Robust-$\mathcal{H}_\infty$ Forecasting and Asset Pricing Anomalies

Aaron Tornell[*]

UCLA and NBER

First draft: August 1998

This draft: January 2001

JEL Classification No. C61, E44, G12

Abstract

We present an alternative expectation formation mechanism that helps

rationalize well known asset pricing anomalies, such as the predictability of

excess returns, excess volatility, and the equity-premium puzzle. As with rational expectations (RE), the expectation formation mechanism we consider is based on a rigorous optimization algorithm that does not presume *misperceptions* – it simply departs from some of the implicit assumptions that underlie RE. Agents fear the existence of *misspecification* and design strategies that will be robust against a very large class of misspecifications. The new element is that uncertainty cannot be modeled via probability distributions. We consider an asset pricing model where uncertainty is represented by unknown disturbance sequences, as in the $\mathcal{H}_\infty$-control literature. Agents must filter the 'persistent' and 'transitory' components of a sequence of observations in order to make consumption and portfolio decisions. We find that $\mathcal{H}_\infty$ forecasts are more sensitive to news than RE forecasts and equilibrium prices exhibit the anomalies previously mentioned.

1. Introduction

There are several stylized facts in macroeconomics that cannot easily be explained by rational expectations models with standard time-additive preferences. Prominent examples are the excess volatility of asset prices, the predictability of excess returns, and the equity premium and risk-free rate puzzles.

In order to rationalize these anomalies researchers have extended the benchmark model in different directions. One strategy has been to keep rational expectations, but to consider more complex preferences that generate time-varying risk premia, or to introduce imperfections in credit markets.[1] Another strategy has been to consider alternative expectation formation mechanisms. A growing behavioral literature postulates small *misperceptions* on the part of otherwise optimizing agents. Specifically, if agents use Bayes law to form expectations but *misperceive* shocks to be more persistent than what they actually are, then forecasts will be more sensitive to news than rational expectations (RE) forecasts. As a result, asset prices will exhibit the previously mentioned macroeconomic anomalies.[2]

---

[1]Barberis, et. al. (1999), and Campbell and Cochrane (1999) consider more complex preferences. Constantinides and Duffie (1996), and Heaton and Lucas (1996) consider incomplete markets.

[2]See Barberis, et. al. (1998), Cecchetti, et. al. (1997), Daniel, et al. (1998), Gourinchas and Tornell (2000), and Mullainathan (1999). An attractive feature of this literature is that the misperceptions which are considered reflect stylized facts established by psychological experiments (e.g. Kahneman and Tversky (1972)).

This paper considers an alternative expectation formation mechanism that departs from the assumptions implicit in RE models. The updating formulas derived under this alternative mechanism generate the same excess sensitivity to news as the behavioral forecasting formulas. However, they are derived by a rigorous optimizing algorithm, without postulating the existence of misperception or irrationality. The main departure from RE is the elimination of the assumption that all uncertainty can be parametrized by a stochastic process whose properties are known, or can be learned.

In the robust-$\mathcal{H}_\infty$ approach considered in this paper, optimizing agents take seriously into account that it is impossible to eliminate all *misspecifications* from their description of the economy, and so design *'robust'* filters and policies. We will see that, in an otherwise standard asset pricing model with time-additive preferences, robust agents behave *as if* they misperceive shocks to be more persistent than what they actually are. As a result, robust forecasts are more sensitive to news than the RE forecasts associated with the same 'nominal' model of the economy. This in turn implies that equilibrium prices exhibit the macroeconomic anomalies mentioned above.

We consider the familiar setup in which an agent must filter the persistent and transitory components of a sequence of observations in order to estimate the

unobservable state of the economy and make forecasts and portfolio decisions.[3] The point of departure of the robust approach is the recognition by agents that although they may know the model that generates payoffs with a high degree of accuracy, they are not perfectly sure about this. There might be misspecifications in the process followed by disturbances, or in the formulation of the model.[4] Furthermore, these misspecifications might follow complicated dynamic patterns that cannot be parametrized with a probability distribution. The robust agent designs forecasting and decision rules that will work well in the face of all misspecifications that satisfy a certain norm bound. $\mathcal{H}_\infty$-control operationalizes this idea by modeling uncertainty as totally unknown disturbance sequences with a bounded l_2-norm (i.e., the disturbances are square summable). An attractive feature of $\mathcal{H}_\infty$-control is that it leads to closed-form solutions.

Several approaches to robust control were developed during the 1980s. Its use in economics has been pioneered by Tom Sargent and Lars Hansen. In a series of papers they have shown that significant insight can be gained by relaxing some assumptions made in RE models.[5] This paper is most closely related to Hansen, et. al. (1999), who rationalize macroeconomic anomalies by considering

---

[3] This is the same setup as the one considered by Lucas (1973) and Muth (1960).

[4] For instance, there are unmodelled nonlinearities, the distribution of disturbances changes continuously in unpredictable ways, etc.

[5] See Anderson, et. al. (1999), Hansen and Sargent (1998), and Hansen, et.al. (1999).

'risk-sensitive' preferences. There are two main differences. First, they model uncertainty in terms of normally distributed shocks, while a major feature of our setup is that uncertainty cannot be parametrized. Second, in their model the state is perfectly observed, so there is no signal-extraction problem. In the literature review below, we expand on these issues and describe other work on robust control and filtering.

When applying $\mathcal{H}_\infty$-control to economic problems one must take into consideration that, in equilibrium, prices and agents' decisions are interdependent. Thus the existing $\mathcal{H}_\infty$ formulae are not applicable. A contribution of this paper is to construct a competitive equilibrium of an exchange economy, and provide closed-form filtering and forecasting formulae, as well as asset pricing equations. Although the formulas we derive are specific to the asset pricing problem we consider, the framework and the method to construct equilibria can be applied to other economic problems in which agents must solve a signal extraction problem in order to make forecasts and decisions.

The results in this paper are also of interest from a estimation perspective. If one forecasts a noisy observation (say, dividends), and the performance index is a function of the forecast's mean square error, then the $\mathcal{H}_\infty$-forecasts do not exhibit excess sensitivity to news relative to RE. In this paper we consider a different

4

problem. Dividend forecasts and portfolio policies are *jointly* determined in a problem whose performance index is induced by a quadratic discounted utility and the agent's budget constraints. As we shall see, the $\mathcal{H}_\infty$-forecasts associated with this problem do exhibit excess sensitivity to news. In other words, in our simple exchange economy being robust to misspecification entails putting more weight on current innovations relative to RE. That is, if two agents have the same 'nominal' model of the economy, the robust agent will put less weight on his prior, when updating beliefs, than the agent that uses RE. It is interesting to note that in the behavioral literature excess sensitivity to news arises from misperception about the importance of transitory and persistent shocks, while in the $\mathcal{H}_\infty$ approach it arises from a robustness concern against misspecification.

We perform some simulations to illustrate how, if the degree of robustness is high, $\mathcal{H}_\infty$ prices are more volatile than dividends, excess returns are predictable by past dividend yields, and the equity premium is high while maintaining a low risk-free interest rate. We also feed the $\mathcal{H}_\infty$ and RE price formulas with actual US dividend data for the period 1871-1996 and find that $\mathcal{H}_\infty$ prices track US. stock prices better than RE prices. As we mentioned earlier, $\mathcal{H}_\infty$ prices can generate these anomalies because in our model $\mathcal{H}_\infty$ forecasts are more sensitive to news than RE forecasts.

The trade-off between $\mathcal{H}_\infty$ and rational expectations strategies is the following. $\mathcal{H}_\infty$ strategies are designed to perform well under any norm-bounded misspecification, while RE strategies are designed to attain the best performance conditional on no misspecification. That is, in the special case in which there is no misspecification, and the data is generated by a known stochastic process, the RE approach generates the best performance. However, in the presence of misspecification, the $\mathcal{H}_\infty$ approach may provide better results. Figure 1 makes this point clear.

A departure from the solution method commonly used in economics is that forecasts of exogenous variables are not formed independently of agents' policies by applying Bayes law to a prior distribution. Instead, forecasts and portfolio policies are jointly determined in a problem that can be analyzed as a dynamic game between the agent and nature. Interestingly, although the problems solved by agents in a RE and an $\mathcal{H}_\infty$ setups are different, the forecasting formulas are quite similar. Furthermore, in the limiting case in which agents do not care about robustness, forecasting formulas are identical. The $\mathcal{H}_\infty$ approach is not so wild after all!

An important property of the approach we use is that the degree of robustness that agents can attain is closely related to the existence of equilibria. Robustness cannot be increased arbitrarily. There is a point beyond which an increase in the

6

degree of robustness implies that a solution to the dynamic game between the agent and nature ceases to have a solution. Thus, beyond the breakdown point equilibria do not exist.

We decompose the original $\mathcal{H}_\infty$-control problem into several sub-problems that are solved using methods familiar to economists, and we explain in detail every step involved in deriving the equilibrium. This means that the paper can be followed without any prior exposure to the robust control literature. Furthermore, to make clear what are the differences and similarities between RE and $\mathcal{H}_\infty$ formulae, we also present a rational expectations version of the problem.

The structure of the paper is as follows. In the next two subsections we present a brief review of the literature, and an overview of the model. In Section 2 we formulate the problems solved by agents in RE and in $\mathcal{H}_\infty$ setups, and provide the formulas for equilibrium asset prices and dividend's forecasts. In section 3 we derive the equilibrium in an $\mathcal{H}_\infty$ economy. In Section 4 we present simulation exercises. In section 5 we present the conclusions. Finally, in the Appendix we present some proofs and the methodology used in the simulations.

1.1. Related Literature

The development of the rational expectations approach is related to the development of optimal control. By the late 1970s there was widespread dissatisfaction among practitioners with the robustness properties of optimal control. Although minimizing an expected cost function leads to the best performance in the case of no misspecification, it leads to bad results and even instabilities in the face of small misspecifications. And when you plan to launch a missile, you would rather be sure this does not occur.[6]

Robust control was developed in the 1980s in response to these shortcomings. This approach directly addressed the issue of how accurate the model and the description of uncertainty should be, and how the performance index should be defined to guarantee the robustness of the controlled system against misspecifications. Various approaches to robust control have developed. In this paper we consider the $\mathcal{H}_\infty$ specification pioneered by Zames (1981). Important developments can be found in Basar and Bernhard (1991), and Zhou, Doyle, and Glover

---

[6]This point is nicely made by Peter Huber in an influential book on classical robust statistcs. Referring to the assumptions one makes when using probability distributions he says (Huber (1981), pg. 1): "The assumptions are not supposed to be exactly true –they are mathematically convenient rationalizations of an often fuzzy knowledge... one justifies their use by appealing to a vague continuity or stability principle: a minor error in the mathematical model should cause only a small error in the final conclusions. Unfortunately, this does not always hold. During the past decade one has become increasingly aware that some of the most common statistical procedures are excessively sensitive to seemingly minor deviations from the assumptions..."

(1996). Hansen and Sargent (1998) show how to treat robust control problems when there is discounting.

$\mathcal{H}_\infty$-control was developed to address engineering problems, and the existing control and forecasting formulas cannot be applied directly to economic problems. First, the performance indexes are different. In economics they are induced by the underlying utility function and budget constraints. Second, in economics one is typically interested in characterizing equilibria where prices clear markets. A contribution of this paper is to construct a competitive equilibrium for a simple exchange economy, and to provide closed-form forecasting and asset pricing formulae.

In the $\mathcal{H}_\infty$-filtering literature the performance index includes the mean square error of the 'unobservable state.' In contrast, in economics, payoffs are generally functions of the 'noisy observation' (e.g., dividends, consumption).[7] It turns out that the $\mathcal{H}_\infty$ filter of the state exhibits excess sensitivity to news relative to RE. However, this is not true for the $\mathcal{H}_\infty$ dividend's forecasts associated with a standard performance index that includes the forecast's MSE. As we mentioned earlier, an attractive property of our model is that the associated dividend's fore-

---

[7]Here we are referring to the familiar setup in which the noisy observation equals the unobservable state plus some noise.

9

casts exhibit excess sensitivity to news.

Independent of $\mathcal{H}_\infty$-control, risk sensitive control was developed in the 1970s (see Whitle (1990)). This approach considers a standard Gaussian setup, with no misspecification. The twist is that the utility index is $\frac{1}{\sigma} \log E_t \left(\sigma u_{t+1} \right)$ instead of $E_t(u_{t+1})$. It has been found that in some cases risk sensitive control yields the same formulas as the ones obtained with $\mathcal{H}_\infty$-control (Doyle, Glover, Khargonekar and Francis (1989)). Hansen, et. al. (1999) use this fact in a very interesting way. They consider a risk-sensitive version of the permanent income model with habit persistence and quadratic utility. In their model uncertainty takes the form of endowment and preference shocks. These shocks are i.i.d. normally distributed, with known mean and variance. Each period, agents observe the 'state' (past consumption, capital stock and the shocks), and choose consumption and investment. They then consider the asset pricing implications, and show that a sufficiently high preference for robustness generates a market price of risk that is compatible with US. data. Anderson, et. al. (1999) consider more general utility functions and a larger class of perturbations. Manehout (1999) applies the model of Anderson, et. al. to a portfolio choice problem.

One difference between our work and these papers resides in the existence of a filtering problem. In our model agents need to filter the transitory and persis-

tent components from sequences of past observations in order to make forecasts. Furthermore, these forecasts must be made jointly with the portfolio decision. In the papers just mentioned, filtering plays no role as agents observe perfectly the state of the economy. Another difference is that we consider an $\mathcal{H}_\infty$ approach, while they consider a risk-sensitive approach. Onatski and Stock (2000) also use the robust approach to tackle an economics problem.

In a robust setup, even if agents were to update their model of the economy they would still fear misspecification. Thus, their estimators will not converge to the RE estimators associated with the same nominal model of the economy. In this sense the present model is different from learning models, such as Marcet and Sargent (1989) and Timmerman (1996).

Finally, we would like to make two comments about the use of the word 'robust' in the literature. First, in the robust control literature, robust refers to guaranteed performance for any disturbance sequence that satisfies a certain norm bound. In the statistical literature it generally signifies insensitivity to small deviations from the assumptions.[8] Second, we would like to note that robust control procedures have little overlap with non-parametric, distribution-free, and adaptive

---

[8]In classical statistics researchers are mainly concerned with distributional robustness (see Huber (1981)). In Bayesian statistics the main concern is with possible misspecification of the prior or of the likelihood function, and with the so called Bayes risk robustness (see Berger (1985)).

procedures.

1.2. Overview

We consider an exchange economy with one consumption good and two one-period assets: a safe asset that pays one unit of the consumption good, and a risky asset whose future dividend is uncertain. Dividends (y_t) have a persistent component (x_t) and a transitory component $(v_t) : y_t = x_t + \sigma_v v_t$. The key is that the persistent component is unobservable and it is autocorrelated: $x_t = ax_{t-1} + \sigma_w w_{t-1}$.

We compute the equilibria in an RE setup as well as in an $\mathcal{H}_\infty$ setup in order to make comparisons. In both setups agents filter the transitory and persistent components from dividend observations $\{y_s\}_{s=1}^t$ (i.e., estimate x_t), and then make forecasts of future payoffs. In the RE setup agents use Bayes law to generate the distribution of future dividends. The important point to note is that in the RE setup we derive forecasts of future dividends without taking into consideration the consumption and portfolio strategies. This 'separation' does not exist in an $\mathcal{H}_\infty$ setup because one cannot speak of a probability distribution of the disturbances that captures all that is unknown about the environment. Disturbances are simply unknown square-summable sequences. This implies that forecasts and portfolio policies must be determined jointly. It turns out that this problem can be analyzed

12

as a dynamic game between the representative agent and nature. A significant part of the paper is devoted to solve this game.

Theorem 3.1 characterizes the solution of this game and establishes the conditions for existence of an equilibrium. The equilibrium is constructed using the '$\mathcal{H}_\infty$ certainty equivalence principle' (Basar and Bernhard (1991)). This principle breaks the original problem into three simple sub-problems. The first is a familiar *backward* dynamic programing problem that characterizes consumption and portfolio strategies for a given value of the state $x_{t+1} = x$. The second sub-problem is a *forward DP* problem that allows the agent to extract the persistent component from past dividend's observations, conditional on $x_{t+1} = x$. The third sub-problem derives the $\mathcal{H}_\infty$ estimates of x_{t+1} and y_{t+1}, as well as equilibrium asset prices. This is done by using the value functions associated with the forward and backward DP problems ((3.14) and (3.19), respectively).

Proposition 2.2 uses Theorem 3.1 to characterize a Markov perfect equilibrium, and exhibits the $\mathcal{H}_\infty$ forecasting formulas and equilibrium asset prices that are used in the simulations (formulas (2.10)-(2.11)). Surprisingly, these formulas turn out to be very similar to the familiar RE formulas (2.8)-(2.9). Furthermore, in the limiting case in which agents do not care about robustness, the $\mathcal{H}_\infty$ and RE forecasts are equal.

A key parameter throughout the paper is γ, which measures the degree of disturbance attenuation (i.e., the lower γ, the greater the degree of robustness). It will turn out that the dynamic game between the agent and nature has a solution if and only if γ is greater than a certain threshold $\underline{\gamma}$. This implies that the degree of robustness that can be attained in equilibrium cannot be arbitrarily large.

It is interesting to note that in the exchange economy we are considering, the lower γ, the greater the sensitivity to news. This property is what allows our model to generate excess volatility, predictability of excess returns, and the equity premium puzzle for a high degree of robustness. In future research we plan to consider a setup in which γ varies through time in order to generate other anomalies that have been identified in the finance literature.

2. Model

In order to concentrate on the essential features of the $\mathcal{H}_\infty$ approach, and to highlight its differences with standard rational expectations (RE) models, we will consider a streamlined Lucas (1978) exchange economy. As we shall see, in both setups agents optimize an induced objective using all available information. The main difference resides in the way uncertainty is modeled.

The model is an overlapping generations economy with one perishable con-

sumption good and two one-period assets: a safe and a risky. A time-t safe asset pays one unit of the consumption good at $t+1$, while a time-t risky asset pays y_t units of the consumption good at t, and y_{t+1} at $t+1$. The economy is populated by a sequence of representative agents that live two periods. A $t-$young agent is endowed with one unit of the risky asset. At time t, taking prices as given, he chooses consumption c_t, and the amount of risky and safe assets (q_t^r, q_t^s) he wishes to hold. It follows that his budget constraint is

$$c_t + p_t^r q_t^r + p_t^s q_t^s \leq p_t^r + y_t \tag{2.1}$$

At $t+1$, when old, the $t-$agent consumes all his wealth

$$c_{t+1}^o = y_{t+1} q_t^r + q_t^s \tag{2.2}$$

We will use a state-space representation for dividends:

$$y_j = x_j + d_j + \sigma_{vj} v_j, \qquad j \geq 1$$
$$x_{j+1} = a_j x_j + \sigma_{wj} w_j, \qquad x_0 = 0 \tag{2.3}$$

Deviations of dividends from trend $(y_j - d_j)$ have a transitory component $\sigma_{vj} v_j$,

as well as a persistent component x_j. The persistent component, which we will call 'the state,' follows an autoregressive process.[9]

At time t a young agent observes the history of dividends $\{y_j\}_{j=1}^t$, but does not observe the transitory and persistent components of dividends' deviations from trend. We will assume that the parameters $\{a_j\}_{j=1}^t$, $\{\sigma_{wj}\}_{j=0}^t$, $\{\sigma_{vj}\}_{j=1}^{t+1}$ and $\{d_j\}_{j=1}^{t+1}$ are known, while the disturbances $\{v_j\}_{j=1}^{t+1}$, $\{w_j\}_{j=0}^t$ are unknown. We will also assume that the initial value of the unobservable state x_1 is unknown. We represent this uncertainty by setting $x_0 = 0$, so that x_1 is proportional to the first element of the disturbance sequence: $x_1 = \sigma_{w0} w_0$.

A key difference between the RE and the $\mathcal{H}_\infty$ approaches is the way in which uncertainty is modeled. Under RE it is assumed that uncertainty can be summarized by a stochastic process, and that an agent has a unique prior probability distribution over all possible distributions that might govern the stochastic process. In contrast, the $\mathcal{H}_\infty$ approach models uncertainty as totally unknown disturbance sequences with a bounded l_2 norm. Below we will be more specific.

Agents have time-additive utility: $u(c_t) + \beta u(c_{t+1})$, where $u(c_t) = -[m - c_t]^2$ and m is a constant. An instantaneous quadratic utility function will allow us

---

[9]In the special case that the state is a constant ($x_j = x$) and $d_j = 0$, (2.3) collaspses to the familiar setup in which agents try to estimate an unobservable from a noisy observation: $y_j = x + \sigma_{vj} v_j$.

to obtain closed-form solutions for the $\mathcal{H}_\infty$ problem that are comparable to the solution of the RE problem. As will become clear latter, this utility function induces an objective for the $\mathcal{H}_\infty$ problem which is an indefinite-quadratic form in the disturbances. Ensuring the positivity of such an indefinite form is the key to the existence of equilibria.

In what follows we formulate the problems solved in a Rational Expectations economy, and in an $\mathcal{H}_\infty$ economy. Then in the following subsections we characterize the equilibria.

In a RE economy a young agent at time t solves the following problem.

Problem RE. *Suppose that the disturbances* $\{v_j\}_{j=1}^{t+1}$, $\{w_j\}_{j=0}^{t}$ *are i.i.d.* $N(0,1)$ *random variables. Given prices* (p_t^r, p_t^s), *choose* (c_t, q_t^r, q_t^s) *to maximize*

$$E\left(-[m - c_t]^2 - \beta[m - c_{t+1}^o]^2 \mid I_t\right) \tag{2.4}$$

subject to budget constraints (2.1), (2.2), to the process followed by dividends (2.3), and to the information set $I_t = \{y_1, ..., y_t, p_t^r, p_t^s\}$.

In the robust approach the representative agent takes seriously into account the fact that his dividend model might be misspecified, and he has no way to eliminate this misspecification. For instance, either the value of the autoregressive coefficient

17

in (2.3) is slightly off, the variances of the disturbances may change constantly in unpredictable ways, there might be some unmodelled non-linearities, etc. In the standard $\mathcal{H}_\infty$ setup misspecification patterns are modeled as unknown disturbance sequences that are square summable (i.e., $\{v_j, w_j\}_{j=0}^{t+1} \in l_{2\,[0,t+1]}$). In other words, 'not anything goes.' We will denote by Ω the set of all such sequences whose elements are not all zero[10]

$$\Omega \equiv \left\{ \{v_j, w_j\}_{j=0}^{t+1} \neq 0 \;\middle|\; \sum_{j=0}^{t+1} \left[v_j^2 + w_j^2\right] < \infty \right\} \tag{2.5}$$

This description of uncertainty is quite different from the probabilistic one used in standard rational expectations models. We will take a brief detour to comment on it. In RE models, and in optimal control problems in general, one describes uncertainty in terms of probability distributions, and optimizes the expected value of certain objective function. During the 1970s researchers noted that, when the state must be estimated, the resulting optimal controllers do not posses good robustness properties, and might induce instabilities in the controlled system. In

---

[10]Note that since the first dividend observation is y_1, it follows that $v_0 = 0$. Note also that the norm considered in (2.5) corresponds to the specification in (2.3), in which we have set $x_0 = 0$ and w_0 belongs to the disturbance sequence. Alternatively, we could specificy x_1 as unknown and set $w_0 = 0$. In this case the l_2-norm of the disturbances would be: $\| \omega \|_{2,[1,t+1]}^2 = \sigma_x^2 x_1^2 + \sum_{j=1}^{t+1} \left[v_j^2 + w_j^2\right].$

the 1980s $\mathcal{H}_\infty$-control was developed in order to be able to tackle problems in which one cannot model via a probability distribution all that is unknown about the system in question. The issue then became what would be an appropriate objective and description of uncertainty that would generate a tractable problem, and would allow the designer to attain robustness? In a seminal paper George Zames (1981) proposed to describe uncertainty in terms of disturbances that satisfy a condition such as (2.5), and to consider as the objective the $\mathcal{H}_\infty$ norm of the dynamic system that maps disturbances to output. Since then, $\mathcal{H}_\infty$-control has become a very active area of research and has found several applications.

Let us return to our exchange economy. In order to keep the presentation simple we will not make reference to the term $\mathcal{H}_\infty$ norm.[11] The problem of an agent in an $\mathcal{H}_\infty$ economy is to device a consumption and portfolio strategy that will perform well in the face of all possible disturbance sequences $\{v_j, w_j\}_{j=0}^{t+1}$ with bounded l_2 norm. Now the issue becomes, how should the induced objective be defined to guarantee robustness? Since in an $\mathcal{H}_\infty$ setup disturbances are not restricted to have any specific probability distribution, we cannot postulate that the agent maximizes an *expected utility* as in the *RE* Problem. Instead, we will consider a 'normalized' utility: the ratio of discounted utility to the l_2-norm of

---

[11]In Section 3 we will define this term, and make clear its link with Problem $\mathcal{H}$ stated below.

disturbances. That is, the induced objective is $-\left\{ \frac{[m-c_t]^2+\beta[m-c^o_{t+1}]^2}{\sum_{j=0}^{t+1}\left[w_j^2+v_j^2\right]} \right\}$. Clearly, the smaller the ratio in braces, the better the performance of the consumption and portfolio strategies.

As we shall see, although the induced objectives and the descriptions of uncertainty under an RE and an $\mathcal{H}_\infty$ economy are different, the forecasting formulas that agents use in equilibrium will turn out to be very similar. The problem solved at time t by an agent in an $\mathcal{H}_\infty$ economy is the following.

Problem $\mathcal{H}$. *Given asset prices* (p_t^r, p_t^s) *and dividends' history* $\{y_j\}_{j=1}^t$

1. *Choose* (c_t, q_t^r, q_t^s) *to ensure*

$$\frac{[m-c_t]^2 + \beta[m-c^o_{t+1}]^2}{\sum_{j=0}^{t+1}\left[w_j^2+v_j^2\right]} \leq \gamma_{t+1}^2 \tag{2.6}$$

for all nonzero disturbances $\{v_j, w_j\}_{j=0}^{t+1} \in l_{2[0,t+1]}$ *that are consistent with dividend history* $\{y_j - d_j\}_{j=1}^t \neq 0$. *The triplet* (c_t, q_t^r, q_t^s) *must satisfy budget constraints (2.1) and (2.2).*

2. *Find the lowest value of the disturbance attenuation parameter* $\underline{\gamma}_{t+1} > 0$, *such that (2.6) holds for all* $\gamma_{t+1} > \underline{\gamma}_{t+1}$.

In an $\mathcal{H}_\infty$ setup the induced objective of an agent is to design consumption

20

and portfolio strategies that will attenuate the effects of all disturbance sequences, with bounded l_2 norm, on the l_2 gain (2.6). Loosely speaking, when there is little noise (i.e., the denominator is small), one would like to attain high discounted utility (i.e., the numerator should be small). When there is a lot noise one cannot expect to attain high discounted utility. However, one would like to bound the rate at which utility deteriorates as the l_2 norm of the disturbances increases. This rate is measured by the so called 'disturbance attenuation parameter' γ_{t+1}, which will play a key role in what follows. Note that it would not make sense to simply consider the numerator in (2.6) as the objective because it is dependent, through the budget constraint, on the l_2 norm of the disturbances.

Note that γ_{t+1} is an index of robustness. The greater γ_{t+1}, the smaller the degree of robustness. We will show that in the limit, when agents do not care at all about robustness (i.e., $\gamma_{t+1} \to \infty$), the $\mathcal{H}_\infty$ forecasting formulas coincide with the RE formulas. We will also show that in equilibrium γ_{t+1} cannot be arbitrarily small. In fact, part 2 of Problem $\mathcal{H}$ consists in determining the maximum disturbance attenuation rate that can be attained.[12]

---

[12]In a classical statistical setup one could introduce distributional robustness by assuming that certain parameters of the probability distribution are unknown, and can be drawn from a large class (e.g. Huber (1981)). In a Bayesian setup one ensures posterior robustness through sensitivity analysis or with more formal methods, such as the use of 'inherently robust priors,' or by considering different priors that belong to either a class of a given functional form, or belong to the so called $\varepsilon-$ contamination class of priors (e.g. Berger (1985)). Note that these

In order to close the model we assume that the supply of the risky asset is one, and that of the safe asset is zero. An equilibrium is defined as follows.

Definition 2.1. *An equilibrium of an RE (respectively, $\mathcal{H}_\infty$) economy is a collection $\{\widehat{c}_t, \hat{q}_t^r, \hat{q}_t^s, \hat{p}_t^r, \hat{p}_t^s\}$ such that: (i)given dividend history and prices, the representative agent solves Problem RE (respectively, $\mathcal{H}$); and (ii)assets and goods markets clear: $\widehat{c}_t = y_t$, $\hat{q}_t^r = 1$, $\hat{q}_t^s = 0$.*

In the remainder of this section we will characterize equilibria in an RE and an $\mathcal{H}_\infty$ economy. Then we will compare the properties of the forecasting formulas.

2.1. Equilibrium in an RE Economy

The solution to Problem RE is well known (e.g. Lucas (1978)). Here we will simply emphasize some aspects of the solution where the key differences between rational expectations and $\mathcal{H}_\infty$ lie. This subsection can be skipped without loss of continuity.

In order to solve Problem RE a young agent computes the distribution of next period's dividends using Bayes law. Since at time t the sequence of past dividends $\{y_j\}_{j=1}^t$ belongs to the information set, a young agent forecasts that

---

procedures do not guarantee robust stability to 'any norm-bounded misspecification' because the designer is still making strong assumptions about the nature of uncertainty.

y_{t+1} is Normally distributed with mean and variance given by

$$E(y_{t+1}|I_t) = \hat{x}_{t+1} + d_{t+1} \qquad\qquad var(y_{t+1}|I_t) = Z_{t+1} + \sigma^2_{vt+1} \qquad (2.7)$$

where $\hat{x}_{t+1}$ and Z_{t+1} are the mean and variance of the unobservable state x_{t+1}, conditional on information $\{y_j\}_{j=1}^t$. They are given by the tth-elements of the familiar recursion

$$
\begin{aligned}
\hat{x}_{j+1} &= a_j \hat{x}_j + h_j [y_j - d_j - \hat{x}_j] \\
Z_{j+1} &= \frac{a_j^2}{Z_j^{-1} + \sigma_{vj}^{-2}} + \sigma^2_{wj}, \qquad & h_j &= \frac{a_j Z_j}{Z_j + \sigma^2_{vj}} \\
Z_1 &= \sigma^2_{w0}, \qquad & \hat{x}_1 &= 0. \qquad (2.8)
\end{aligned}
$$

The intuition is as follows: since transitory disturbances have zero mean, predicted dividends equal the trend component d_{t+1} plus the expected value of the unobservable state, conditional on $\{y_j\}_{j=1}^t$: $\hat{x}_{t+1}$. Formula (2.8) shows that $\hat{x}_{t+1}$ is obtained recursively by adding to $\hat{x}_j$ the innovation $[y_j - d_j - \hat{x}_j]$ multiplied by the 'gain' h_j. Note that h_j is decreasing in the variance of the transitory disturbance. In the limit, as $\sigma^2_{vj} \to \infty$, h_j converges to zero. Since all disturbances are transitory, dividends' observations convey no information about the persistent

23

component x_j. Thus, they are useless in predicting future dividends.[13]

The recursion that generates $\hat{x}_{t+1}$ is derived using Bayesian updating in the following way. Conditional on x_j, the observation $y_j - d_j$ is normally distributed with mean x_j and variance σ_{vj}^2. Let the prior distribution of x_j be Normal with mean $x_{j|j-1}$ and variance $Q_{j|j-1}$. Using the familiar Bayesian updating formula we see that the posterior distribution of x_j is Normal with mean $x_{j|j} = x_{j|j-1} + k_j[y_j - d_j - x_{j|j-1}]$, and variance $Q_{j|j} = [Q_{j|j-1}^{-1} + \sigma_{vj}^{-2}]^{-1}$. The gain is given by the familiar ratio of variances $k_j = \frac{Q_{j|j-1}}{Q_{j|j-1} + \sigma_{vj}^2}$. To derive the next prior note that since persistent disturbances have zero mean, the prior of x_{j+1} is Normal with mean $x_{j+1|j} = a_j x_{j|j}$, and variance $Q_{j+1|j} = a_j^2 Q_{j|j} + \sigma_{wj}^2$. To obtain the recursion in (2.8) simply let $\hat{x}_{j+1} := x_{j+1|j}$ and $Z_{j+1} := Q_{j+1|j}$.

As is well known, the price of asset i satisfies: $p_t^i = E_t \left(\frac{\beta[m - c_{t+1}^o]}{m - c_t} \pi_{t+1}^i \right)$, where the pricing kernel $\beta \frac{m - c_{t+1}^o}{m - c_t}$ is the intertemporal marginal rate of substitution, and π_{t+1}^i is the payoff of asset i at time $t + 1$. In equilibrium, asset prices must induce the representative agent to hold the supply of each asset and to consume all the dividends. That is, $\hat{q}_t^r = 1$, $\hat{q}_t^s = 0$, $\widehat{c}_t = y_t$ and $\widehat{c}_{t+1}^o = y_{t+1}$. Since the risky asset pays y_{t+1} at $t + 1$, and the safe asset pays one unit of the consumption good, their

---

[13] The recursion (2.8) is also the celebrated Kalman filter, which minimizes the mean square error of the state estimation error among all causal estimators (e.g. Kwakernaak and Sivan (1972)).

equilibrium prices are

$$\hat{p}_t^r(y^t) = E\left[\beta\frac{y_{t+1}-m}{y_t-m}\cdot y_{t+1}\,|I_t\right] = \beta\left[\frac{[\hat{x}_{t+1}+d_{t+1}-m]\,[\hat{x}_{t+1}+d_{t+1}]+Z_{t+1}+\sigma_{vt+1}^2}{y_t-m}\right]$$

$$\hat{p}_t^s(y^t) = E\left[\beta\frac{y_{t+1}-m}{y_t-m}\,|I_t\right] = \beta\left[\frac{\hat{x}_{t+1}+d_{t+1}-m}{y_t-m}\right] \qquad (2.9)$$

The second equality in the first row of (2.9) follows from (2.7) and $E_t(y_{t+1}^2) = E_t(y_{t+1})^2 + var_t(y_{t+1})$.

2.2. Equilibrium in an $\mathcal{H}_\infty$ Economy

In this subsection we will present the equilibrium of an $\mathcal{H}_\infty$ economy, and discuss the intuition behind it. To facilitate reading we will present the derivation of the results in Section 3.

In order to solve Problem $\mathcal{H}$ an agent must come up with a forecast of future dividends (we will denote it by $\mathcal{F}_t(y_{t+1})$). As we shall see, he uses all available information ($I_t = \{y_1, ..., y_t, p_t^r, p_t^s\}$) to derive such an estimate. However, he does not use the procedure of the previous subsection, in which dividends' forecasts are constructed using the underlying probability distributions, and without considering the portfolio strategy. Since in an $\mathcal{H}_\infty$ economy the agent has no information

25

about the disturbances, other than they have a bounded l_2 norm, his forecasts and portfolio choices are jointly determined.

As we shall see, equilibrium dividends' forecasts and portfolio strategies solve problem (3.2). Interestingly, this problem can be analyzed as a dynamic game between the agent and nature. In this game the agent forecasts dividends, and chooses his consumption and portfolio, while nature chooses disturbance sequences. The solution of this game is the subject of Section 3. The next Proposition summarizes the results.

Proposition 2.2. *Let $\underline{\gamma}_{t+1} = \sqrt{\beta\left[Z_{t+1} + \sigma_{vt+1}^2\right]}$, where Z_{t+1} is the tth element of recursion (2.8). Then*

- *If $\gamma_{t+1} \leq \underline{\gamma}_{t+1}$, there exists no equilibrium.*

- *If $\gamma_{t+1} > \underline{\gamma}_{t+1}$, there exists an equilibrium in which the $\mathcal{H}_\infty$ one-period-ahead forecast of dividends is*

$$\mathcal{F}_t(y_{t+1}) = x_{t+1}^* + d_{t+1} + \frac{x_{t+1}^* + d_{t+1} - m}{\gamma_{t+1}^2 \beta^{-1} \sigma_{vt+1}^{-2} - 1}, \qquad (2.10)$$

26

and the estimate of the unobservable state is:

$$x_{t+1}^* = \frac{\hat{x}_{t+1} + Z_{t+1}[d_{t+1} - m]\left[\gamma_{t+1}^2\beta^{-1} - \sigma_{vt+1}^2\right]^{-1}}{1 - Z_{t+1}\left[\gamma_{t+1}^2\beta^{-1} - \sigma_{vt+1}^2\right]^{-1}}, \qquad (2.11)$$

where $\hat{x}_{t+1}$ and Z_{t+1} are the tth elements of recursion (2.8). The equilibrium prices of the risky and safe assets are:

$$p_t^{r*}(y^t) = \beta\frac{\mathcal{F}_t(y_{t+1}) - m}{y_t - m} \cdot \mathcal{F}_t(y_{t+1}), \qquad p_t^{s*}(y^t) = \beta\frac{\mathcal{F}_t(y_{t+1}) - m}{y_t - m} \qquad (2.12)$$

The proof is in Section 3. Here we simply make some remarks. First, in an $\mathcal{H}_\infty$ setup agents arrive at an estimate of the unobservable state x_{t+1}^* using a recursion that has the same 'observer' form as the one we used in the RE economy. To see this note that x_{t+1}^* is a linear function of $\hat{x}_{t+1}$, which is obtained recursively by adding to $\hat{x}_j$ a term proportional to the innovation (i.e., $\hat{x}_{j+1} = a_j\hat{x}_j + h_j[y_j - d_j - \hat{x}_j]$). The forecast of dividends ($\mathcal{F}_t(y_{t+1})$) is then obtained by adding to x_{t+1}^* the trend coefficient d_{t+1}, and the forecast of the disturbance v_{t+1}. A non-zero estimate of v_{t+1} reflects the fact that robust agents take into account that the disturbance might be highly correlated with the state.

Recall that the parameter γ_{t+1} is an index of robustness. The lower γ_{t+1}, the

greater the degree of robustness (see (2.6)). Note that when agents in an $\mathcal{H}_\infty$ economy do not care about robustness, their forecasts are equal to the rational expectations forecasts. That is, if $\gamma_{t+1} \to \infty$, the $\mathcal{H}_\infty$ one-period-ahead dividend forecast $\mathcal{F}_t(y_{t+1})$ coincides with the expected value of dividends in an RE economy $E_t(y_{t+1})$. To see this take the limit of (2.10) and (2.11) as $\gamma_{t+1} \to \infty$ to get

$$\lim_{\gamma_{t+1} \to \infty} \mathcal{F}_t(y_{t+1}) - d_{t+1} = \lim_{\gamma_{t+1} \to \infty} x^*_{t+1} = \hat{x}_{t+1}$$

A comparison with (2.7) reveals that $\lim_{\gamma_{t+1} \to \infty} \mathcal{F}_t(y_{t+1}) = E_t(y_{t+1})$. This similarity is noteworthy and surprising, given that the problems and induced objectives are quite different.

The second point we would like to emphasize is that the degree of robustness that can be attained in equilibrium cannot be arbitrarily large. Proposition 2.2 says that an equilibrium exists only if $\gamma^2_{t+1} > \underline{\gamma}^2_{t+1}$. Since $\underline{\gamma}^2_{t+1} = \beta \left[Z_{t+1} + \sigma^2_{vt+1} \right]$, it follows that ceteris paribus, the maximum degree of robustness that can be attained is inversely related to the 'intensities' of the transitory disturbance σ_{vt+1}, and of the persistent disturbance σ_{wt} (recall that $y_{t+1} = d_{t+1} + x_{t+1} + \sigma_{vt+1} v_{t+1}$, and $Z_{t+1} = \frac{a_t^2}{Z_t^{-1} + \sigma_{vt}^{-2}} + \sigma^2_{wt}$). As it will become clear in Section 3, the reason for the existence of this lower bound on γ_{t+1} follows from the fact that the value of

28

the game between the agent and nature is bounded only if $\gamma_{t+1} > \underline{\gamma}_{t+1}$.

Finally, we would like to emphasize that the formal similarities we have noted hide important differences in the way solutions are derived. In the RE economy we made forecasts of uncertain variables without reference to the agents' control policies. That is, we derived the probability distribution of y_{t+1} without taking into consideration the agents' portfolio and consumption policies. This 'separation' does not hold in an $\mathcal{H}_\infty$ setup. Since disturbance sequences are totally unknown, forecasts and agents' control policies must be jointly determined. In Section 3 we will make clear the nature of this interdependency.

2.3. Trade-off: Robustness vs. Conditional Optimality

In order to illustrate the trade-off involved in using the RE and $\mathcal{H}_\infty$ forecasting formulas, we will do some simulations and compare their performance under a familiar metric: the mean square error (MSE) of one-period-ahead dividend forecasts. We will see that if there is no-misspecification, RE forecasting formulas generate the lowest MSE, as expected. However, in the presence of a large misspecification $\mathcal{H}_\infty$ formulas generate a lower MSE.

Consider the following counterfactual experiment in which disturbances are drawn from a specific white noise process. Clearly, if the parameters (denote

29

them by $\tilde{\theta}$) used in the forecasting formulas were equal to the parameters that characterize the data generating process (denote them by θ^{true}), then the MSE of the RE forecasts would be lower than the MSE of the $\mathcal{H}_\infty$ forecasts. This is because the RE forecasting formulas (2.8) correspond to the Kalman filter, which in this case minimizes the MSE of the state estimation error. Suppose instead that $\mathcal{H}_\infty$ and RE forecasts are constructed using $\tilde{\theta}$ in the forecasting formulas, but that the data generating process is characterized by $\theta^{true} = \tilde{\theta} + \Delta^{true}$. The simulations we perform show that if the misspecification $|\Delta^{true}|$ is large, then $\mathcal{H}_\infty$ forecasts have a lower MSE than RE forecasts.

Panel (a) in Figure 1 considers the case in which the intensity of the persistent disturbance might be misspecified ($\sigma_w^{true} = \tilde{\sigma}_w + \Delta^{true}$), but RE and $\mathcal{H}_\infty$ dividend forecasts are constructed setting $\sigma_w = \tilde{\sigma}_w$ in forecasting formulas (2.8) and (2.10). As we can see, when there is no misspecification (i.e. $\Delta^{true} = 0$, so that $\sigma_w^{true} = \tilde{\sigma}_w$), the RE forecasts produce a lower MSE than the $\mathcal{H}_\infty$ forecasts. However, if $|\Delta^{true}|$ is big, the MSE of the $\mathcal{H}_\infty$ forecasts is lower that the MSE of RE forecasts. Panels (b) and (c) show that the same holds true in the case where the intensity of the observation noise is misspecified ($\sigma_v^{true} = \tilde{\sigma}_v + \Delta^{true}$), or the autoregressive parameter is misspecified ($a^{true} = \tilde{a} + \Delta^{true}$), respectively.
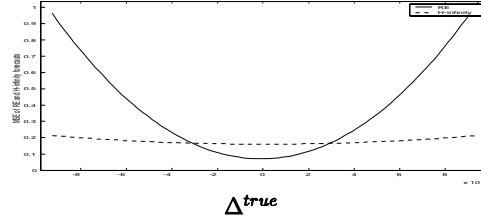
It is interesting to note that although the $\mathcal{H}_\infty$ forecasting formula (2.10) has

been designed to solve Problem $\mathcal{H}$, it has good robustness properties under the MSE metric.
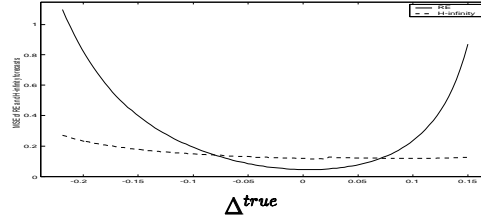
Figure 1: Mean Square Error of Dividends' Forecasts in the Presence of Misspecification
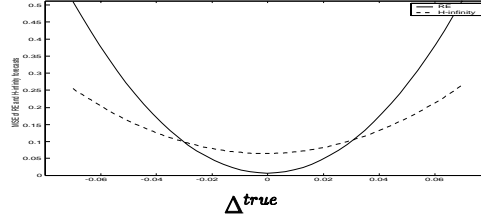
(a) $\sigma_w^{true} = \tilde{\sigma}_w + \Delta^{true}$

(b) $\sigma_v^{true} = \tilde{\sigma}_v + \Delta^{true}$

(c) $a^{true} = \tilde{a} + \Delta^{true}$

Note: In panel (a) forecasts are made setting the intensity of persistent shocks equal to $\tilde{\sigma}_w$, while the 'true' intensity is $\sigma_w^{true} = \tilde{\sigma}_w + \Delta^{true}$. For each value of Δ^{true} we generate 100 artificial dividend sequences. Then for each dividend sequence h we compute the mean square error of the forecast: $MSE_h(\tilde{\sigma}_w + \Delta^{true}) = \sum_{j=1}^{99}(y_{j+1} - \mathcal{F}_j(y_{j+1}))^2/100$. The $\mathcal{H}_\infty$ curve plots the average MSE for different values of Δ^{true} : $\overline{MSE}(\tilde{\sigma}_w + \Delta^{true}) = \sum_{h=1}^{100} MSE_h(\tilde{\sigma}_w + \Delta^{true})/100$. The RE curve is constructed in the same way. In panel (b) we let $\sigma_v^{true} = \tilde{\sigma}_v + \Delta^{true}$, and set $\sigma_w^{true} = \tilde{\sigma}_w$. Lastly in panel (c) we let $a^{true} = \tilde{a} + \Delta^{true}$. For details see the Appendix.

3. Derivation of Equilibrium in an $\mathcal{H}_\infty$ Economy

In this section we convert Problem $\mathcal{H}$ into a dynamic game between the agent and nature, and show how to solve it. We also determine necessary and sufficient conditions for an equilibrium to exist, and derive the equilibrium defined in Proposition 2.2. On a first pass the reader might wish to skip ahead to the simulations of Section 4. However, to fully appreciate the intuition behind the $\mathcal{H}_\infty$ approach the reader should come back to this section.

In order to simplify notation we will express disturbances, prices and assets' demands as follows:

$$\omega_j := (w_j, v_j) \qquad p := (p_t^r, p_t^s) \qquad q := (q_t^r, q_t^s)$$

Also, for any sequence $\{z_j\}_{j=0}^t$ we will often use the following notation: $z := \{z_j\}_{j=0}^t$ and $z^n := \{z_j\}_{j=0}^n$ for any $n < t$. Admissible portfolio strategies can depend only on available information $I_t = \{y_1, ..., y_t, p_t^r, p_t^s\}$. Furthermore, the

quantity demanded of each asset must lie in [0,1]. That is,

$$q = \chi(y_1, ..., y_t, p_t^r, p_t^s)$$

where $\chi : \mathcal{R}^{t+2} \longmapsto [0,1]^2$. We will denote by $\mathcal{Q}$ the space of such portfolio policies.

In equilibrium the representative agent solves Problem $\mathcal{H}$ taking prices as given, and asset and goods markets clear. Using budget constraints (2.1) and (2.2) to substitute out c_t and c_{t+1}^o, it follows that (2.6) in Problem $\mathcal{H}$ is equivalent to:

$$J(q, \omega; p) := L(q, y_{t+1}; p) + \sum_{j=0}^{t+1} g_j(\omega_j) \leq 0 \qquad (3.1)$$

where

$$L(q, y_{t+1}; p) \quad : \quad = [m - y_t - p_t^r[1 - q_t^r] + p_t^s q_t^s]^2 + \beta[m - q_t^r y_{t+1} - q_t^s]^2$$

$$g_j(\omega_j) \quad : \quad = -\gamma_{t+1}^2 \left[w_j^2 + v_j^2\right]$$

Since the denominator in (2.6) is positive, for given prices, a portfolio policy $q = \chi(I_t) \in \mathcal{Q}$ solves Problem $\mathcal{H}$ if and only if $J(q, \omega; p) \leq 0$ for all nonzero disturbance sequences $\omega \in \Omega$ that are consistent with information set I_t. Clearly,

34

this holds if and only if $\Gamma(p)$, defined below, is bounded above by zero:

$$\Gamma(p) := \quad \inf_{\chi \in \mathcal{Q}} \, \sup_{\omega \in \Omega} \, J(q, \omega; p) \leq 0 \qquad\qquad (3.2)$$
$$\text{subject to} \quad v_j = \frac{y_j - d_j - x_j}{\sigma_{vj}}, \qquad j = 1, ..., t$$

Note that the constraint imposed by dividends' observations $\{y_j\}_{j=1}^{t}$ on distur-

bance sequences is given by the sequence $\{v_j = \frac{y_j - d_j - x_j}{\sigma_{vj}}\}_{j=1}^{t}$. To see the close link

between the existence of a solution to Problem $\mathcal{H}$ and the existence of a saddle-

point solution to (3.2), denote by $(\chi^*(I_t), \omega^*(I_t))$ a stationary point of $J(q, \omega; p)$.

Then it is straightforward that $\Gamma(p) \leq 0$ is necessary for Problem $\mathcal{H}$ to have

a solution because $J(\chi(I_t), \omega^*(I_t); p) \geq \Gamma(p)$ for any $\chi(I_t) \neq \chi^*(I_t)$. Thus, if

for $q^* = \chi^*(I_t)$ the supremum in (3.2) is unbounded, there exists a disturbance

sequence for which the ratio in (2.6) is greater than γ_{t+1}. Next, $\Gamma(p) \leq 0$ is suf-

ficient for Problem $\mathcal{H}$ to have a solution because $q = \chi^*(I_t)$ is a portfolio policy

that ensures the ratio in (2.6) is less or equal than γ_{t+1} for all square summable

disturbance sequences (i.e., $\Gamma(p) = J(\chi^*(I_t), \omega^*(I_t); p) \leq J(\chi^*(I_t), \omega; p))$.

It follows that an equilibrium exists if and only if there is a price vector

$p^*(y_1, ..., y_t)$ that induces the representative agent to demand the available net

35

supply of assets (i.e., $\chi^*(I_t) = (1,0)$), and $\Gamma(p^*(y_1, ..., y_t))$ is bounded. As we shall see, since $J(q, \omega; p)$ is an indeterminate quadratic form in ω, the value of the disturbance attenuation parameter γ_{t+1} is the key to whether $\Gamma(p^*(y_1, ..., y_t))$ is bounded. This observation should make clear the convenience of a quadratic utility.

Note that problem (3.2) can be analyzed as a dynamic game between the representative agent and nature. In this game the agent chooses his portfolio, while nature chooses disturbance sequences. Clearly, this is simply an analogy, as not many of us believe that nature actually cares about our portfolio strategies. However, this observation has been useful because it has allowed researchers to solve complicated $\mathcal{H}_\infty$-control problems using dynamic game theory.

As it stands, problem (3.2) seems quite complicated. First, portfolio strategies can be selected from a very large class of functions ($q = \chi(y_1, ..., y_t, p_t^r, p_t^s)$). Second, disturbance sequences ($\{v_j\}_{j=0}^{t+1}$, $\{w_j\}_{j=0}^{t+1}$) are not restricted to follow any specific process, and can be highly correlated. Theorem 3.1 below says that we can break Problem (3.2) into three simple sub-problems. This insight, which we might call the '$\mathcal{H}_\infty$ certainty equivalence principle,' is due to Basar and Bernhard (1991).

The key point is to assume temporarily that players are acting _as if_ they knew the state x_{t+1} equals a certain value x. The first sub-problem is to find

the maximally malevolent sequence of past disturbances (from the perspective of objective (3.2)) that are consistent with dividend history $\{y_j\}_{j=1}^{t}$ and that bring the unobservable state from $x_0 = 0$ to a certain value x at time $t+1$. The second sub-problem determines the consumption and portfolio strategies, as well as the disturbance ω_{t+1} under the assumption that the unobservable state x_{t+1} takes the value x. Lastly, the third sub-problem generates the estimate of the state (x_{t+1}^*) and the price vector that clears asset markets $p^*(x_{t+1}^*)$.

This decomposition can be carried out because the dynamic system we are considering is Markovian. Recall that the unobservable state satisfies $x_{t+1} = a_t x_t + \sigma_{wt} w_t$, and that dividends are $y_{t+1} = d_{t+1} + x_{t+1} + \sigma_{vt+1} v_{t+1}$. Loosely speaking, conditional on $x_{t+1} = x$, past disturbances $\{\omega_j\}_{j=0}^{t}$ do not affect $L(q, y_{t+1}; p) + g_{t+1}(\omega_{t+1})$ in (3.2). Similarly, conditional on $x_{t+1} = x$, the choice of (q, ω_{t+1}) does not affect $\sum_{j=0}^{t} g_j(\omega_j)$. Therefore, conditional on $x_{t+1} = x$, the first two subproblems can be solved independently of each other. The third subproblem uses the value functions associated with the first two sub-problems to generate the estimate of x_{t+1}.

We proceed by formally stating the three subproblems. In order to define the first sub-problem we will denote the state's and dividends' trajectories generated

by the disturbance subsequence $\omega^s := \{\omega_j\}_{j=0}^s$ as follows:

$$x_s = X_s(\omega^{s-1}), \qquad\qquad y_s = Y_s(\omega^s), \quad s \geq 1 \qquad (3.3)$$

We can then define the following subsets of Ω (see (2.5))

$$\Omega_s := \{\{\omega_i\}_{i=0}^{t+1} \in \Omega \mid Y_j(\omega^j) = y_j \quad \text{for all } j \leq s\} \qquad (3.4)$$

That is, Ω_s is the set of disturbance sequences $\{\omega_j\}_{j=0}^{t+1}$ that are compatible with observed dividends $\{y_j\}_{j=1}^s$. Using this notation and (3.2) we can express the first sub-problem in terms of the following 'cost-to-come function'

$$
\begin{aligned}
W_{t+1}(x) = \sup_{\{\omega_j\}_{j=0}^t \in \Omega_t} -\gamma_{t+1}^2 \sum_{j=0}^{\mathbf{t}} \left[w_j^2 + v_j^2\right] \\
\text{subject to} \quad x_{t+1} = x
\end{aligned}
\qquad (3.5)
$$

The only information an agent has about the disturbances $\{\omega_j\}_{j=0}^t$ is that they are square summable sequences, and that they have generated a dividend history $\{y_j\}_{j=1}^t$ according to the dynamic system (2.3). Solving problem (3.5) is an intermediate step that allows the agent to estimate x_{t+1}. It characterizes the disturbance sequences that make (3.1) less likely to hold, given that they bring the state from $x_0 = 0$ to $x_{t+1} = x$, and are consistent with dividend history.

38

Consider now the second sub-problem. In an $\mathcal{H}_\infty$ setup even if an agent knew the value of the unobservable state x_{t+1}, he would not forecast that dividends (y_{t+1}) will be $x_{t+1} + d_{t+1}$. This is because there is no reason to expect that the disturbance v_{t+1} will be identically zero. In fact the agent forecasts v_{t+1} jointly with the choice of his portfolio and consumption. This is formalized by the following problem in which the unobservable state x_{t+1} is again assumed to take the value x.

$$V_{t+1}(x;p) = \inf_{\tilde{\chi} \in \mathcal{Q}_x} \sup_{\tilde{v} \in \mathcal{V}_x} \left\{ L(q, y_{t+1}; p) - \gamma_{t+1}^2 v_{t+1}^2 \right\}$$

$$\text{subject to} \qquad x_{t+1} = x \qquad\qquad (3.6)$$

$$y_{t+1} = x_{t+1} + d_{t+1} + \sigma_{vt+1} v_{t+1}$$

The set of admissible portfolio strategies $\mathcal{Q}_x$ consists of the 'Markov strategies' that are conditional on the state x_{t+1} being equal to x. That is, $q = \tilde{\chi}(x,p)$ with $\tilde{\chi} : \mathcal{R}^3 \longmapsto [0,1]^2$. The set of admissible strategies for the disturbance $\mathcal{V}_x$ consists of the Markov strategies $v_{t+1} = \tilde{v}(x,q)$ with $\tilde{v} : \mathcal{R}^3 \times [0,1]^2 \longmapsto \mathcal{R}$.[14]

Four observations are in order. First, a major simplification relative to the original problem (3.2) is that strategies are functions of (x,p), not of the entire

---

[14]Markov strategies are also known as feedback strategies. These strategies are closed-loop strategies in which history matters only through its effect on the current state. See Basar and Olsder (1995).

history of observations $(\{y_j\}_{j=1}^t, p)$. The name '$\mathcal{H}_\infty$ certainty equivalence principle' derives from the fact that in this sub-problem agents make forecasts, as well as consumption and portfolio decisions, acting _as if_ the state x_{t+1} takes the value of x. Second, note that the disturbance v_{t+1} has access to the realization of q in order to ensure that the portfolio policy is robust. Third, no hard bound has been imposed on the disturbance v_{t+1}. As we shall see, if problem (3.6) has a solution, the disturbance will be bounded in equilibrium. Finally, note that w_{t+1} does not appear in (3.6). Recall that throughout the paper we have set $w_{t+1} = 0$ because it affects neither x_{t+1} nor y_{t+1}.

The third sub-problem is to determine the $\mathcal{H}_\infty$ forecast of x_{t+1} (call it x_{t+1}^*), and of dividends y_{t+1}^*. To solve this sub-problem we will denote the price vector that clears assets and goods markets' when $x_{t+1} = x$ by $p^*(x) := (p_t^{r*}(x), p_t^{s*}(x))$, and the value of the "cost-to-go" value function (3.6) evaluated at these prices by $V_{t+1}(x)$. That is,

$$V_{t+1}(x) := V_{t+1}(x; p^*(x)) \tag{3.7}$$

As the next Theorem states, if it exists, the $\mathcal{H}_\infty$ estimate of the state is

$$x_{t+1}^* \in \arg\max \{V_{t+1}(x) + W_{t+1}(x)\} \tag{3.8}$$

Replacing x_{t+1}^* in the policies that solve (3.6) we will obtain the price vector that clears assets and goods markets, and the $\mathcal{H}_\infty$ forecasts of dividends

$$\mathcal{F}_t(y_{t+1}) = x_{t+1}^* + d_{t+1} + \sigma_{vt+1} v^*(x_{t+1}^*). \tag{3.9}$$

The following Theorem provides necessary and sufficient conditions for an equilibrium to exist, and states that if an equilibrium exists, it can be constructed using the algorithm we have just described.

Theorem 3.1. *Consider an economy in which disturbances satisfy (2.5), the representative agent observes $(y_1, ..., y_t, p_t^r, p_t^s)$, and his objective is (2.6). Then, for a given degree of robustness γ_{t+1}:*

i. *There exists an equilibrium if and only if there are bounded functions $V_{t+1}(x)$ and $W_{t+1}(x)$, given by (3.6) and (3.5), which satisfy*

$$\sup_x \Lambda(x) := \sup_x \{V_{t+1}(x) + W_{t+1}(x)\} < \infty \tag{3.10}$$

ii. *(Certainty Equivalence) If (3.10) holds and $\Lambda(x)$ is strictly concave, there exists a unique Markov perfect equilibrium. Consumption (c_t^*), asset demands (q_t^{r*}, q_t^{s*}) and forecasted dividends $\mathcal{F}_t(y_{t+1})$ are jointly determined by the solu-*

41

tion to problem (3.6) taking the state to be $x_{t+1}^* = \arg\max \Lambda(x)$. Equilibrium

asset prices are such that $c_t^* = y_t$, $q_t^{r*} = 1$ and $q_t^{s*} = 0$.

This Theorem implies that one can determine whether an equilibrium exists by considering only Markov strategies (i.e., strategies that only depend on the estimate of the state x_{t+1}, and prices). If a Markov perfect equilibrium (MPE) of the game between the agent and nature does not exist, then there exists no other closed-loop equilibrium in which portfolio and disturbance strategies are more complicated functions of history $\{y_j\}_{j=1}^t$. Note, however that this Theorem does not say that there is a unique closed-loop equilibrium. If an MPE exists, there might exist other equilibria in which strategies have the same open-loop representation as the Markov strategies.[15]

In order to verify the conditions stated in Theorem 3.1, and obtain a closed-form formula for the dividend's forecast and for asset prices we need to derive in closed form the value functions $W_{t+1}(x)$ and $V_{t+1}(x)$. We will proceed by first

---

[15] In this footnote we make a brief detour and explain why the name $\mathcal{H}_\infty$ is used. Let $\mathcal{G}_\chi$ be a linear operator that maps an input sequence $\{\omega_j\}_{j=0}^{t+1}$ to a certain objective under control policy χ. The $\mathcal{H}_\infty$ induced norm of the operator $\mathcal{G}_\chi$ is then defined as

$$\parallel \mathcal{G}_\chi \omega \parallel_\infty = \sup_{\omega \in l_{2,[0,t+1]}, \omega \neq 0} \frac{\parallel \mathcal{G}_\chi \omega \parallel_{2[0,t+1]}}{\parallel \omega \parallel_{2[0,t+1]}}$$

where $\parallel \omega \parallel_{2[0,t+1]}$ is the l_2-norm of the sequence ω. If we let the numerator be equal to the first two terms in $J(\chi(I_t), \omega; p)$ in (3.1), we can restate Problem $\mathcal{H}$ as finding portfolio policies that achieve $\parallel \mathcal{G}_\chi \omega \parallel_\infty < \gamma$, and finding $\underline{\gamma} = \inf_{\chi \in \mathcal{Q}} \parallel \mathcal{G}_\chi \omega \parallel_\infty$.

deriving these functions, and then applying part (ii) of the Theorem to derive the formulas stated in Proposition 2.2. As we shall see, existence condition (3.10) holds if and only if the disturbance attenuation parameter satisfies: $\gamma_{t+1} > \underline{\gamma}_{t+1}$.

3.1. Derivation of the Cost-to-Come Value Function $\mathbf{W}_{t+1}(x)$

We will derive $W_{t+1}(x)$ by representing (3.5) as a recursive problem. In order to do this let $\Omega_s(x|y^s)$ be the set of admissible disturbance sequences that bring the state to level x at time $s+1$, and that are consistent with dividend history $\{y_j\}_{j=1}^s$

$$\Omega_s(x|y^s) := \left\{ \omega \in \Omega_s \mid x = X_s(\omega^{s-1}) \right\}, \quad s \in \{1, ..., t\} \tag{3.11}$$

Analogously to (3.5) we can then define the cost-to-come value function, conditional on information up to time s as

$$W_{s+1}(x) = \sup_{\omega^s \in \Omega_s(x|y^s)} -\gamma_{t+1}^2 \sum_{j=0}^{s} \left[w_j^2 + v_j^2 \right], \quad s \in \{1, ..., t\} \tag{3.12}$$

If (3.12) has a finite solution, it satisfies the following *forward dynamic programing* equation

43

$$W_{s+1}(x) = \begin{cases} \max_{(\xi, w, v)} \left\{ W_s(\xi) - \gamma_{t+1}^2 \left[w^2 + v^2 \right] \right\} \\ \\ \text{subject to} \quad \begin{aligned} x &= a_s \xi + \sigma_{ws} w \\ y_s &= \xi + d_s + \sigma_{vs} v \end{aligned} \end{cases} \qquad s \in \{1, ..., t\} \qquad (3.13)$$

$$W_1(x) = -x^2 \gamma_{t+1}^2 \sigma_{w0}^{-2}$$

Note that (ξ, w, v) corresponds to (x_s, w_s, v_s) and that x corresponds to x_{s+1}.

Forward dynamic programing problems are solved in a similar way to standard *backward* DP problems. The difference is that they are solved starting at initial time, not terminal time. To solve problem (3.13) we need to find a closed form solution for the value function $W_{s+1}(x)$. Since $W_{s+1}(x)$ is the supremum of a quadratic function subject to an affine constraint, it is quadratic in x. The next Lemma, which is proved in the Appendix, provides the solution.

Lemma 3.2 (Cost-to-Come Value Function). *For any* $\gamma_{t+1} > 0$ *the solution to (3.13) is*

$$W_{s+1}(x) = -[x - \check{x}_{s+1}]^2 K_{s+1} + l_{s+1}, \qquad s \in \{1, ..., t\} \qquad (3.14)$$

44

where l_{s+1} is independent of x, and $(K_{s+1}, \check{x}_{s+1})$ satisfy the recursion: $K_1 = \gamma_{t+1}^2 \sigma_{w0}^{-2}$, $\check{x}_1 = 0$, and for $s \in \{1, ..., t\}$

$$\check{x}_{s+1} = a_s \check{x}_s + \frac{a_s}{1 + \sigma_{vs}^2 \gamma_{t+1}^{-2} K_s} [y_s - d_s - \check{x}_s]$$

$$K_{s+1} = \left[\frac{a_s^2}{P_s} + \gamma_{t+1}^{-2} \sigma_{ws}^2 \right]^{-1}, \qquad \text{with} \quad P_s = \gamma_{t+1}^2 \sigma_{vs}^{-2} + K_s \qquad (3.15)$$

Notice the similarity between recursion (3.15) and filtering formula (2.8) that we used in the RE economy.

3.2. Derivation of the Cost-to-Go Value Function $\mathbf{V}_{t+1}(x)$

We can think of problem (3.6) as a zero-sum game between the agent and nature in which the agent chooses q_t^r and q_t^s, while nature chooses the disturbance v_{t+1}. In this game the state $(x_{t+1} = x)$ is perfectly observed by both players, and nature has access to q_t^r and q_t^s. In a Markov perfect equilibrium of this game the strategies satisfy:

$$\tilde{v}_{t+1}^*(x, q; p) = \begin{cases} \frac{[m - q_t^r d_{t+1} - q_t^r x - q_t^s] q_t^r \sigma_{vt+1}}{q_t^{r2} \sigma_{vt+1}^2 - \beta^{-1} \gamma_{t+1}^2} & \text{if } \gamma_{t+1}^2 > \beta q_t^{r2} \sigma_{vt+1}^2 \\ \infty & \text{otherwise} \end{cases} \qquad (3.16)$$

45

$$\tilde{\chi}_t^{r*}(x; p) = \frac{p_t^r \left[y_t + p_t^r - m - \tilde{\chi}_t^{s*} p_t^s \right] + \beta \left[m - \tilde{\chi}_t^{s*} \right] y_{t+1}^*(x)}{p_t^{r2} + \beta y_{t+1}^*(x)^2} \qquad (3.17)$$

$$\tilde{\chi}_t^{s*}(x; p) = \frac{p_t^s \left[y_t + p_t^r - m - \tilde{\chi}_t^{r*} p_t^r \right] + \beta \left[m - \tilde{\chi}_t^{r*} y_{t+1}^*(x) \right]}{p_t^{s2} + \beta} \qquad (3.18)$$

For any given value of x_{t+1} (say x), equilibrium prices $p^*(x)$ induce the representative agent to set $\tilde{\chi}_t^{r*}(x, p^*(x)) = 1$ and $\tilde{\chi}_t^{s*}(x, p^*(x)) = 0$. It follows from (3.16) that at equilibrium prices the disturbance is $v_{t+1}^*(x) = \frac{[m - d_{t+1} - x]\sigma_{vt+1}}{\sigma_{vt+1}^2 - \beta^{-1}\gamma_{t+1}^2}$. Replacing this in (3.6) we have that at equilibrium prices the cost-to-go value function $V_{t+1}(x; p_t^{r*}(x), p_t^{s*}(x))$ is equal to

$$V_{t+1}(x) = \begin{cases} \frac{[m - d_{t+1} - x]^2}{\beta^{-1} - \gamma_{t+1}^{-2}\sigma_{vt+1}^2} + [m - y_t]^2 & \text{if } \gamma_{t+1}^2 > \beta\sigma_{vt+1}^2 \\ \infty & \text{otherwise} \end{cases} \qquad (3.19)$$

A key point to notice is that the kernel in (3.6) is an indefinite quadratic form in v_{t+1}. Depending on the value of γ_{t+1}, it might or might not be concave in v_{t+1}. This indefiniteness is key in pinning down the best performance $\underline{\gamma}_{t+1}$. If the condition on γ_{t+1} specified in (3.16) is not satisfied, then in equilibrium the kernel of (3.6) is not concave in v_{t+1}, and Problem $\mathcal{H}$ has no solution. Since at equilibrium prices $q_t^{r*} = 1$ and $q_t^{s*} = 0$, it follows from (3.16) that a necessary condition for the $\mathcal{H}_\infty$ disturbance v_{t+1}^* and the value function $V_{t+1}(x)$ to be bounded in equilibrium is

that $\gamma_{t+1}^2 > \beta\sigma_{vt+1}^2$.

3.3. Derivation of x_{t+1}^* and Proof of Proposition 2.2

The first two sub-problems were solved under the assumption that the state x_{t+1}

had a particular value x. The value functions associated with these sub-problems

are $V_{t+1}(x)$ and $W_{t+1}(x)$. The third sub-problem is to obtain an estimate of x_{t+1}.

Theorem (3.1) says that $x_{t+1}^* = \arg\max\{V_{t+1}(x) + W_{t+1}(x)\}$. Using (3.14) and

(3.19) we have that

$$
x_{t+1}^* = \begin{cases} \dfrac{m - d_{t+1} - \left[\beta^{-1} - \gamma_{t+1}^{-2}\sigma_{vt+1}^2\right]K_{t+1}\check{x}_{t+1}}{1 - \left[\beta^{-1} - \gamma_{t+1}^{-2}\sigma_{vt+1}^2\right]K_{t+1}} & \text{if } \gamma_{t+1}^2 > \underline{\gamma}_{t+1}^2 := \dfrac{\sigma_{vt+1}^2}{\beta^{-1} - K_{t+1}^{-1}} \\[4mm] \infty & \text{if } \gamma_{t+1}^2 \le \underline{\gamma}_{t+1}^2 \end{cases} \tag{3.20}
$$

Note that $[V_{t+1}(x) + W_{t+1}(x)]$ is concave in x if and only if $K_{t+1} > [\beta^{-1} -$

$\gamma_{t+1}^{-2}\sigma_{vt+1}^2]^{-1}$. Recall that the function $V_{t+1}(x)$ is bounded if and only if $\gamma_{t+1}^2 >$

$\beta\sigma_{vt+1}^2$. Since this last condition is implied by the concavity of $[V_{t+1}(x) + W_{t+1}(x)]$,

it follows that x_{t+1}^* is finite if and only if $\gamma_{t+1}^2 > \underline{\gamma}_{t+1}^2$.

Note that the expression for x_{t+1}^* in (3.20) has different terms than (2.11)

in Proposition 2.2. The former contains $(K_{t+1}, \check{x}_{t+1})$, while the latter contains

$(Z_{t+1}, \hat{x}_{t+1})$. In order to show that (3.20) and (2.11) are indeed identical simply

make the substitution $Z_s := \gamma_{t+1}^2 K_s^{-1}$ in recursion (3.15) and in (3.20). It is

straightforward to verify that $\check{x}_s = \hat{x}_s$ for all $s \in \{1, ..., t\}$ and that $\gamma_{t+1}^2 > \frac{\sigma_{vt+1}^2}{\beta^{-1} - K_{t+1}^{-1}}$

is equivalent to $\gamma_{t+1}^2 > \beta[Z_{t+1} + \sigma_{vt+1}^2]$.

Next, we verify that $\gamma_{t+1}^2 > \underline{\gamma}_{t+1}^2$ ensures that the value functions $V_{t+1}(x)$

and $W_{t+1}(x)$ are bounded. We show in the proof of Lemma 3.2 that for all

$\gamma_{t+1} > 0$ the value function $W_{t+1}(x)$ is bounded and that $K_{t+1} > 0$. To show

that $V_{t+1}(x)$ is bounded if $\gamma_{t+1}^2 > \underline{\gamma}_{t+1}^2$ note that $Z_{t+1} = \gamma_{t+1}^2 K_{t+1}^{-1} > 0$. Thus,

$\underline{\gamma}_{t+1}^2 := \beta \left[Z_{t+1} + \sigma_{vt+1}^2 \right] > \beta \sigma_{vt+1}^2$. Since $V_{t+1}(x)$ is bounded if and only if $\gamma_{t+1}^2 >$

$\beta \sigma_{vt+1}^2$, the result follows.

Equilibrium prices are determined by substituting x_{t+1}^* in (3.17) and (3.18)

and imposing asset market clearing. That is, equilibrium prices are such that

$\chi_t^{r*}(x_{t+1}^*, p^*(x_{t+1}^*)) = 1$ and $\chi_t^{s*}(x_{t+1}^*, p^*(x_{t+1}^*)) = 0$. This system of equations has

a unique solution given by (2.12) in Proposition 2.2. Dividends' forecast in (2.10)

are obtained by replacing (3.20) and $v_{t+1}^*(x) = \frac{[m - d_{t+1} - x] \sigma_{vt+1}}{\sigma_{vt+1}^2 - \beta^{-1} \gamma_{t+1}^2}$ in (2.3).

Summing up, we have shown that for every $\gamma_{t+1}^2 > \underline{\gamma}_{t+1}^2$ there is a unique

Markov price vector $p^*(x_{t+1}^*)$, such that $\tilde{\chi}_t^{r*}(x_{t+1}^*, p^*(x_{t+1}^*)) = 1$ and $\tilde{\chi}_t^{s*}(x_{t+1}^*, p^*(x_{t+1}^*)) =$

0 solve Problem $\mathcal{H}$. Furthermore, given market clearing, the condition $\gamma_{t+1}^2 > \underline{\gamma}_{t+1}^2$

is necessary and sufficient for the disturbance sequence $\{w_j, v_j\}_{j=0}^{t+1}$ to be square

summable. This completes the proof of Proposition 2.2.

4. Asset Pricing Puzzles

In this Section we simulate dividend data by drawing shocks from a random number generator, and constructing artificial RE and $\mathcal{H}_\infty$ price sequences. We then verify whether these price sequences exhibit predictability of excess returns, excess volatility, and a high equity premium while maintaining a low risk-free interest rate. Lastly, instead of using simulated dividends, we feed the RE and $\mathcal{H}_\infty$ price formulae with historical US dividend data, and analyze how well they track the actual US stock market.[16]

Throughout this section we will assume that the data generating process is given by (2.3), that the disturbances $\{v_j\}_{j=0}^{t+1}$, $\{v_j\}_{j=0}^{t+1}$ are i.i.d. $N(0,1)$, and that the parameters of the data generating process $(a, d, \sigma_w, \sigma_v)$ are constant. We set them equal to their maximum likelihood estimates associated with US stock market data for the period 1871-1997 (see (6.6)). We then construct sequences of forecasts and prices using formulas (2.9) and (2.12) setting the discount factor $\beta = 0.95$, and the disturbance attenuation parameter equal to a constant $\gamma > \underline{\gamma}_{t+1}$.

---

[16] As we mentioned in the Introduction, these anomalies can be rationalized without invoking $\mathcal{H}_\infty$ forecasting. For instance, Campbell and Cochrane (1999) reproduce the above-mentioned puzzles by adding an external time-varying habit to the standard power utility function. Barberis, et. al. (1999) consider preferences that are dependent on prior outcomes, and that are more sensitive to reductions in wealth than to increases. Constantinides and Duffie (1996), and Heaton and Lucas (1996) consider the case of market incompletness, while Timmerman (1996) considers a lerning setup.

Lastly, we choose the value of the utility function's parameter m so that the average simulated risk-free rate in our RE economy equals the average 4-6 month commercial paper rate over the period 1871-1997. The details are in the Appendix.

4.1. Predictability of Excess Returns

A finding in the empirical macroeconomics literature is that stock returns can be predicted using past information. Several researchers have regressed excess returns, over various horizons, onto past ratios of dividends or earnings to prices. They have found that the estimates of the coefficients corresponding to these ratios are significantly different from zero at the usual significance levels (see, for instance Fama and French (1988)). In our simulations we consider the following regression equation

$$\rho_{t+1} = \alpha_0 + \alpha_1 \frac{y_t}{p_t^r} + u_{t+1} \tag{4.1}$$

where excess returns (ρ_{t+1}) and the safe return $(1 + r_t^s)$ are given by

$$\rho_{t+1} \equiv \frac{y_{t+1} - [1 + r_t^s]p_t^r}{p_t^r}, \qquad 1 + r_t^s = \frac{1}{p_t^s} \tag{4.2}$$

Using the simulated dividend sequences we construct 100 sequences of prices and excess returns. We then regress excess returns onto the past ratio of dividends

to prices. The regressions that correspond to rational expectations are based on price formulas (2.9), while those that correspond to $\mathcal{H}_\infty$ are based on (2.12).

Table 1 reports the simulation results. We find that using the RE price sequences the estimated coefficient $\hat{\alpha}_1$ is different from zero, at the 1% significance level, in only 6% of the regressions. The average $t-$statistic across the 100 regressions is only 0.77, while the $t-$statistic of the regression that uses actual US data is 3.71.

The regression results using $\mathcal{H}_\infty$ forecasts with a disturbance attenuation parameter $\gamma = \infty$ are similar to those reported for the rational expectations case. This should come as no surprise, given the equivalence of the RE and the $\mathcal{H}_\infty$ forecasting formulae when $\gamma = \infty$. Note, however, that when γ is set at a sufficiently low level (say $\gamma = 0.5$), the proportion of cases in which $\hat{\alpha}_1$ is significant increases to 80%, and the average R^2 is 40%. Recall that γ cannot be set arbitrarily small. Proposition 2.2 states that its infimum is given by $\underline{\gamma}_{t+1} := [Z_{t+1} + \sigma_v^2]\beta = 0.32$. If $\gamma \leq \underline{\gamma}_{t+1}$, an equilibrium does not exist.

Table 1
Predictability of Excess Returns

Regression: $\rho_{t+1} = \alpha_0 + \alpha_1 \frac{y_t}{p_t} + u_{t+1}$

	$\gamma = 0.33$	$\gamma = 0.5$	$\gamma = 0.6$	$\gamma = 1.7$	$\gamma = \infty$	RE	U.S. data
% of cases $\hat{\alpha}_1$ is significant at the 1% level	98	80	43	22	6	6	Yes
average $t(\hat{\alpha}_1)$	59.48	32.71	3.29	0.89	0.77	0.77	3.71
average R^2	0.78	0.40	0.15	0.05	0.04	0.04	0.10

Note: All variables are in levels. Excess returns are defined as $\rho_{t+1} \equiv \left(\frac{y_{t+1}-(1+r_t^s)p_t}{p_t}\right)$ where $r_t^s = \frac{1}{p_t^s} - 1$. We generate 100 dividend series of 100 periods each. The parameters of the data generating process (a,d,σ_w,σ_v) are set equal to their maximum likelihood estimate given in the Appendix. For each γ in the table, we construct sequences of prices and excess returns setting $\beta = 0.95$ and $m = 0.944877$. This value of m makes the RE risk free rate equal to the average risk free rate in the US data (2.4%). We then perform 100 regressions. In each regression, we discarded the first 20 observations to minimize the effect of initial values. The first row in the table reports the percentage of cases in which the estimated α_1 is different from 0 at the 1% significance level. The second row reports the average t statistic across regressions. The third row reports the average R^2. For the RE case, we follow the same procedure using the RE price formulae. In the last column, we report the results of a regression that uses US data for the period $1871 - 1997$. Data sources are in the Appendix.

4.2. Excess Volatility

LeRoy and Porter (1981) and Shiller (1981) have shown that stock prices are too volatile to be generated by rational expectations forecasts with a constant discount factor. They compare the variance of actual prices with the variance of the price sequences that would prevail if the discount factor were constant and agents had perfect knowledge of future dividends (denote it by $\tilde{p}_t$). Since a rational expectations forecast cannot have a greater variance than the quantity it is forecasting, $var(p_t^{re}) \leq var(\tilde{p}_t)$.

Several researchers have found that the variance of actual stock prices in the U.S. has been several times greater than $var(\tilde{p}_t)$. This violation of variance bounds, is not a rejection of the hypothesis of efficient markets, as it can be caused by the existence of time varying discount factors, learning, etc. The simulations of this subsection show that, if the disturbance attenuation parameter γ is low enough, $\mathcal{H}_\infty$ prices violate the variance bounds (i.e., $var(p_t^\gamma) > var(\tilde{p}_t)$).

Table 2 reports the results of our simulations. The data generating process is the same as the one in the previous subsection. Again, we generate 100 dividend sequences. Then, for each value of γ shown in the table, we construct 100 price

53

sequences using formula (2.12), and compute the mean variance $\sigma^2(p^\gamma)$. We also compute 100 sequences of 'Shiller prices,' under the assumption that agents have perfect knowledge of next period's dividends, by substituting y_{t+1} for $\mathcal{F}_t(y_{t+1})$ in (2.12). The mean variance of these price sequences is denoted by $\sigma^2(\tilde{p})$ in the table. For the rational expectations column we use price formula (2.9).

As expected, under rational expectations, the variance ratio $\left(\frac{\sigma^2(p^{re})}{\sigma^2(\tilde{p}_t)} \right)$ is much smaller than the ratio generated by US data: 1.15 versus 3.26. Furthermore, in all simulations $\sigma^2(p^{re}) < \sigma^2(\tilde{p}_t)$ as found by Shiller (1981) and several other researchers. Similar results obtain when we use the $\mathcal{H}_\infty$ pricing formula with $\gamma = \infty$ (i.e., agents do not care about robustness). However, as γ is reduced, the ratio $\frac{\sigma^2(p^\gamma)}{\sigma^2(\tilde{p}_t)}$ increases monotonically. For $\gamma = 0.5$, the average variance ratio equals 2.24. Furthermore, for $\gamma = 0.33$ this ratio is 8.87 and the proportion of cases in which $\sigma^2(p^\gamma) \geq 4\sigma^2(\tilde{p}_t)$ equals 90%.

Table 2
Excess Volatility of Stock Prices

	$\gamma = 0.33$	$\gamma = 0.36$	$\gamma = 0.5$	$\gamma = 0.8$	$\gamma = \infty$	RE	U.S. data
$\sigma^2(p)/\sigma^2(\widetilde{p})$	8.87	3.32	2.24	1.47	1.15	1.15	3.26
% of cases $\sigma^2(p) > \sigma^2(\widetilde{p})$	98	76	52	28	5	5	N/A
% of cases $\sigma^2(p) > 2\sigma^2(\widetilde{p})$	96	39	14	5	4	4	N/A
% of cases $\sigma^2(p) > 3\sigma^2(\widetilde{p})$	93	23	7	4	4	4	N/A
% of cases $\sigma^2(p) > 4\sigma^2(\widetilde{p})$	90	15	5	4	3	3	N/A
% of cases $\sigma^2(p) > 5\sigma^2(\widetilde{p})$	89	12	4	3	0	0	N/A

Note: For each γ shown in the Table we construct 100 price sequences following the same proce-dure as in Table 1. For each price sequence, we compute the variance as $\sigma_i^2(p^\gamma) = \frac{\sum_{t=1}^{100}(p_t^\gamma - \bar{p}^\gamma)^2}{99}$. For each γ, we report the mean variance: $\sigma^2(p^\gamma) = \frac{\sum_{i=1}^{100}\sigma_i^2(p^\gamma)}{100}$. We compute the variance of "Shiller's prices" $\sigma^2(\widetilde{p}^\gamma)$ using the same procedure, but substituting y_{t+1} for $\mathcal{F}_t(y_{t+1})$ in the price formula. For the RE case, we follow the same procedure using the RE price formula. The first row reports the ratio of the mean variance of prices to the mean variance of "Shiller's prices". The second row reports the percentage of cases that the mean variance of prices ex-ceed 2 times the mean variance of "Shiller's prices", and so on. The last column is constructed using actual US data; $\sigma^2(p^{actual})$ is the variance of US stock prices for the period $1871 - 1997$; $\sigma^2(\widetilde{p})$ is the variance of the series $\widetilde{p}_t = \beta \frac{m - y_{t+1}^{actual}}{m - y_t^{actual}} y_{t+1}^{actual}$. We discard the first 20 observations to minimize the effect of initial values.

4.3. The Equity Premium Puzzle

Historically, the average excess return on US stocks over short-term debt has been around 6% per year. The 'equity premium puzzle' of Mehra and Prescott (1985) is that, in a standard time-additive power utility model, an equity premium of 6% is consistent with the highly smooth U.S. consumption sequences only if the coefficient of relative risk aversion is unreasonably large. The associated 'risk-free rate puzzle' is that, in the standard model, the risk-free rate has to be very large in order to generate observed equity premia.

Table 3 reports the results of our simulations. The data generating process is the same as the one in previous subsections. Again, we generate 100 dividend sequences. For each value of γ shown in Table 3, we construct 100 price sequences using formula (2.12). Then, we compute 100 sequences of risk free rates $r_t^s = \frac{1}{p_t^s} - 1$, of rates of return on equity $r_t = \frac{y_{t+1} - p_t^r}{p_t^r}$, and of equity premia $(r_t - r_t^s)$. In the rational expectations case we follow the same procedure using (2.9) instead of (2.12).

In Table 3 we report, for each γ, the average equity premium $(\bar{r} - \bar{r}^s)$ and the average risk free rate $\bar{r}$. As we explained before, we chose m so that, in the rational

expectations case, the simulated average risk-free interest rate equals the average interest rate on commercial paper over the period 1871-1997. As expected, in the RE case the equity premium is much smaller than the one observed in the data: 1.17% versus 5.75%. In the $\mathcal{H}_\infty$ case the average equity premium is also very small if agents do not care about robustness (i.e., $\gamma = \infty$). However, the equity premium increases as γ falls. For the case $\gamma = 0.5$ it attains a value of 5.9%, while maintaining a low risk-free interest rate.

Table 3

Equity Premium and Risk Free Interest Rate

	$\gamma = 0.33$	$\gamma = 0.5$	$\gamma = 0.8$	$\gamma = 1.7$	$\gamma = \infty$	RE	U.S.data
$(r - r^s) * 100\%$	6.17	5.90	3.29	2.83	1.17	1.17	5.75
$r^s * 100\%$	0.16	1.49	2.06	2.34	2.40	2.40	2.40

Note: For each γ shown in the table we construct 100 price sequences using the same procedure as in Table 1. Then, for each value of γ shown in the table, we construct 100 sequences of risk free rates $r_t^s = (\frac{1}{p_t^s} - 1)$ and 100 sequences of rates of return on equity $(\frac{y_{t+1} - p_t}{p_t})$. For each γ, we report in the first row the average equity premium $(\bar{r} - \bar{r}^s) * 100\%$. For the RE case, we follow the same procedure using the RE price formulae. For the last column, we use actual US data for the period 1871-1997, and we use the following formula: $r_t^{actual} = (\frac{P_{t+1} + D_t - P_t}{P_t} - \frac{PC_{t+1} - PC_t}{PC_t})$ where PC_t is the consumption deflator series and r_t^s is the real interest rate on $4 - 6$ month commercial paper.

4.4. Tracking the US Stock Market

In this subsection we feed the RE and $\mathcal{H}_\infty$ price formulae with historical US dividend data, and analyze how well they track the actual US stock market over the period 1871-1996.

The solid line in Figure 2 plots the actual US stock price. The dotted line corresponds to the price sequence that would have been observed if prices were obtained using the $\mathcal{H}_\infty$ formula (2.12) setting $\gamma = 0.33$. The broken line corresponds to the rational expectations price formula (2.9). We construct the RE and $\mathcal{H}_\infty$ price sequences by feeding formulas (2.9) and (2.12), respectively, with *past observed dividends in the U.S. stock market*. We set the parameters $(a, d, \sigma_w, \sigma_v, \beta, m)$ equal to the values they take in Tables 1-3.

Panel (a) in Figure 2 plots the price sequences for the period 1871-1996, while panel (b) plots prices for the sub-period 1975-1996. As can be seen, the $\mathcal{H}_\infty$ price sequence tracks the actual S&P500 index better than the RE price sequence. Figure 3 makes this point clear. It plots the mean square difference between the S&P 500 index and $\mathcal{H}_\infty$ prices for different values of $\gamma > \underline{\gamma} = 0.32$ over the period 1871-1996. The MSE is minimized at $\gamma = 0.33$, which is the value used in Figure

2.

Figure 2: Tracking the US Stock Market
### (a) 1872-1996
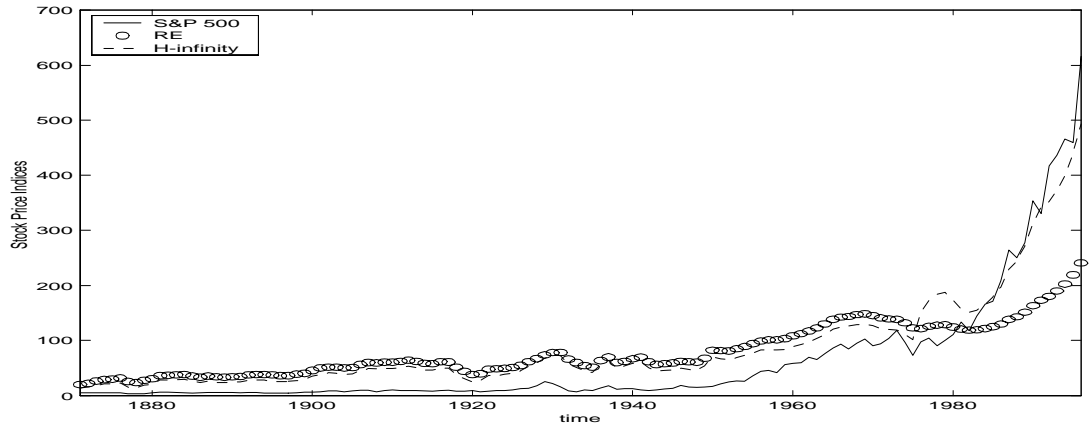
(b) 1975-1996
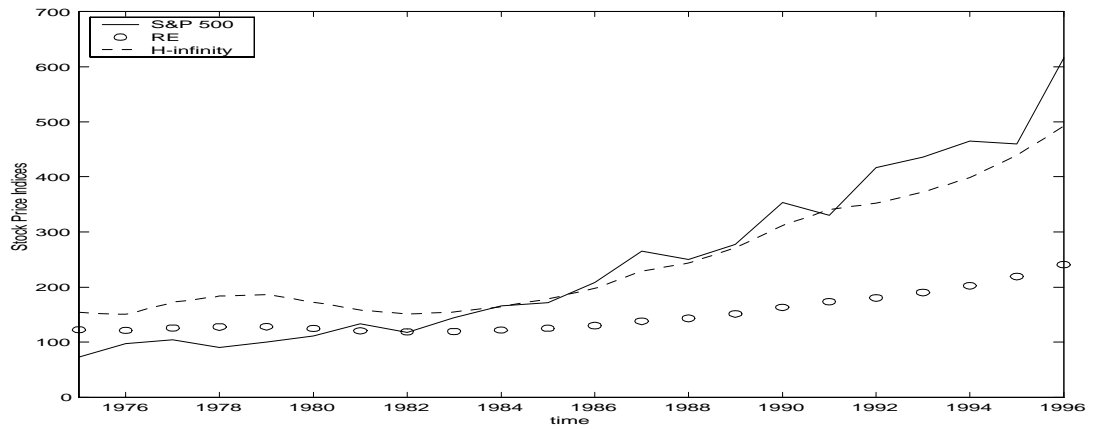
Note: The $\mathcal{H}^{\infty}$ prices and RE prices are generated using actual real US dividends. The initial point of the forecast is 1872. We set the initial simulated price for RE and $\mathcal{H}^{\infty}$ equal to the actual stock price in 1871. The parameters used in producing the forecasts are the same we used in Table 1, 2 and 3. See the Appendix for details.

Figure 3: How Well do $\mathcal{H}^\infty$ Prices Track the US Stock Market$(1871-1996)$?
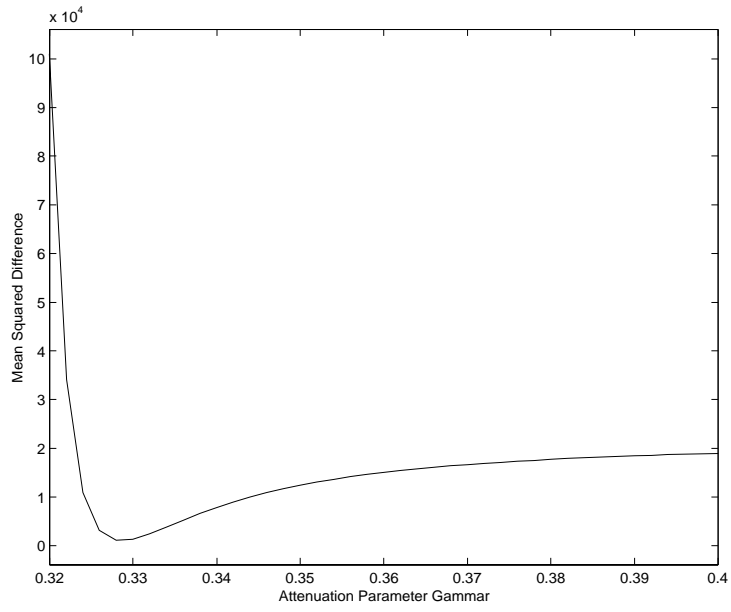
Note: For each value of γ, we produce $\mathcal{H}^\infty$ prices for the period $1871-1996$ with actual US real dividends using the same methodology as in the previous Figure. Then we plot, for each value of γ, the mean square difference between the actual S&P 500 index and $\mathcal{H}^\infty$ prices: $\frac{\sum_{t=1871}^{1996}(p_t^{actual}-p_t^\gamma)^2}{126}$ for each γ.

4.5. Sensitivity of Forecasts to News and the Behavioral Literature

The ability of $\mathcal{H}_\infty$ prices to generate anomalies exhibited by U.S. macroeconomic data reflects the fact that, in our exchange economy, $\mathcal{H}_\infty$ forecasts are more sensitive to dividend news than RE forecasts. Furthermore, this sensitivity is greater, the greater is the degree of robustness (i.e., the smaller is the disturbance attenuation parameter γ).

The link between sensitivity to news and the degree of robustness can be appreciated in the impulse response functions of Figure 4. This Figure depicts, for different values of γ, the evolution of the one-period-ahead dividend forecast $\mathcal{F}_{t-1}(y_t)$ in response to a persistent shock at time 0. As can be seen, at all time horizons the forecasts' response corresponding to $\gamma = 0.7$ are greater than those for $\gamma = 0.9$, which in turn are greater than the RE forecasts (i.e., $\gamma = \infty$).

In what follows we would like to link the excess sensitivity of $\mathcal{H}_\infty$ forecasts to the behavioral literature. This literature has rationalized a number of anomalies by postulating that agents are boundedly rational. That is, although agents optimize an objective function and filter information using Bayes law, they *misperceive* the true model of the economy. An appealing feature of this literature is

that it considers misperceptions that reflect regularities established by psychological experiments (e. g. Kahneman and Tversky (1972).[17]

For instance, in order to rationalize excess volatility and long-term reversals, researchers have assumed that agents misperceive shocks as being more persistent than what they actually are. As a result, their forecasts are more sensitive to news than rational expectations' forecasts (i.e., they overreact). This behavioral assumption captures psychological regularities such as *overconfidence* and *associativeness*.[18]

The point we want to make is that, in a simple asset pricing model, excess sensitivity to news can result from either: (a)*misperception* of the duration of shocks in a behavioral setup, or (b)from a desire for *robustness* in an $\mathcal{H}_\infty$ setup. To illustrate this point consider the model of Daniel, Hirshleifer and Subrahmanyam (1998, henceforth DHS). In DHS agents are *overconfident* about the precision of their private signals. After observing a signal agents estimate, using Bayes law, the unobservable state that will determine the payoff of the asset at terminal time

---

[17]See Barberis, et.al. (1998), Cecchetti, et. al. (1997), Daniel, et al. (1998), and Mullainathan (1999).

[18]Overconfident agents view themselves as more able to value assets than what they actually are, so that they underestimate their forecast error variane. Associativeness implies that a signal affects beliefs not only through the information it conveys, but also through the memories it evokes.

(call it x). The price at time t is the estimate of x at time t.[19]

To see the link between our setup and that of DHS note that in their setup agents observe a sequence of signals and estimate the unobservable state, which is a constant. Using our notation this is equivalent to $y_t = x + \sigma_v v_t$ (see (2.3)), where $x \sim N(0, \sigma_w)$ and $v_t \sim N(0, 1)$. It follows that their overconfidence assumption is equivalent to a situation in which agents perceive the noise-to-signal ratio $(\tilde{n} = \frac{\tilde{\sigma}_v^2}{\tilde{\sigma}_w^2})$ to be lower than what it actually is (i.e., $\tilde{n} < n := \frac{\sigma_v^2}{\sigma_w^2}$). Consider now an economy like the one of subsection 2.1, but where agents are overconfident (i.e., $\tilde{n} < n$). It then follows from the RE forecasting formulae (2.8) that the gain will be greater (i.e., $h(\tilde{n}) > h(n)$), and there will be excess sensitivity to news.[20] To complete the analogy, recall that the $\mathcal{H}_\infty$ dividend forecasts are decreasing in the disturbance attenuation parameter γ, and that they are equal to the RE forecast if $\gamma \to \infty$.

Consequently, we can suggest that greater sensitivity to news reflects either greater robustness in an $\mathcal{H}_\infty$ setup, or missperception in a behavioral setup. A

---

[19]Overreaction to news occurs in other papers for similar reasons. In Mullainathan (1999) current events evoke memories. This reduces the noise-to-signal ratio perceived by agents. In Barberis, et. al. (1998) overconfidence comes about because, after observing several signals pointing in the same direction, an agent beleives that earnings follow a steady trend, while they actually follow a random walk.

[20]For concreteness, consider the forecast of y_2 and assume that $\tilde{n} < n$. In an RE economy an agent forms his one-period ahead dividend forecast (i.e., $\widehat{y}_2$) using the filter (2.8). Since the expected value of v_2 is zero, his forecast is: $\widehat{y}_2(\tilde{n}) = \hat{x}_2(\tilde{n}) = \frac{a_1}{1+\tilde{n}} y_1$. Clearly, since the perceived noise-to-signal ratio is smaller than the true one ($\tilde{n} < n$), it follows that $\widehat{y}_2(\tilde{n}) > \widehat{y}_2(n)$.

key difference is that in an $\mathcal{H}_\infty$ setup agents use the same 'nominal' model as RE agents, while in a behavioral setup agents use a different nominal model (i.e., $\tilde{n} \neq n$).

We would like to point out that the positive relationship between the degree of robustness and sensitivity to news is not an inherent property of $\mathcal{H}_\infty$ forecasts. It should be clear, from the derivation in Section 3, that $\mathcal{H}_\infty$ forecasting formulas are jointly determined with the portfolio strategy, and are dependent on the agent's preferences and budget constraints. Thus, one could obtain different response patterns by changing the economic environment in suitable ways.

Finally, note that there are several other asset pricing anomalies that we have not considered. Our purpose has been simply to illustrate how the robust approach can help explain some anomalies. In this respect, a promising direction for future research is to consider environments in which the $\mathcal{H}_\infty$ approach generates richer price dynamics.[21]

---

[21]Daniel, et. al. (1998) present an extensive survey of asset pricing anomalies. A prominent anomaly, present mainly in cross-sections, is the short-term momentum and long-term reversals of returns (i.e., the underreaction-overreaction anomaly). That is, a positive autocorrelation of returns over horizons shorter than 12 months, and a negative autocorrelation over longer horizons. DHS generate this pattern by assuming that overconfidence increases over time. This generates a response of forecast to news that increases over time (i.e., undereaction relative to the future). In our setup this corresponds to a reduction of the disturbance attenuation parameter over time. Barberis et. al. (1998) assume that after observing several signals that point in the same direction that agents switch their perception about the data generating process from mean reverting to a steady trend. However, the true process is a random walk.

Figure 4: Impulse Response of Dividend Forecasts
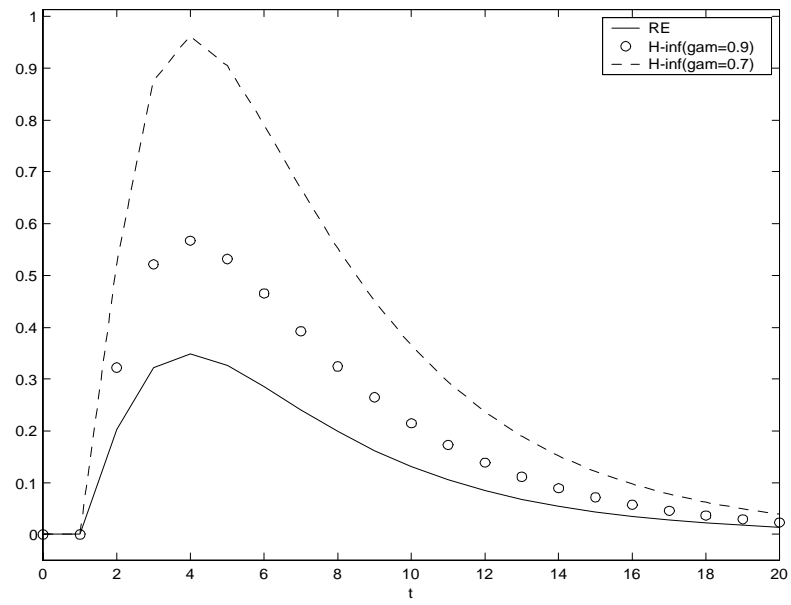
Note: Each curve shows the change in the one-period-ahead dividend forecast in response to a persistent shock to dividends at date 0 (i.e. $y_1 = \sigma_w w_0 = 1$).

5. Conclusions

The signal-extraction problem cuts across several fields in economics. Typically, an agent has available a sequence of observations and tries to untangle the 'persistent' and 'transitory' components (e.g., Muth (1960)), or he might want to separate the 'fundamental' and 'noisy' components (e.g. Lucas (1973)). This problem has been tackled by postulating that observations are generated by sequences of disturbances that follow a well specified stochastic process and that agents either know this process or can learn it. Furthermore, if disturbances are i.i.d. Normal with known variances, one applies the familiar Bayesian updating formula that weights innovations in accordance to the variances of persistent and transitory disturbances.

In macroeconomics there are several instances where agents seem to form expectations in ways that are not consistent with the predictions of models that combine Bayes law with standard time-additive preferences. This has lead some researchers to consider boundedly rational agents that act *as if* they confuse the transitory and persistent components of a given sequence of observations. These agents form their expectations applying Bayes law but using the wrong 'nominal'

model. That is, they use the wrong ratio of variances when applying the familiar Bayesian updating formula referred to above. This small deviation from full rationality has proven to be quite useful in rationalizing several anomalies.

In this paper we have shown that forecasting formulas that feature excess sensitivity to news can be obtained in a setup where agents have the same nominal model as rational expectations (RE) agents, but fear misspecification. The straightjacket of optimal estimation imposed by RE models is not loosened; it is simply replaced by another straightjacket. Robust agents estimate unobservable variables using all available information. However, to derive the optimal strategies, instead of working with probability distributions one works with a dynamic game between the agent and nature. This is because the assumption that uncertainty can be modelled in terms of probability distributions is eliminated. Instead, $\mathcal{H}_\infty$-control models uncertainty as totally unknown sequences of disturbances. Of course, not anything goes, as these sequences must be square summable.

We have combined this description of uncertainty with a simple Lucas type exchange economy. Forecasts of dividends in this economy exhibit excess sensitivity to news, relative to RE forecasts. As a result, if the degree of robustness is high, equilibrium prices are more volatile than dividends, returns are predictable, and the equity premium is large, while keeping a low risk-free interest rate.

There are still many open questions. What determines the degree of robustness chosen by agents? How do agents learn through time? What are sensible ways to extend the model in order to rationalize other anomalies? This paper is just one example of the potential of robust control to address economic problems.

6. Appendix

Proof of Theorem 3.1. An equilibrium exists if and only if there exists a price vector such that Problem $\mathcal{H}$ has a solution, and asset and goods markets clear. We have shown that Problem $\mathcal{H}$ has a solution if and only if the value function associated with (3.2) is bounded. In what follows we will determine the conditions under which this is true at equilibrium prices. Using the definition of Ω_s in (3.4) we can rewrite (3.2) as

$$\Gamma(p) = \inf_{\chi \in \mathcal{Q}} \sup_{\omega \in \Omega_t} \left\{ L(q, y_{t+1}; p) + \sum_{j=0}^{t+1} g_j(\omega_j) \right\} \leq 0 \qquad (6.1)$$

Let $\hat{p}(y)$ be an equilibrium price vector given dividend history $\{y_j\}_{j=1}^t := y$. Since the supplies of risky and safe assets are 0 and 1, respectively, it follows from (6.1)

70

that at equilibrium prices Problem $\mathcal{H}$ has a solution only if

$$\Gamma(\hat{p}(y)) = \sup_{\omega \in \Omega_t} \left\{ L(1, 0, y_{t+1}; \hat{p}(y)) + \sum_{j=0}^{t+1} g_j(\omega_j) \right\} < \infty \qquad (6.2)$$

The dynamic system (2.3) followed by dividends implies that past disturbances $\{\omega_j\}_{j=0}^{t}$ affect y_{t+1} only through x_{t+1}. Thus, conditional on $x_{t+1} = x$: (i)$L(q, y_{t+1}, p)+$ $g_{t+1}(\omega_{t+1})$ is independent of $\{\omega_j\}_{j=0}^{t}$, and (ii) $\sum_{j=0}^{t} g_j(\omega_j)$ is independent of (q, ω_{t+1}). Thus, if $\Gamma(\hat{p}(y))$ is finite, we can carry out the following expansion of (6.2):

$$\begin{aligned}
\Gamma(\hat{p}(y)) &= \max_x \max_{(\tilde{v} \in \mathcal{V}_x)} \max_{(\omega^t \in \Omega_t | x_{t+1}=x)} \left\{ L(1, 0, y_{t+1}; \hat{p}) + g_{t+1}(\omega_{t+1}) + \sum_{j=0}^{t} g_j(\omega_j) \right\} \\
&= \max_x \left\{ \max_{\tilde{v} \in \mathcal{V}_x} \left(L(1, 0, y_{t+1}; \hat{p}) + g_{t+1}(\omega_{t+1}) \right) + \max_{\omega^t \in \Omega_t | x_{t+1}=x} \sum_{j=0}^{t} g_j(\omega_j) \right\} \\
&= \max_x \{ V_{t+1}(x) + W_{t+1}(x) \} := \max_x \Lambda(x) \qquad (6.3)
\end{aligned}$$

The functions $W_{t+1}(x)$ and $V_{t+1}(x)$ are defined in (3.14) and (3.19), respectively. Recall that the set $\mathcal{V}_x$ of admissible strategies for v_{t+1} consists of functions of (x, q, p), see (3.6). The set Ω_t, defined in (3.4), consists of all disturbance sequences that belong to the set Ω and that are consistent with dividend history $\{y_j\}_{j=1}^{t}$.

The 'only if' in part (i) of Theorem 3.1 follows from the fact that $\Gamma(\hat{p}) < \infty$ if and only if $V_{t+1}(x)$ and $W_{t+1}(x)$ are bounded, and $\Lambda(x)$ is concave. Hence, if one

71

of these conditions does not hold, there exists no price vector such that Problem $\mathcal{H}$ has a solution with $q = (1, 0)$. Thus, an equilibrium does not exist.

To prove part (ii) and the 'if' in part (i), it suffices to show that there is a unique x^* and a unique Markov price vector $(p^*(x^*))$ that induce $q^* = (1, 0)$. Since $\Gamma(p^*)$ is finite (because part i of Theorem 3.1 holds), we can carry out the following expansion of (6.1):

$$\Gamma(p^*) = \max_x \min_{(\tilde{\chi} \in \mathcal{Q}_x)} \max_{(\tilde{v} \in \mathcal{V}_x)} \max_{(\omega^t \in \Omega_t | x_{t+1} = x)} \left\{ L(q, y_{t+1}; p) + g_{t+1}(\omega_{t+1}) + \sum_{j=0}^{t} g_j(\omega_j) \right\}$$
(6.4)

Fix $x_{t+1} = x$ and recall that $y_{t+1} = x_{t+1} + \sigma_{vt+1} v_{t+1}$ (by (2.3)). Conditional on $x_{t+1} = x$, the problem $\min_{\tilde{\chi} \in \mathcal{Q}_x} \max_{\tilde{v} \in \mathcal{V}_x} \{ L(q, y_{t+1}, p) + g_{t+1}(\omega_{t+1}) \}$ can be analyzed as a standard zero-sum game in which the state is perfectly observed. To determine whether a solution exits recall that $w_{t+1} = 0$, and note that the kernel is an indefinite quadratic form in v_{t+1}, and that it is strictly convex in q. It follows that for a given pair (x, p) there is a solution if and only if $L(q, y_{t+1}(x), p) + g_{t+1}(0, v_{t+1})$ is concave in v_{t+1}. This is the case if and only if $\gamma_{t+1}^2 > \beta q_t^{r2} \sigma_{vt+1}^2$. If this condition holds, the solution is unique and given by (3.16)-(3.18).

To derive the $\mathcal{H}_\infty$ estimate of the terminal state (x^*) denote by $p^*(x)$ the price vector that equilibrates asset markets, conditional on $x_{t+1} = x$. This price

vector is uniquely determined by $\tilde{\chi}_t^{r*}(x, p^*(x)) = 1$ and $\tilde{\chi}_t^{s*}(x, p^*(x)) = 0$, where asset demands are given by (3.17) and (3.18). Substitute $q^*(p^*(x)) = (0, 1)$ and $\tilde{v}^*(x, p^*(x), (1, 0))$ in (6.4) and use (3.14) to obtain $\Gamma(p^*(x)) = \max_x \{V_{t+1}(x) + W_{t+1}(x)\} := \max_x \Lambda(x)$. Note that $\tilde{v}^*(x, p^*(x), (1, 0))$ is bounded and unique if and only if $\gamma_{t+1}^2 > \beta \sigma_{vt+1}^2$. This condition is implied by part (i) of Theorem 3.1. To see this note that $\Lambda(x)$ is concave if and only if $\gamma_{t+1}^2 \geq \underline{\gamma}_{t+1}^2 = \beta[\sigma_{vt+1}^2 + Z_{t+1}]$, and $Z_{t+1} > 0$ (see the proof of Lemma 3.2). The next step is to note that (6.4) implies that $x^* \in \arg\max \Lambda(x)$.

Lastly, if $\Lambda(x)$ is strictly concave, there is a unique MPE because there is a unique pair $(x^*, p^*(x^*))$ that clears assets and goods markets (i.e., $q^* = (1, 0)$ and $c_t^* = y_t$). The Markov equilibrium price vector $p^*(x^*)$ is given by (2.12), and the associated *forecasted dividends* are $y_{t+1}^* = x^* + d_{t+1} + \sigma_{vt+1} v_{t+1}^*(x^*, p^*(x^*), (1, 0))$. $\square$

Proof of Lemma (3.2). Since $W_{s+1}(x)$ is the supremum of a quadratic function under an affine constraint, it must be quadratic in x if a maximum exists. That is, it must be of the form $W_{s+1}(x) = -K_{s+1}[x - \check{x}_{s+1}]^2 + l_{s+1}$, where K_{s+1}, $\check{x}_{s+1}$ and l_{s+1} are undetermined coefficients. To derive $W_1(x)$ note that there is no observation y_0. Since the state satisfies $x_{t+1} = a_t x_t + \sigma_{wt} w_t$, and $x_0 = 0$ (by

(2.3)), we have that $x_1 = \sigma_{w0}w_0$. Thus, for $s = 0$ Problem (3.12) becomes

$$W_1(x) = \max_{\omega_0} -\gamma_{t+1}^2 \left[w_0^2 + v_0^2\right] \qquad \text{s.t.} \quad x = \sigma_{w0}w_0$$

The solution to this problem is $w_0^* = x\sigma_{w0}^{-1}$, and $v_0^* = 0$. Therefore, $W_1(x) = -x^2\gamma_{t+1}^2\sigma_{w0}^{-2}$. This implies that $\check{x}_1 = 0$ and $K_1 = \gamma_{t+1}^2\sigma_{w0}^{-2}$.

Replacing $W_{s+1}(x) = -K_{s+1}[x - \check{x}_{s+1}]^2 + l_{s+1}$ in (3.13) and regrouping terms we have that, for $s \geq 1$, $W_{s+1}(x)$ is the solution to the problem

$$W_{s+1}(x) = \max_{\xi,\omega} \left\{-P_s\xi^2 + 2[y_s\gamma_{t+1}^2\sigma_{vs}^{-2} + K_s\check{x}_s]\xi - \gamma_{t+1}^2 w^2 + l_s\right\}$$

$$\text{subject to } a_s\xi + \sigma_{ws}w = x \tag{6.5}$$

where the independent terms are included in l_s, and P_s is given by (3.15). If $P_s \neq 0$, the maximum is finite only if the bordered Hessian of the associated Lagrangian is non-negative: $a_s^2\gamma_{t+1}^2 + \sigma_{ws}^2 P_s \geq 0$. In this case $W_{s+1}(x)$ is given by Lemma (3.2).

Lastly, since $K_1 > 0$, we have that $P_1 = \gamma_{t+1}^2\sigma_{v1}^{-2} + K_1 > 0$. This in turn implies that $K_2 = \left[\frac{a_1^2}{P_1} + \gamma_{t+1}^{-2}\sigma_{w1}^2\right]^{-1} > 0$. Thus, $P_s > 0$ and $K_{s+1} > 0$ for all $s \in [1, t]$. Since $P_s > 0$ for all $s \leq t$, the bordered Hessian of the Lagrangian associated with (6.5) (i.e., $a_s^2\gamma_{t+1}^2 + \sigma_{ws}^2 P_s$) is positive for all $s \leq t$. Hence, Problem (6.5) has a

74

bounded solution for all $s \leq t.\square$

Simulations for Figures 1, 2 and 3.

To construct Figure 1a we generate artificial dividends' observations using (2.3). The disturbances $\{v_j\}_{j=0}^{t+1}$ and $\{w_j\}_{j=0}^{t+1}$ are i.i.d. $N(0,1)$, and the parameters are set equal to: $a = 0.0607$, $d = 0.483$, $\sigma_v = 0.0097$, and $\tilde{\sigma}_w = $ .... [22] We select 45 values of Δ^i in the interval $[-0.0698, 0.0687]$, and for each value of $\sigma_w^{true} = \{\tilde{\sigma}_w + \Delta^1, ..., \tilde{\sigma}_w + \Delta^{45}\}$ we generate 100 dividend sequences, each consisting of 100 observations. For each value of Δ^i we then construct RE and $\mathcal{H}_\infty$ one-period-ahead dividend forecasts (i.e., $E_j(y_{j+1})$ and $\mathcal{F}_j(y_{j+1})$) using formulas (2.8) and (2.10) with $\Delta_i = 0$. The other parameters are set equal to the values they take in data generating process, while $m = 0.78$ and $\gamma = 0.3$. The curve corresponding to the $\mathcal{H}_\infty$ forecasts in Figure 1a is constructed in the following way. For each value of Δ^{true} we compute 100 mean square errors of the $\mathcal{H}_\infty$ forecasts using the formula: $MSE_h(\tilde{\sigma}_w + \Delta^{true}) = \sum_{j=1}^{99}(y_{j+1} - \mathcal{F}_j(y_{j+1}))^2/100$. Then we compute the average MSE as $\overline{MSE}(\tilde{\sigma}_w + \Delta^{true}) = \sum_{h=1}^{100} MSE_h(\tilde{\sigma}_w + \Delta^{true})/100$. Figure 1a plots $\overline{MSE}(\tilde{\sigma}_w + \Delta^{true})$ for different values of Δ^{true}. The curve corresponding to the RE forecasts is constructed in the same way using (2.8).

---

[22]We draw disturbances from a Normal distribution to ensure that RE forecasting formulas (2.8) correspond with the rational expectations forecasts in the absence of missperception.

Figure 1b is constructed in the same way as Figure 1a. The only difference is that $\sigma_w^{true} = \tilde{\sigma}_w$, and $\sigma_v^{true} = \tilde{\sigma}_v + \Delta^{true}$, where $\tilde{\sigma}_v = 0.0097$ and Δ_i takes values on $[-0.0093, 0.0095]$. In Figure 1c $a^{true} = \tilde{a} + \Delta^{true}$, where $\tilde{a} = 0.0607$ and Δ_i takes values on $[-0.2187, 0.15]$.

Figure 2 is constructed in three steps. First, we detrend real annual dividends per share using a fourth degree polynomial. Second, we compute $\mathcal{H}_\infty$ and RE price sequences by feeding price equations (2.9) and (2.12) with detrended dividends. We set the disturbance attenuation parameter $\gamma = 0.33$. The other parameters are set equal to their values in Tables 1, 2 and 3. Lastly, we estimate a time trend for the actual stock prices, and add it to the prices we generated in the second step. To construct Figure 3 we compute $\mathcal{H}_\infty$ prices for different values of γ. Then we compute, for each value of γ, the mean square difference between the actual S&P500 index and the $\mathcal{H}_\infty$ price sequence.

Data

Data for the period 1871-1985 is from Shiller (1989). For the period 1986-1997 we used DataStream and the International Financial Statistics from the IMF. We used the following series:

- The nominal stock price index (P_t) is the annual average Standard and Poor's composite stock price index.

76

- The consumption series corresponds to real consumption of non-durables and services. We denote the consumption deflator by PC_t.

- The real stock price index (p_t^{actual}) is obtained by dividing P_t by PC_t.

- D_t are the nominal annual dividends of the firms in the S&P index.

- y_t^{actual} denotes real detrended annual dividends corresponding to the firms in the S&P index.

- r_t^{actual} denotes the real return on stocks. It is given by $r_t^{actual} = \frac{P_{t+1}+D_t-P_t}{P_t} - \frac{PC_{t+1}-PC_t}{PC_t}$.

- $r_t^{s,actual}$ denotes the real riskless interest rate. It is computed by subtracting the inflation rate from the nominal interest rate on 4-6 months commercial paper (R_t). That is, $r_t^{actual} = R_t - \frac{PC_{t+1}-PC_t}{PC_t}$.

Maximum Likelihood Estimates of Parameters

In the simulations we assume that the parameters $(a, \sigma_v, \sigma_w, d)$ are constant. We obtain the maximum likelihood estimates by assuming that the data generating process is given by (2.3), with $v_t \sim N(0,1)$, $w_t \sim N(0,1)$. We set x_1 equal to the value of real dividends in 1871, which is the initial observation in our data set.

Defining the information set as $I_t = \{y_{t-j}\}_{j=0}^{t-1}$ and $\hat{x}_{t+1} = E[x_{t+1}|I_t]$, we can write the conditional likelihood of y_{t+1} as:

$$\log(f_{y_{t+1|t}}(y_{t+1}|I_t)) = \log\left((2\pi)^{-\frac{T}{2}}[Z_{t+1} + \sigma_v^2]^{-\frac{1}{2}}\right) - \frac{1}{2}\frac{[y_{t+1} - d - x_{t+1}]^2}{Z_{t+1} + \sigma_v^2}$$

$$= \log\left([2\pi]^{-\frac{T}{2}}\left[\frac{a^2}{Z_t^{-1} + \sigma_v^{-2}} + \sigma_w^2 + \sigma_v^2\right]^{-\frac{1}{2}}\right) - \frac{1}{2}\frac{[y_{t+1} - d - a\hat{x}_t - \frac{a}{1+Z_t^{-1}\sigma_v^2}(y_t - d - \hat{x}_t)]^2}{\frac{a^2}{Z_t^{-1}+\sigma_v^{-2}} + \sigma_w^2 + \sigma_v^2}$$

where $\hat{x}_t$ is generated by recursion (2.8). We replace $\{y_j\}_{j=1}^T$ by actual U.S. dividend data for the period 1871-1997 and choose $(a, \sigma_v, \sigma_w, d)$ to maximize $\sum_{t=1}^T \log(f_{y_{t+1|t}}(y_{t+1}|I_t))$, using the MLE function of TSP. The initial guesses for the unknown parameters are: $a = 0.6566$, $\sigma_v^2 = 0.1409$, $\sigma_w^2 = 0.0506$ and $d = 0.557$. TSP maximizes the log likelihood function using the standard gradient method. The maximum likelihood estimates and standard errors are:

$$a = \begin{matrix} 0.70233 \\ (0.0233) \end{matrix}, \quad d = \begin{matrix} 0.50436 \\ (0.0328) \end{matrix}, \quad \sigma_w^2 = \begin{matrix} 0.07888 \\ (0.0151) \end{matrix}, \quad \sigma_v^2 = \begin{matrix} 0.02149 \\ (0.0104) \end{matrix} \tag{6.6}$$

References

Anderson, E., L. Hansen, and T. Sargent, 1999, "Risk and Robustness in Equilibrium," working paper, Stanford University.

Barberis, N., A. Shleifer, and R. Vishny, 1998, "A Model of Investor Sentiment," *Journal of Financial Economics,* 49(3), pp. 307-43.

Barberis, N., M. Huang, and T. Santos, 1999, "Prospect Theory and Asset Prices," mimeo, University of Chicago.

Basar, T. and P. Bernhard, 1991, $\mathcal{H}^{\infty}$-*Optimal Control and Related Minimax Design Problems*, Birkhauser.

Basar, T. and G. Olsder, 1999, *Dynamic Noncooperative Game Theory,* SIAM.

Berger, J., 1985, *Statistical decision Theory and Bayesian Analysis*, Springer-Verlag.

Campbell, J. and J. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy.*

Cecchetti, S., P. Lam, and M. Nelson, 1997, "Asset Pricing with Distorted Beliefs: Are Equity Returns to Good to be True?," mimeo.

Constantinides, G. and D. Duffie, "Asset Pricing with Heterogenous Consumers," *Journal of Political Economy, 104,* 219-240.

Daniel, K., Hirshleifer, D., and A. Subrahmanyam, 1998, "Investor Psychology and Security Market Under- and Overreactions," *The Journal of Finance,* 53(6), pp.1839-1885.

Doyle, J., K. Glover, P. Khargonekar and B. Francis, 1989, "State Space Solutions to Standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control Problems," *IEEE Transactions on Automatic Control,* vol. AC-34, 831-847.

Fama E., and K. French, 1988, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics,* 22, 3-27.

Gourinchas, P. and A. Tornell, 2000, "Exchange Rate Dynamics and Misperception," mimeo UCLA.

Hansen, L., T. Sargent and T. Tallarini, 1999, "Robust Permanent Income and Pricing," *Review of Economic Studies,* 66(4), pp. 873-907.

Hansen, L., and T. Sargent, 1998, "Discounted Robust Linear Decision and Estimation in the Frequency Domain," mimeo.

Heaton J., and D. Lucas, 1996, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy, 104,* *668-712.*

Huber, P. 1981, *Robust Statistics,* John Wiley.

Kwakernaak, H. and R. Sivan, 1972, *Linear Optimal Control Systems*, John Wiley.

Kahneman, D. and A. Tversky, "On the Psychology of Prediction," *Psychological Review,* 80(4), pp.237-251.

LeRoy and Porter, 1981, "The Present Value Relation: Tests Based on Variance Bounds," *Econometrica*, 49, 555-577.

Lucas, R., 1973, "Some International Evidence on Output-Inflation Trade-offs," *The American Economic Review*, Vol. 63, No. 3, pp. 326-334.

Lucas, R., 1978, "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429-1445.

Manehout, P., 1999, "Robust Portfolio Rules and Asset Pricing," mimeo, Harvard University.

Marcet, A., and T. Sargent, 1989, "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," *Journal of Political Economy*, 97(6), pp. 1306-22.

Mehra, R. and E. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-161.

Mullainathan, S., 1998, "A Memory Based Model of Bounded Rationality," mimeo, MIT.

Muth, J., 1960, "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, Vol. 55, No. 290, pp. 299-306.

Onatski, A. and J. Stock, 2000, "Robust Monetary Policy Under Model Uncertainty in a Small Model of the U.S. Economy," NBER working paper No. 7490.

Shiller, R., 1981, "Do Stock Prices Move Too Much to be Justified By Subsequent Changes in Dividends?, *American Economic Review*, 71, 421-436.

Shiller, R., 1989, *Market Volatility*, MIT Press.

Timmerman, A., 1996, "Excess Volatility and Predictability of Stock Prices in Autorregressive Dividend Models with Learning," *Review of Economic Studies, 63(4),* pp. 523-557.

Whitle, P., 1990, *Risk Sensitive Optimal Control*, J. Wiley.

Zames, G., 1981, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," *IEEE Transactions on Automatic Control*, vol AC-26, 301-320.

Zhou, K., J. Doyle and K. Glover, 1996, *Robust and Optimal Control*, Prentice-Hall.