reduced information Brownian motion $d\hat{B}_t$, 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. This vector adds a drift to the full information Brownian motion. 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

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

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:

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

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 HMM's, 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

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

Table 2: Distortions for two representations with hidden information.

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 imposed on the relation between distortions can be defended from a more primitive starting point. (See Cagetti, Hansen, Sargent, and Williams (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.

4.3 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 $\hat{p}$ exceeds one half and to the high growth state when $\hat{p}$ is greater than one.

5 Value Functions

We have described representations of information for a 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.

5.1 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.

5.2 Hamilton-Jacobi-Bellman Equations

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

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

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

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

where ∇ is the gradient. This equation formalizes the fact that the value function W^1 should have a local mean given by negative of the discounted instantaneous contribution

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

appropriately optimized. The θ parameter controls the degree of robustness that is sought. Setting θ to ∞ gives the efficient resource allocation without concerns about model misspecification.

For Game II we replace (9) with (10). Write the stochastic evolution as:

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

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:

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

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.

5.3 Computations

The value functions for both of these games with hidden information will be linear in the stochastic discount factor. That is, they will 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 (12) and (13) are both linear in Y^*.

For both games, consumption will satisfy:

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

Moreover, objective (12) is quadratic in the scalar h and (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 is described in Appendix C.

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.

6 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 particular, our solution to Game I (and of course the solution to the growth model without robustness) could be reinterpreted as a *full information* game with a singular multivariate diffusion model for the technology shock. 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.

6.1 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 to 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.

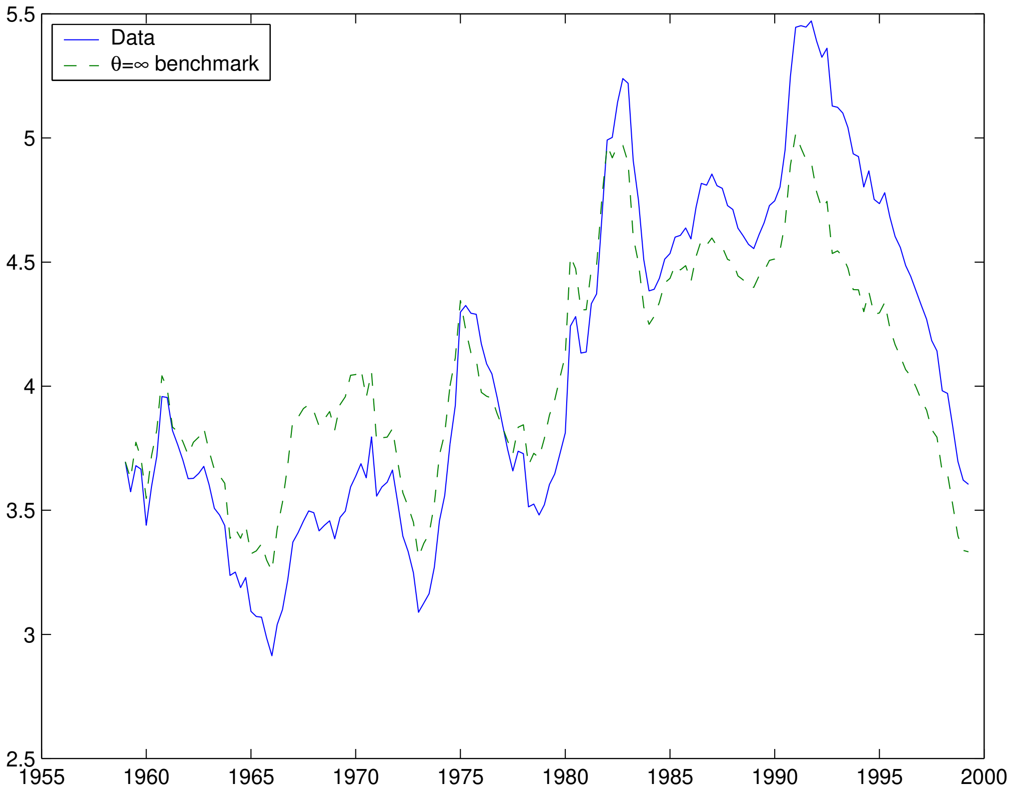

Figure 2: This figure displays the observed time series of the capital/technology ratio k and the ratio implied by a HMM benchmark model with $\theta = \infty$ (i.e., no preference for robustness).

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 no surprise since the parameter configurations were selected in part to match the growth features of the model.[9]

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 non-robust 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

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

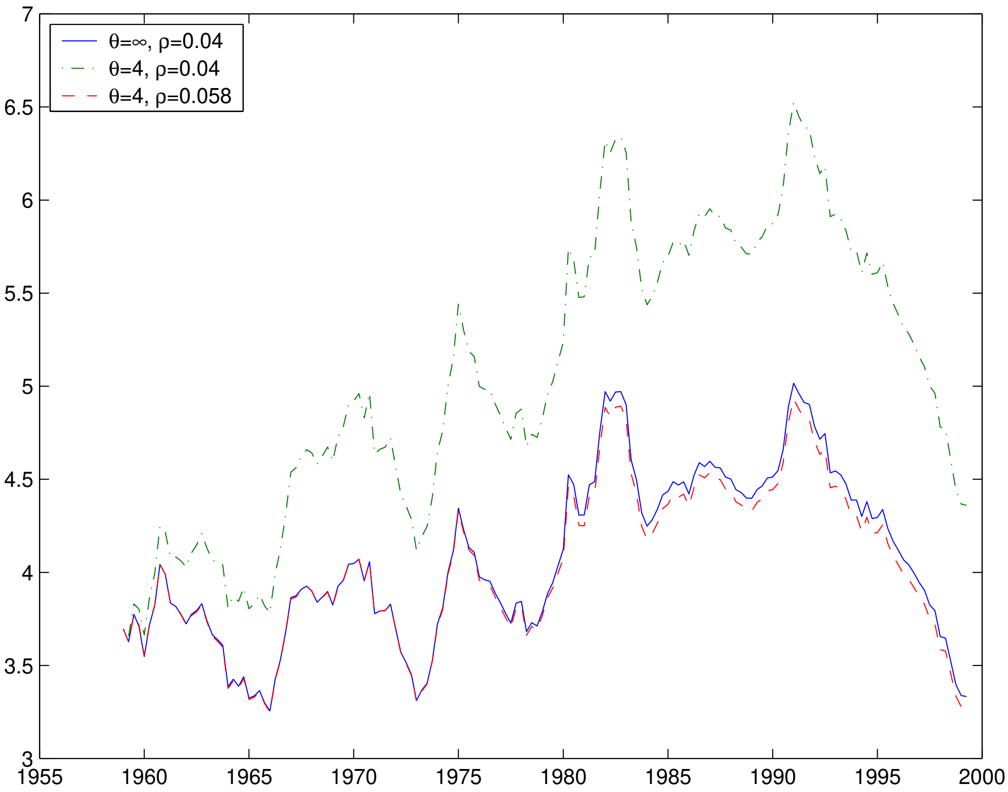

Figure 3: This figure displays the capital technology trajectories implied by different values of the subjective discount rate (ρ) and the robustness parameter (θ). Introducing robustness results in additional capital accumulation holding fixed the discount rate. This is seen by comparing the $(\theta, \rho) = (\infty, 0.04)$ 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.

by the possibility of bad shocks coming from a given and pre specified probability distribution. These effects are known to be small when calibrated to macroeconomic measures of uncertainty.[10] In our robust setup, the decision-maker considers also the potential misspecification of his approximating 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 baseline 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.

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

Macroeconomic calibrations of parameters based on mean growth rates and average returns typically ignore the impact of precautionary savings.[11] 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 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 (2000) 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 non-robust ($\theta = \infty$) 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).

6.2 Other Quantity Implications

The baseline 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 model suggests that the risk premia are likely to be larger than in a model with more plausible consumption variability. The aim of these exercise is to explore how a concern about robustness changes the implications of a well known and pedagogically valuable model. Unfortunately, we inherit some empirical deficiencies of the baseline stochastic growth model that can only be fixed by exploring other complications in the economic environment. While the HMM framework provides us with an interesting laboratory to study how robustness alters the dynamic equilibrium models, it does not repair some of the well known empirical deficiencies in macroeconomic and financial models.

[11] 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.

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

Table 3: 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 average of the conditional entropies are: .00368 for HMM Game I, .00372 for HMM Game II, and .00375 for the full information benchmark.

7 Tuning Robustness

The parameter θ is used to govern the extent of robustness in both Games I and II. Thus we have two families of hidden information games, each of which is indexed by θ. The parameter θ does not necessarily have the same meaning across games (also see Hansen, Sargent, and Wang (2000)). Since the asset pricing implications we will report are numerical in nature, we are compelled to focus on a limited number of parameter configurations.

In our calculations, we specify θ for each game based on detection-error considerations. Anderson, Hansen, and Sargent (2000) describe a link between robustness and detection for robust decision-problems in which the Markov state is fully observed. Their analysis applies for Game I but not Game II. Here we adopt a more informal approach. Consider first HMM Game I and the pair wise comparison between the $h = 0$ model and the $h = h^*$ model. The entropy term $\frac{1}{2}h^2$ can be viewed as the conditional expectation of the instantaneous contribution to a log-likelihood ratio. Thus cumulating this measure over the sample interval provides a statistical measure of discrimination.

For HMM Game II, matters are more complicated. 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.

Based on detection-error considerations, we employed the θ and ρ pairs reported in Table 3 in our subsequent calculations. These choices make the sample mean of the entropy remain about the same across games. We report the time series of entropy measures for both HMM games in Figure 4. To get a rough idea of statistical

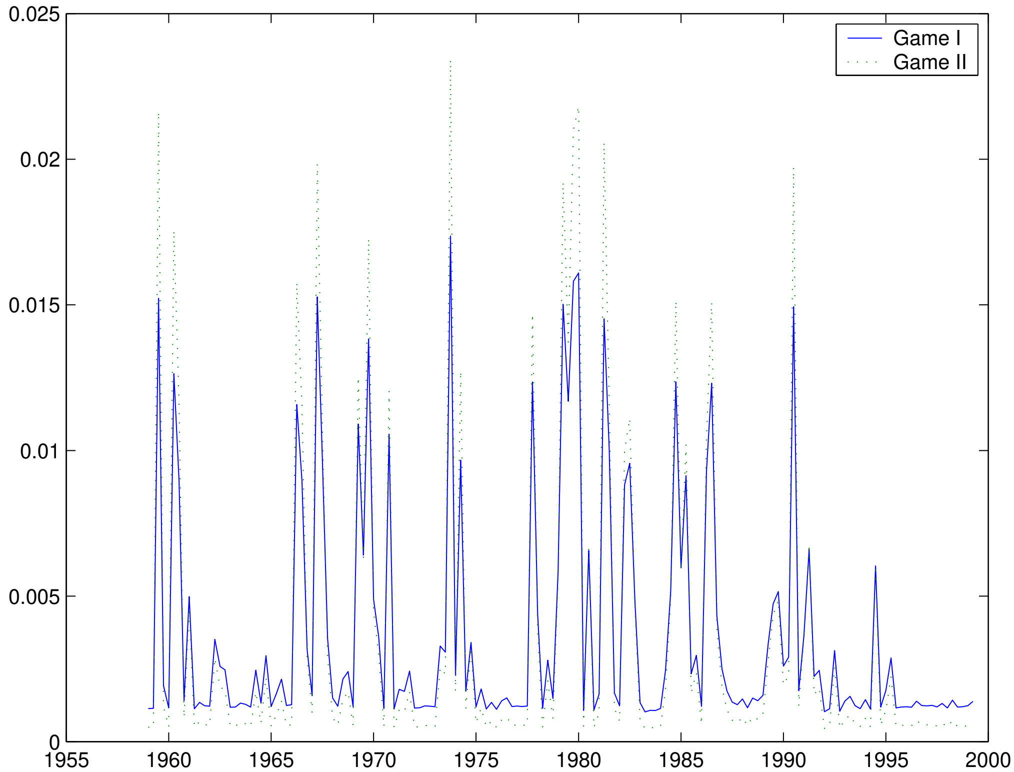

Figure 4: The top panel gives the time series for the conditional relative entropies for the two HMM games. For HMM Game I we report $\frac{(h^*)^2}{2}$ and for HMM Game II we report $\frac{\sum_i (g_i^*)^2 \hat{p}_i}{2}$.

discrimination, the sample mean can be multiplied by the number of time periods of data available for discrimination. The implied time unit for this figure is one quarter. Notice that there is more variability in the discrimination measures for HMM Game II than for Game I. A more formal analysis of discrimination could be based on large deviation calculations similar to those in Anderson, Hansen, and Sargent (2000) or on simulation as in Hansen, Sargent, and Wang (2000). We view our choice of robustness as only a rough guide and as a way to get rough comparability across games. There are other potential approaches for tuning the robustness parameter θ. While detection error analysis eliminates candidate models that should be easy to uncover, 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.

8 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.[12] In the calculations in this subsection 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.

[12] Also see Hansen, Sargent, and Tallarini (1999).

We study the (*local*) pricing of $d\hat{B}$ 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 diffusion coefficient (or factor loading) 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 $d\hat{B}$. Anderson, Hansen, and Sargent (2000) and Chen and Epstein (1999) show that the factor price of the Brownian increments $d\hat{B}$ can be decomposed into two prices: the usual price for risk, and a price of model uncertainty.[13] Thus the factor price λ is the sum

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

The price of risk λ_m is obtained by applying a formula from Breeden (1979)'s analysis of a consumption-based asset pricing model. This risk price is given by (minus) the weighting coefficient on $\hat{B}_t$ 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 formula (14) for the consumption-technology ratio c. The coefficient on $d\hat{B}$ is computed as the sum of partial derivatives with respect to y, $\hat{p}$ and k:

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

The model uncertainty price is $\lambda_u = -h^*$ for Game I and $\lambda_u = -\hat{p} \cdot g^*$ for Game II where h^* and g^* are the malevolent agent's solutions in the respective games. As we will see the extra dependence on the underlying state alters the time series properties for λ_u for Game II by making it more sensitive to the hidden state probability.

[13] Anderson, Hansen, and Sargent (2000) and Chen and Epstein (1999) 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 Game II solution and prices are not special cases of those reported in Anderson, Hansen, and Sargent (2000).

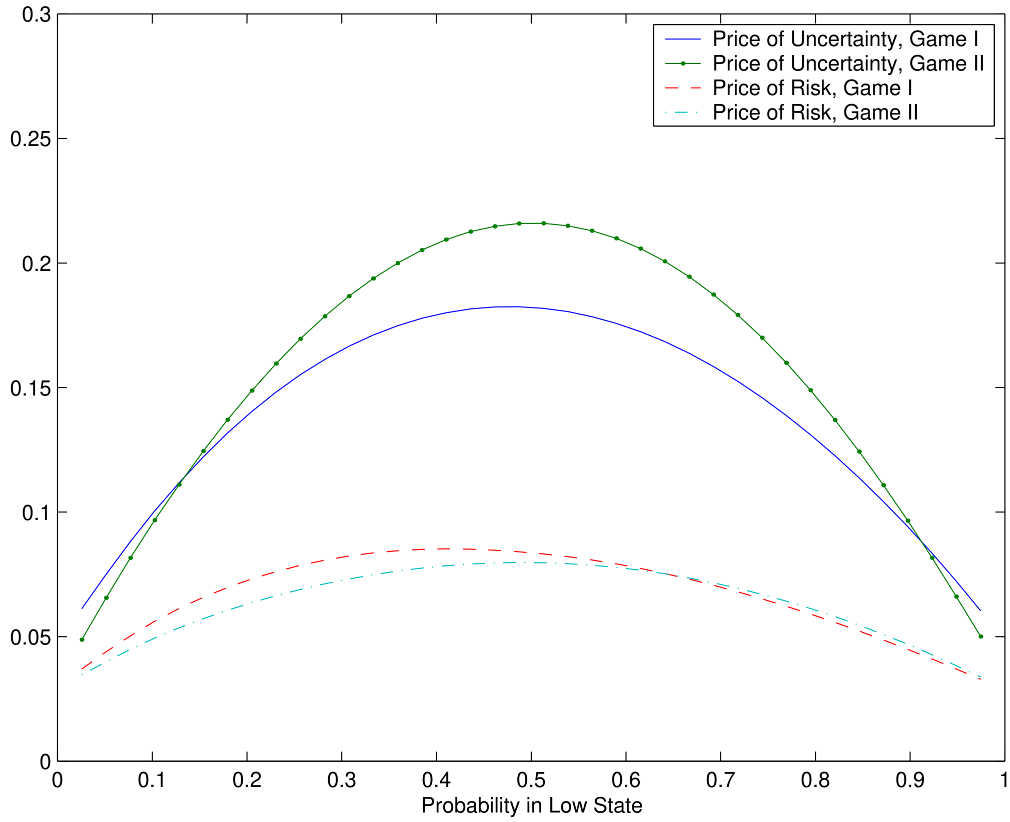

Figure 5: This figure shows the prices of risk and uncertainty as functions of the probability of being in the low growth state. These functions are computed holding k fixed at the sample median.

Figure 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 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 do display cyclical fluctuations, with

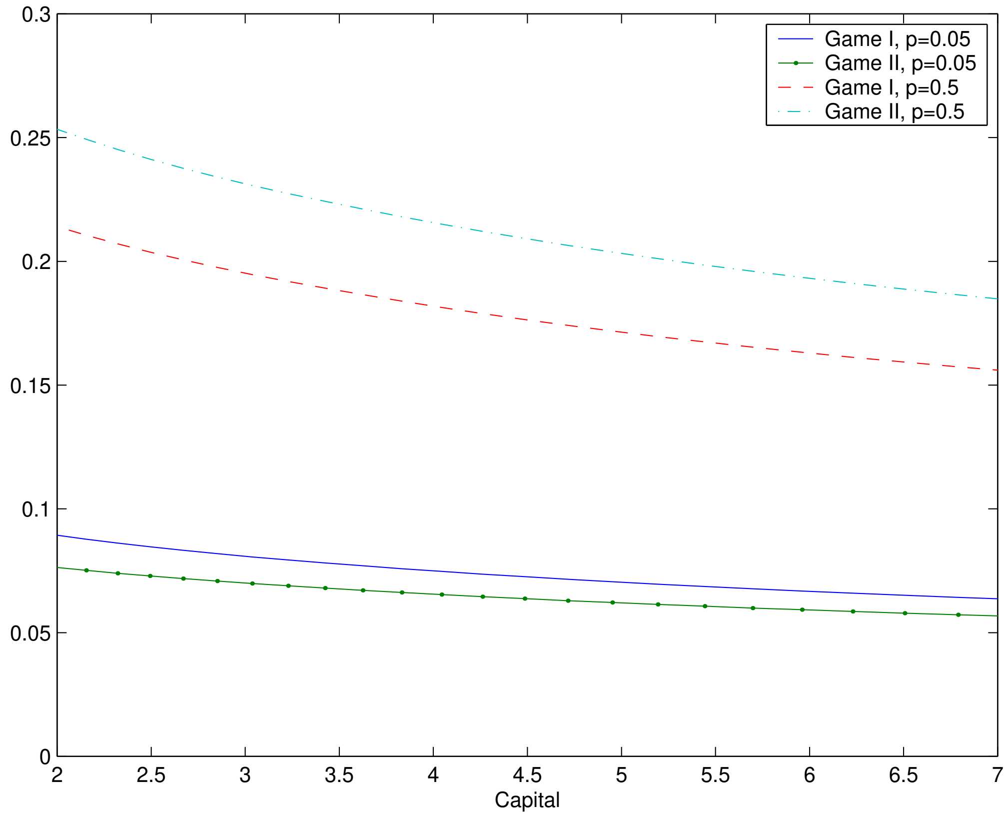

Figure 6: This figure shows the market price of uncertainty as a function of k for Games I and II and for two different values of the probability of being in a low growth state.

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 a 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.

9 Price-Earnings Ratios

To study price-earnings ratios, we apply the HMM model simultaneously to the technology shock process and the earnings process. Growth rates in the technology and earnings respond to the same two-state hidden Markov chain. For the pur-

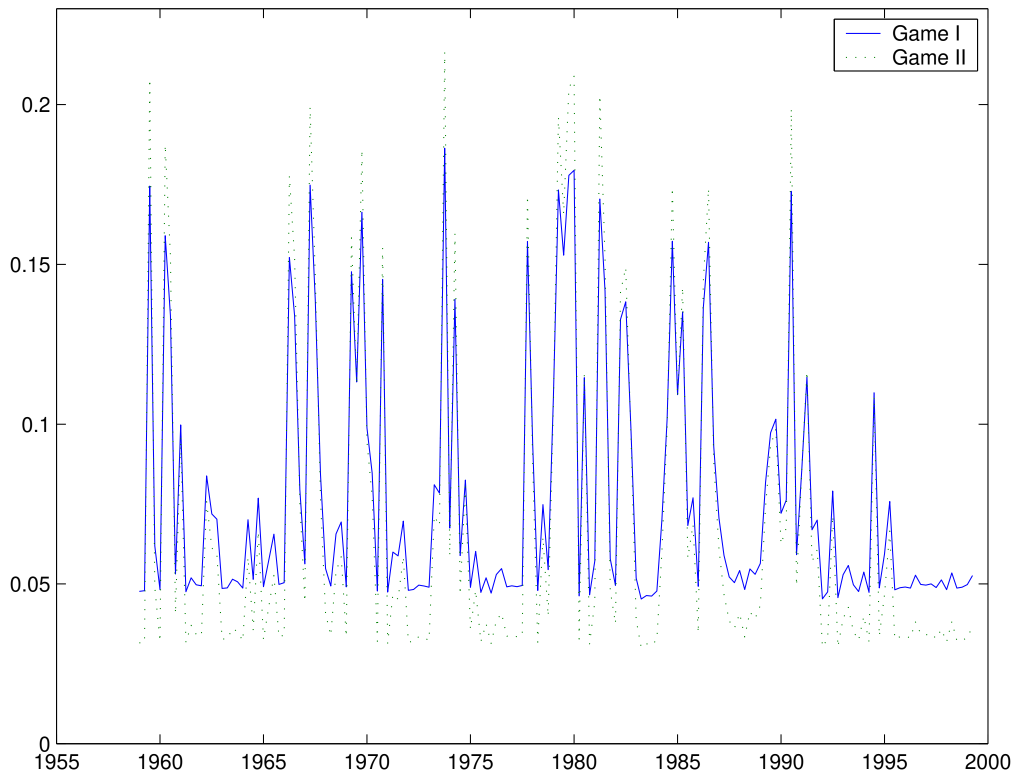

Figure 7: This figure shows the time series for the market price of uncertainty for Games I and II. The market price of uncertainty is measured as $-h^*$ for Game I and as $-g^* \cdot \hat{p}$ for Game II.

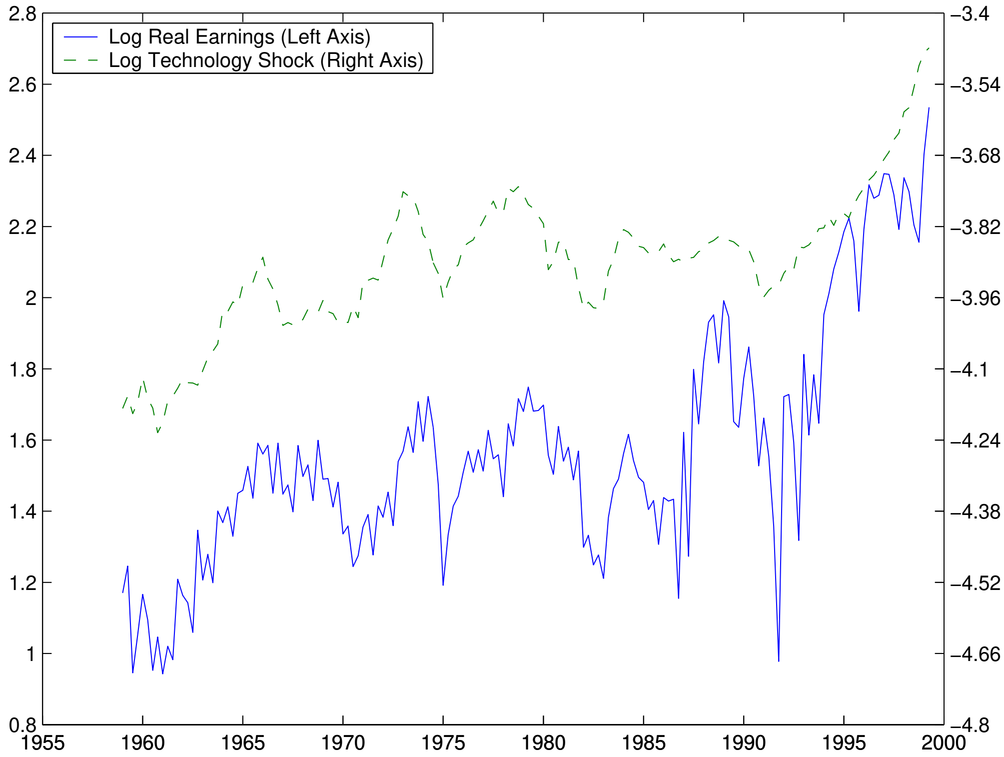

Figure 8: In this figure we report the time series for the cumulative Solow residual and S & P earnings.

poses of pricing, we treat the earnings process as a stream of *dividends*: claims to consumption to be priced. These *dividends* are not extra sources of consumption endowments, but 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 follows closely 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.

The time series for technology and earnings are plotted in Figure 8. We used data on reported earnings on the Standard and Poors Stock Price Index divided by the price deflator for fixed investment. The earnings data were obtained from DRI.

9.1 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 model 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.[14] The estimates of the parameters that govern the technology evolution were very close to those obtained when we estimated the HMM model 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.

9.2 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 $d\hat{B}_t$ associated with the investors' information set. 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:

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

[14] 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.

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 std. dev.	0.0192
σ_e	earnings shock std. dev.	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

Table 4: In this table we report the HMM parameters for the technology-earnings model. 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.

where C^1 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 partial differential 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.[15] 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.

9.3 Results and Figures

We solved for price-earnings functions for the HMM robustness games. For comparisons, 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 model misspecification makes the

[15] 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.

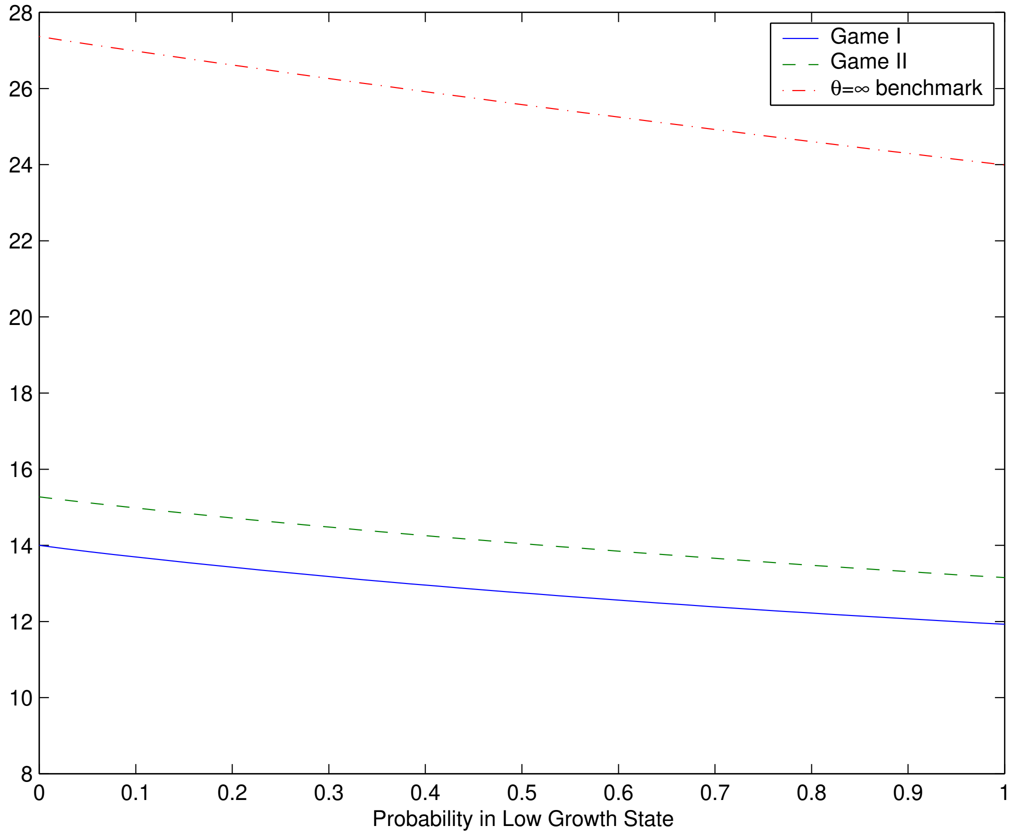

Figure 9: This figure displays the price/earnings ratio as a function of the probability of being in the low state for HMM Games I and II. For comparison, it also displays the price/earnings ratio for the HMM economy that abstracts from robustness ($\theta = \infty$). The functions are plotted by fixing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

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.

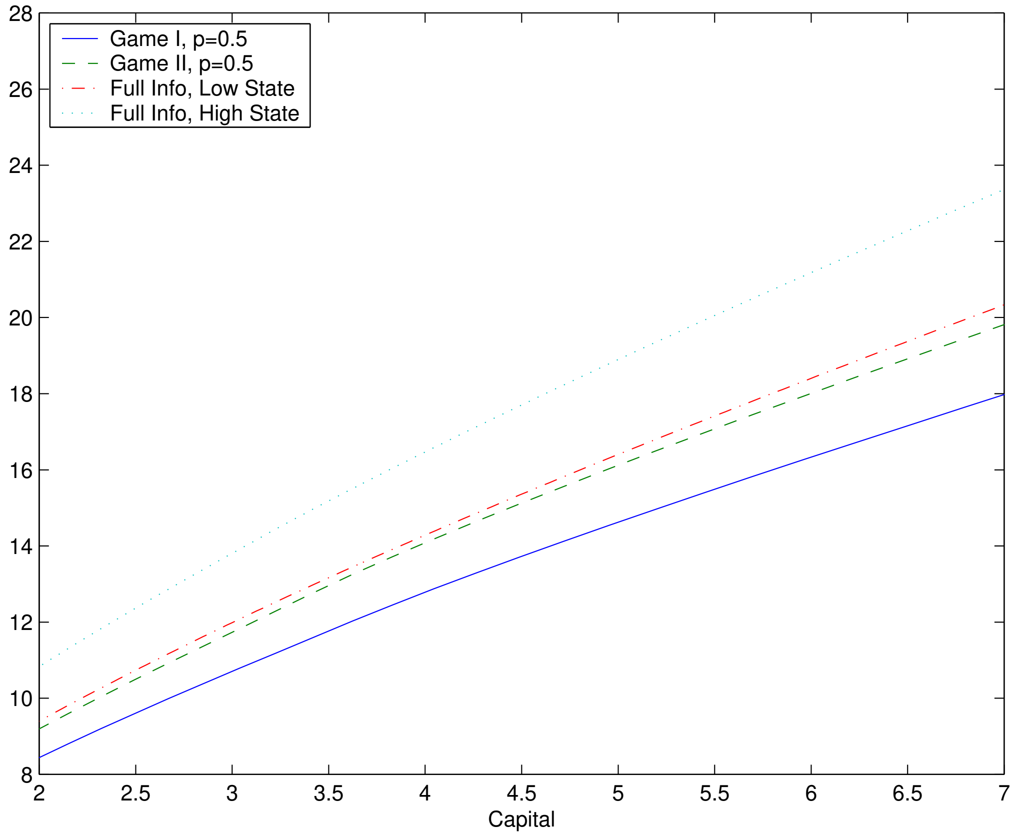

Figure 10: This figure displays price/earnings ratios as functions of the capital/technology ratio k for HMM Games I and II. For comparison, it also reports the corresponding functions for the full information robustness game. The growth state probabilities $\hat{p}$ are fixed at the designated values.

The median price-earnings ratio for our sample is 16.46. As seen in Figures 10 and 9, the robust decision games bring this level more 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. Movements in the price-earnings ratios computed from our models are dominated by changes in the capital-technology ratio k. Recall that the capital/technology ratio declines at the end of our sample starting in 1992. This leads to the counterfactual prediction of 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. Moreover, actual price earnings-ratios are much more variable than those implied by our models.[16] Thus these models, even with robustness added, are not equipped to address the excess volatility puzzle in equity prices. Presumably movements in other information variables than those modeled here are needed better to track the historical time series.

10 Conclusion

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

- The robust motive for precautionary savings leads to an increases in 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 conditional risk-return tradeoff measured by financial econometricians. The robustness component is particularly sensitive to growth-state probabilities and is largest when, under the baseline model, investors are most unsure of the hidden state. There is an intriguing interaction between ambiguity, as reflected by state probabilities, and model uncertainty as encoded in the local factor prices.

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

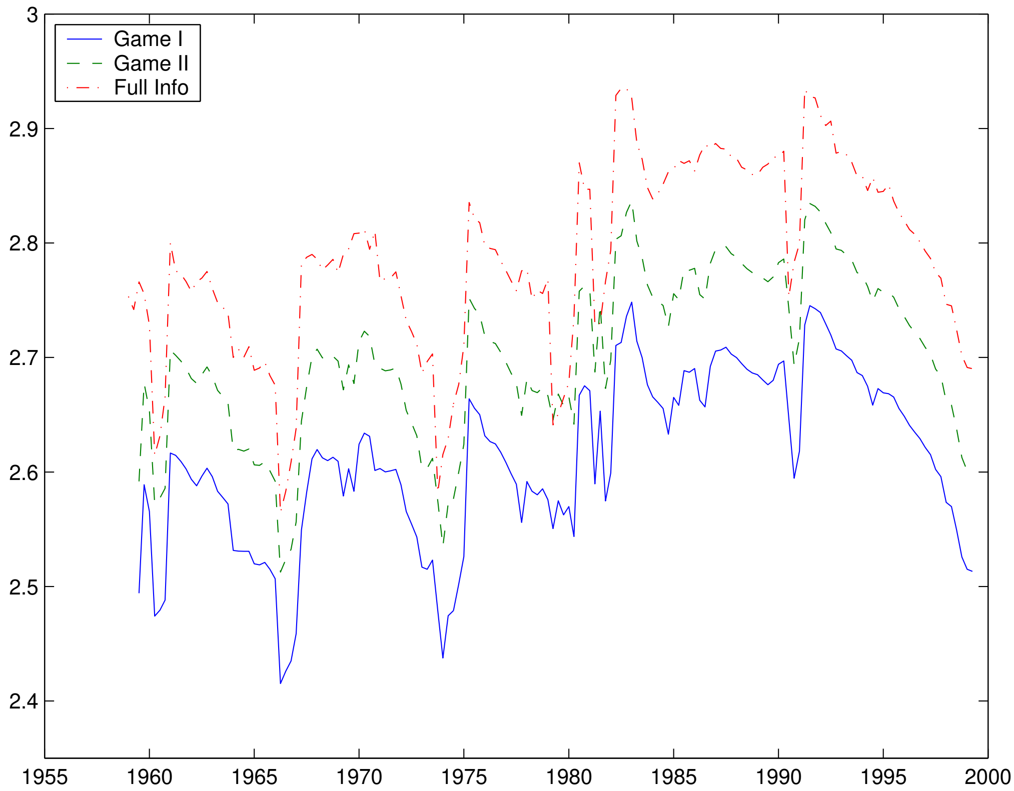

Figure 11: This figure plots the time series for the logarithm of the price/earnings ratio implied by models of robust decision-making.

- A concern about model misspecification causes price-earnings ratios to drop. While the overall level is closer to that in postwar data, the time series trajectories implied by the model show some important departures from the data.

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, a richer learning dynamics may result in quantitatively important aymmetries 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 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. Veldcamp shows how investors' decisions to initiate and terminate projects create more public information about the hidden state in the economy to arrive 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.

References

Aiyagari, S. R. (1994). Unisured idiosyncratic risk and aggregate savings. *Quarterly Journal of Economcics 104*, 219–240.

Anderson, E., L. P. Hansen, and T. Sargent (2000, March). 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 (1999). Habit persistence, asset returns and the business cycle. Mimeo.

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, June). Optimal economic growth and uncertainty: The discounted case. *Journal of Economic Theory 4*(3), 479–513.

Cagetti, M., L. P. Hansen, T. J. Sargent, and N. Williams (2000). Robust recursive prediction and control. unpublished.

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*, pp. 172–194. Oxford: Oxford University Press.

Chalkley, M. and I. H. Lee (1998). Learning and asymmetric business cycles. *Review of Economic Dynamics 1*, 623–645.

Chen, Z. and L. Epstein (1999). Ambiguity, risk and asset returns in continuous time. Mimeo.

Cochrane, J. and L. P. Hansen (1992). Asset pricing explorations for macroeconomics. In O. J. Blanchard and S. Fischer (Eds.), *NBER Macroeconomics Annual*, pp. 115–165. Cambridge, MA: The MIT Press.

Cochrane, J. M. (1991, March). Production-based asset pricing and the link between stock returns and economic fluctuations. *Journal of Finance 46*(1), 209–237.

David, A. (1997, December). Fluctuating confidence in stock markets: Implications for returns and volatility. *Journal of Financial and Quantitative Analysis 32*(4), 457–462.

David, A. and P. Veronesi (1999, November). 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.

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*. New York: Springer-Verlag.

Ethier, S. N. and T. G. Kurz (1986). *Markov Processes. Characterization and Convergence*. John Wiley.

Fellner, W. (1965). *Probability and Profit: A Study of Economic Behavior Along Bayesian Lines*. Irwin Series in Economics. Homewood, Illinois: Richard D. Irwin, Inc.

Hamilton, J. D. (1989, March). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*(2), 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, January). Robust permanent income and pricing. Mimeo.

Hansen, L. P., T. J. Sargent, and N. E. Wang (2000). Robust permanent income and pricing with filtering. *Macroeconomic Dynamics*. forthcoming.

Hansen, L. P. and K. J. Singleton (1983, April). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy 91*(2), 249–265.

Huber, P. J. (1981). *Robust Statistics*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons.

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*(4), 1342–1364.

James, M. R., J. Baras, and R. J. Elliott (1994, April). Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Transactions on Automatic Control 39*(4), 780–791.

Krusell, P. and A. A. Smith (1998). Income and wealth heterogeneity. *Journal of Political Economy 106*, 867–896.

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

Maenhout, P. J. (1999). Robust portfolio rules and asset pricing. Mimeo.

Merton, R. C. (1975, July). An asymptotic theory of growth under uncertainty. *Review of Economic Studies 42*(3), 375–393.

Rouwenhorst, K. G. (1995). Asset pricing implications of equilibrium business cycle models. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Theory*, pp. 294–330. Princeton University Press.

Sclove, S. L. (1983, February). Time-series segmentation: A model and a method. *Information Sciences 29*(1), 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*. Elsevier. Forthcoming.

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, December). Measures of fit for calibrated models. *Journal of Political Economy 101*(6), 1011–1041.

Wonham, W. J. (1964). Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM Journal on Control 2*(3), 347–368.

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.[17] 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 $\hat{C}$. The conditional expectation operator for interval τ can be represented as:

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

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

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

for each of the n values of s.

The generator is composed by 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 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:

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

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

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

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 non degenerate probabilities. The semigroup is again defined using a one-point compactification. Since x is now a diffusion, the resulting generator, denoted $\hat{\mathcal{G}}$, is

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

where ∇f and $\triangle f$ 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 $\hat{\sigma}$, again, depend on the robustness game, and were denoted in the paper 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 however 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 $\hat{\mathcal{G}}f$ to denote the function g used in this construction. $\mathcal{G}$ and $\hat{\mathcal{G}}$ are called the *extended* generators of the corresponding Markov processes (see Ethier and Kurz (1986) and Davis (1993).)

B HJB Equations

In the case of full information and no robustness, using the generator notation previously introduced, the resulting Hamilton-Jacobi-Bellman (HBJ) equation is:

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

For the non robust, 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 penalization. The corresponding equation for Game II is:

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

By substituting the formula for $\widehat{\mathcal{G}}$, we get equation 12 in the paper.

For Game II, we have:

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

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 less 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.

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) and . 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 $\mathrm{HJB}(V, i)$ be the differential equations described above, keeping the decision rule i fixed. Following Candler (1999), we introduce explicitly time, and solve

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

This corresponds to the backwards iteration often used to solve 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:

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

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

  The equation is linear, and we use an implicit, upwind, finite difference method.
- 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 (15),which is then again solved using a finite difference method.