

Asset Return Dynamics and Learning ^{*}

William A. Branch
University of California, Irvine

George W. Evans
University of Oregon

Revised October 01, 2008

Abstract

This paper advocates a theory of expectation formation that incorporates many of the central motivations of behavioral finance theory while retaining much of the discipline of the rational expectations approach. We provide a framework in which agents, in an asset pricing model, underparameterize their forecasting model in a spirit similar to Hong, Stein, and Yu (2005) and Barberis, Shleifer, and Vishny (1998), except that the parameters of the forecasting model, and the choice of predictor, are determined jointly in equilibrium. We show that multiple equilibria can exist even if agents choose only models that maximize (risk-adjusted) expected profits. A real-time learning formulation yields endogenous switching between equilibria. We demonstrate that a real-time learning version of the model, calibrated to U.S. stock data, is capable of reproducing regime-switching returns and volatilities, as recently identified by Guidolin and Timmermann (2007).

JEL Classifications: G12; G14; D82; D83

Key Words: Asset pricing, misspecification, behavioral finance, regime-switching, adaptive learning.

1 Introduction

There is, by now, an established literature that studies financial market anomalies such as excess volatility, Markov switching returns and volatilities, and predictability

^{*}We thank participants at the first annual Learning Week workshop, held at the St. Louis Federal Reserve Bank in July 2006, for helpful comments. Support from National Science Foundation Grant No. SES-0617859 is gratefully acknowledged. George Evans is also Part-time SIRE Professor of Economics and Finance at the University of St. Andrews.

of excess returns. (See Lettau and Ludvigson (2005) for a recent discussion.) One important finding in this literature is evidence of multiple regimes each with distinct return and volatility characteristics. See, for example, Guidolin and Timmermann (2005, 2007a, 2007b, 2008), Ang and Bekaert (2002), Turner, Startz, and Nelson (1989), Bollerslev, Chou, and Kroner (1992), Garcia and Perron (1996), and Perez-Quiros and Timmermann (2000). This paper develops a model of bounded rationality that is able to capture many of the salient features of Markov switching returns.

One popular viewpoint is that empirically observed excess returns cannot be explained by a standard Rational Expectations (RE) model. An explosion of research proposes alternative theoretical foundations for the empirical findings.¹ An offshoot of this literature looks beyond RE and formulates behavioral or boundedly rational channels through which these anomalies might arise (e.g., Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999), Hong, Stein, and Yu (2005), and Lansing (2006)). Bounded rationality, of course, is not only of interest to financial economists. In macroeconomics there is a broad literature that replaces full rationality with agents who behave as econometricians; that is, by agents who estimate and select their models in real-time. (See, for example, Marcet and Sargent (1989), Evans and Honkapohja (2001), and Sargent (1999).)

While similar in spirit, these two approaches differ in the degree to which agents' expectations differ from rational expectations. For example, in Marcet and Sargent (1989) and Evans and Honkapohja (2001) agents typically have correctly specified reduced-form models but update their parameter estimates in real-time. In many models, these expectations converge to rational expectations. In Sargent (1999) and Williams (2004), agents may have misspecified econometric models but within the context of their subjective model they are unable to detect their misspecification. In Branch and Evans (2006a) and Evans and Ramey (2006), degrees of freedom or cognitive limitations force agents to underparameterize their forecasting models. These self-referential models restrict beliefs and the nature of misspecification to be determined in equilibrium.

In this paper, we apply the econometric misspecification approach to asset pricing questions. We develop our results in the context of an asset pricing model in which the stock price depends on expected future returns and on an exogenous process for share supply. Our modeling of share supply is meant to proxy for asset float, as discussed in Ofek and Richardson (2003), Cochrane (2005) and Hong, Scheinkman, and Xiong (2005). We are motivated, in part, by Hong, et al. who demonstrate strong empirical implications from a model of heterogeneous expectations, increasing supply of shares, and short-sales constraints. Following the approach of Branch and Evans (2006a), we assume that agents underparameterize their forecasting model for price:

¹An important counter viewpoint is provided by Fama and French (1996).

agents perceive price as depending on dividends or share supply, but not both.² This simple framework is meant to stand in for a more complex environment in which traders face uncertainty about their model specification and choose parsimonious trading strategies.³ We assume that agents choose only between models, or trading strategies, that yield the highest (or nearly the highest) risk-adjusted trading profits. Within this class of underparameterized models, the key condition restricting beliefs is that model parameters must satisfy a least-squares orthogonality condition. Agents' forecasting models are statistically optimal in the sense that their forecast errors are orthogonal to their predictor.

It is worth expanding on the motivation for our assumption that agents employ well-chosen but underparameterized forecasting models. In the econometric learning approach, economic agents are modeled as econometricians. It follows that one should take seriously the applied econometric problem of choosing a satisfactory specification of the regressors. Although various types of misspecification could be considered, we here, as in our earlier work, focus on the implications of agents choosing between alternative underparameterized models. In doing so we are investigating the implications of agents following the advice of many econometricians to choose "parsimonious" forecasting models, i.e. to omit variables or lags that are not clearly essential for improving the statistical fit. At least as early as Nelson (1972) it has been known that simple, parsimonious models often empirically outperform more complex models in out-of-sample forecasting.

The central reason to expect the use of underparameterized models is that the economic environment is complex relative to the amount of data typically available for estimation of alternative forecasting models. Forecasting models face a degrees of freedom limitation, dictated by the available sample size, which typically force the use of simplified models. Long before the available degrees of freedom are actually exhausted, the quality of estimated forecast models deteriorates because of parameter uncertainty: the mean square error of parameter estimates and of forecasts can be reduced by omitting relevant variables if the sample size is not sufficiently large. Furthermore, if there is actual or perceived ongoing structural change in the economy, the degrees of freedom problem cannot be expected to diminish in severity asymptotically, since the effective sample size, which discounts past data, remains constant over time. These considerations motivate our focus on underparameterized models.⁴

²One could generalize the model further by assuming dividends and asset float follow multivariate stochastic processes with high order lags, and agents are restricted to underparameterize in at least one dimension. The main qualitative findings of this paper would extend to this more general formulation.

³Although we motivate our underparameterization restriction by noting that agents may face degree of freedom or cognitive limitations, we could also appeal to psychological explanations such as investor inattention. The implications of investor inattention are considered by Peng and Xiong (2006) and Hirshleifer and Teoh (2004), in distinct settings.

⁴In Section 4.3 we look at the robustness of our results to expanding the list of forecast models to

Many financial economists embrace bounded rationality as a way of explaining the existence of multiple trading strategies, heterogeneity in expectations and preferences, volatility and under/over reaction to economic news. A gap exists in the literature that we seek to fill by studying the existence of these multiple trading strategies, and evolution over time, as an equilibrium phenomenon. To these ends, this paper makes a number of contributions. We demonstrate that underparameterization and misspecification equilibria can arise in a simple asset pricing model. Depending on the deep parameters of the model, there may exist multiple misspecification equilibria in asset prices. As we show, agents' misspecification regarding the price process affects their perceptions about the return and riskiness of stocks. When there are multiple equilibria, traders will hold different perceptions of return and risk at each equilibrium. This implies that the mean and variance of excess returns will differ in each equilibrium.

In a related approach, Hong, Stein, and Yu (2005) assume that dividends are driven by two exogenous processes and agents can only condition their expectations on one part of the process.⁵ Their model, however, does not fully exploit the self-referential nature of asset pricing models. Instead, they appeal to behavioral and psychological explanations. While behavioral approaches are interesting and important there is still an open question of whether one can address these financial market anomalies and still assert the kind of discipline imposed by rational expectations. In a rational expectations model, the self-referential feature of the model requires that both the forecasts generated from the model and the market outcomes be jointly determined. In the approach presented in this paper, the parameters of the forecasting model, the perceived riskiness of stocks, and the distribution of agents across models are jointly determined. The additional equilibrium feedback effect in our approach makes multiple equilibria possible, which as discussed below, has important empirical implications.

This paper is most closely related to Timmermann (1994,1996), who demonstrates that an asset pricing model, where rational expectations are replaced with adaptive learning, is capable of generating excess volatility. Similarly, Barsky and DeLong (1993) show that a model with adaptive learning can generate autocorrelation patterns consistent with the data. We, like Timmermann, assume agents are boundedly rational, but unlike Timmermann we confront investors with a list of misspecified models. The theoretical novelty of our approach is that we pin down the beliefs and distribution of agents across models as an equilibrium outcome jointly determined

include a bivariate model as well as the univariate models. In future research it would be of interest to allow for an even larger set of possible specifications.

⁵In Barberis, Shleifer, and Vishny (1998) dividends follow a multi-layered Markov chain that proxies for a simple model in which there are two different Markov processes governing dividends, one with high persistence and one with low persistence. Agents, though, only believe in one of the two models; hence the underparameterization.

with stock price. The multiple equilibria, distinct to our theoretical models, has important empirical implications since it provides a potential explanation for Markov switching returns.

The equilibrium in the model described so far implies that trading strategies and expectations are time-invariant. We also consider a real-time learning and dynamic predictor selection version in order to study the model's ability to capture empirical regularities in excess return dynamics. In this extension of the basic model, agents update in real-time their parameter estimates and a geometric average of past trading profits. They then decide on their predictor and holdings of the risky asset conditional on these real-time estimates. We demonstrate that, with the model calibrated to U.S. stock data, the model regime-switching dynamics, as the model switches in real time between equilibria, matches the Markov-switching in returns found in Guidolin and Timmermann (2005a).

In addition to the least squares learning approach, pioneered in finance by Timmermann (1994,1996) and further developed here, there is also a literature on Bayesian, or rational, learning. For example, Brennan and Xia (2001), Lewellen and Shanken (2002), and Pastor and Veronesi (2003) show that introducing uncertainty about the process driving stock prices, such as parameter uncertainty, allowing agents to form priors and update them with real-time data, can generate returns that capture many features of the data. Guidolin and Timmermann (2007b,2008) demonstrate the optimal portfolio implications of Bayesian learning. The novelty of our theoretical model is its ability to generate multiple equilibria which, as we demonstrate, is able to explain the regime switching return dynamics observed in data.

This paper proceeds as follows. Section 2 presents the model. Section 3 presents theoretical results. Section 4 discusses the empirical implications and presents the calibrated version of the model. Section 5 concludes.

2 Asset Pricing Model with Restricted Perceptions

We employ a mean-variance linear asset pricing model, similar to DeLong, Shleifer, Summers, and Waldmann (1990).⁶ We make this assumption for two analytical reasons: first, so that demands remain bounded; second, so that demand is linear. One can justify the assumption based on a log approximation to exponential utility with (perceived) Gaussian returns in an overlapping generations framework.⁷

⁶A similar modeling strategy is pursued in Grossman and Stiglitz (1980), Brock and Hommes (1998), Hong, Scheinkman, and Xiong (2005), among many others, in assuming that agents optimize with respect to the mean-variance efficient frontier.

⁷See Branch and Evans (2008) for details of an overlapping generations model that leads to an exogenous share supply process.

There is a single risky asset that yields a dividend stream $\{y_t\}$, and trades at price p_t , net of dividends. There is a risk-free asset that pays a rate of return $R = \beta^{-1} > 1$. Households, at time t , solve

$$\max_{z_t} E_t^* W_{t+1} - \frac{a}{2} E_t^* \text{Var}_t W_{t+1}$$

subject to

$$W_{t+1} = RW_t + (p_{t+1} + y_{t+1} - Rp_t)z_t$$

where E^* denotes (possibly) non-rational expectations, z_t is the holdings of the risky asset, p is its price, y are dividends, and $R > 1$ is the risk-free rate of return. We assume that dividends follow a stationary AR(1) process,

$$\hat{y}_t = (1 - \rho)y_0 + \rho\hat{y}_{t-1} + \varepsilon_t$$

where ε_t is mean-zero with variance σ_ε^2 . In equilibrium, the demand for shares must equal supply. The usual assumption is that the supply of shares is constant and normalized to one. We, however, assume a stationary AR(1) process representing the supply of shares:

$$\hat{z}_{st} = (1 - \phi)s_0 + \phi\hat{z}_{st-1} + \nu_t$$

The stochastic disturbance ν_t is mean-zero, with variance σ_ν^2 , and is possibly correlated with ε_t , i.e. we allow for $\sigma_{\nu\varepsilon} \neq 0$.

As previously mentioned, we interpret the share process \hat{z}_{st} as a proxy for asset float. Asset float is the change in the supply of shares usually after a lock-up period following an initial public offering. Recent papers by Cochrane (2005), Lamont and Thaler (2003), Hong, Scheinkman, and Xiong (2005) show that float can have an effect on price. We view asset float and the supply of shares more generally so that it also includes stock repurchases. Assuming that the supply of shares follows an AR(1) is an obvious analytic device. This paper is a first step at incorporating equilibrium underparameterization and learning into an asset pricing model and leaves a more comprehensive theory of asset share supply to future research. In the calibrated version of the model, below, we estimate an AR(1) for share supply using U.S. data.

There are two types of agents, each omitting some relevant information from their forecasting model when they solve the above problem. One type omits the role of supply in affecting price while the other omits the dividends process. We make this assumption to bring some realism to the asset pricing model. Because of a preference for parsimony agents are assumed to underparameterize their model. This is the same motivation of Hong, Stein, and Yu (2005) in the case where agents omit a portion of the dividend process from their forecasting model. Heterogeneous expectations also arise in Hong, Scheinkman, and Xiong (2005). A novelty to our approach is that we endogenize the parameters and distribution of agents across these underparameterized models. In contrast to Hong, Stein, and Yu (2005), we emphasize parameter and

trading strategy uncertainty in self-referential models. Although we assume that dividends are a univariate stochastic process, one could easily extend dividends and share supply to bivariate VAR processes along the lines of Branch and Evans (2006a). The main innovation to our approach is that we pin down *both* the forecasting model parameters *and* the distribution of agents across models as an equilibrium object. We then can use real-time learning to study the dynamics and to speculate on the model's ability to address financial market puzzles.

Each agent type j solves

$$\max_{z_{jt}} RW_t + E_t^j (p_{t+1} + \hat{y}_{t+1} - Rp_t) z_{jt} - \frac{a}{2} \sigma_{jt}^2 z_{jt}^2$$

where $\sigma_{jt}^2 = \text{Var}_{jt}(p_{t+1} + \hat{y}_{t+1} - Rp_t)$ is the subjective conditional variance of the excess rate of return. An important feature of our analysis is that the value of σ_j^2 will be pinned down in equilibrium. The first-order condition leads to the demand for type j of,

$$z_{jt} = \frac{1}{a\sigma_{jt}^2} E_t^j (p_{t+1} + \hat{y}_{t+1} - Rp_t)$$

The responsiveness of demands z_{jt} to expected rates of return depends on $a\sigma_{jt}^2$, which it will be convenient for us to call “perceived risk.” Note that perceived risk is a product of the subjective conditional variance, an equilibrium object, and the degree of risk aversion, a .

Financial market equilibrium requires that price adjusts to ensure market clearing. Let n denote the fraction of agents with expectations E_t^1 . In equilibrium,

$$nz_{1t} + (1 - n)z_{2t} = \hat{z}_{st}$$

which leads to the equilibrium process for stock prices,

$$p_t = \beta (y_0 + \rho y_t) + \beta \left(\frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right)^{-1} \left[\frac{n}{a\sigma_{1t}^2} E_t^1 p_{t+1} + \frac{1-n}{a\sigma_{2t}^2} E_t^2 p_{t+1} - \hat{z}_{st} \right] \quad (1)$$

where for convenience we write $\beta = R^{-1}$. To derive (1) we have assumed that $E_t^1 \hat{y}_{t+1} = E_t^2 \hat{y}_{t+1} = y_0 + \rho y_t$, and $y_t = \hat{y}_t - y_0$ are dividends written in deviations from mean form.

We envision underparameterization in part because degrees of freedom constraints prevent agents from regressing price on all available information. Agents know the univariate processes for dividends and supply, but we assume that agents do not incorporate both elements into their forecasting model for *price*. It might appear contradictory that agents know the processes for dividends and supply, yet do not use all known information when forecasting stock price. In this simple setting this is, perhaps, unrealistic. But, if one thinks of all of the factors that might be influencing dividends, share supply, and price, the total number of factors with non-trivial

predictive power would exceed degrees of freedom constraints. If dividends and share supply were actually high order vector autoregressive processes, possibly correlated, then forecasting future dividends and supplies are curtailed by the number of parameters of the model. For example, an n -variable VAR(p) has $n^2 \times p$ coefficients to estimate, plus the parameters of the autocovariance matrix. At the monthly frequency, the degrees of freedom would quickly evaporate.

Hong, Stein, and Yu (2005) and Barberis, Shleifer, and Vishny (1998) also assume underparameterized forecast models, though they impose particular misspecified models exogenously. These authors instead motivate the assumption by appealing to psychology research that suggests people forecast using simple paradigms or reference models. One could also extend their motivations to our approach. Our theoretical interest, though, is to impose some modeling discipline on these deviations from full information: in our framework, within the context of their forecasting models, agents are unable to detect their misspecification. Remarkably, the theoretical and empirical implications of this approach are rich.

Agents forecast by projecting a perceived law of motion (PLM) for price. The set of PLMs, given the underparameterization restriction, are:

$$\begin{aligned} PLM_1 : p_t &= b_0^1 + b_1^1 y_t + \eta_t \\ PLM_2 : p_t &= b_0^2 + b_1^2 z_{st} + \eta_t \end{aligned}$$

where η_t is a perceived exogenous white noise shock, and $z_{st} = \hat{z}_{st} - s_0$. This implies expectations of the form,

$$\begin{aligned} E_t^1 p_{t+1} &= b_0^1 + b_1^1 \rho y_t \\ E_t^2 p_{t+1} &= b_0^2 + b_1^2 \phi z_{st} \end{aligned}$$

Plugging these expectations into (1) leads to the following actual law of motion (ALM) for price,

$$\begin{aligned} p_t &= \beta \left[y_0 + \left\{ \frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right\}^{-1} \left(\frac{n}{a\sigma_{1t}^2} b_0^1 + \frac{1-n}{a\sigma_{2t}^2} b_0^2 - s_0 \right) \right] \\ &+ \beta \left[1 + \left\{ \frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right\}^{-1} \frac{n}{a\sigma_{1t}^2} b_1^1 \right] \rho y_t \\ &+ \beta \left\{ \frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right\}^{-1} \left[\frac{1-n}{a\sigma_{2t}^2} b_1^2 \phi - 1 \right] z_{st}, \end{aligned}$$

or,

$$p_t = \xi_0(n) + \xi_1(n)y_t + \xi_2(n)z_{st} \tag{2}$$

where

$$\begin{aligned}\xi_0(n) &= \beta \left[y_0 + \left\{ \frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right\}^{-1} \left(\frac{n}{a\sigma_{1t}^2} b_0^1 + \frac{1-n}{a\sigma_{2t}^2} b_0^2 - s_0 \right) \right] \\ \xi_1(n) &= \beta \left[1 + \left\{ \frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right\}^{-1} \frac{n}{a\sigma_{1t}^2} b_1^1 \right] \rho \\ \xi_2(n) &= \beta \left\{ \frac{n}{a\sigma_{1t}^2} + \frac{1-n}{a\sigma_{2t}^2} \right\}^{-1} \left[\frac{1-n}{a\sigma_{2t}^2} b_2^2 \phi - 1 \right]\end{aligned}$$

In the sequel, we will suppress the dependence of ξ_j on n . In a rational expectations equilibrium (REE),

$$\begin{aligned}\xi_0 &= \frac{\beta}{1-\beta} (y_0 - a\sigma^2 s_0) \\ \xi_1 &= \frac{\beta\rho}{1-\beta\rho} \\ \xi_2 &= -\frac{\beta a\sigma^2}{1-\beta\phi}\end{aligned}$$

and $\sigma^2 = (1 + \xi_1)^2 \sigma_\varepsilon^2 + \xi_2^2 \sigma_\nu^2$, is the rational expectations equilibrium value for the perceived riskiness of the risky asset.

Although agents in the model are assumed to have underparameterized forecasting models (restricted perceptions), we require that they forecast in a statistically optimal manner. We require that the forecast model parameters are optimal linear projections. That is, the belief parameters b^j , $j = 1, 2$ satisfy the following least-squares orthogonality conditions,

$$E(1, y_t)' (\xi_0 + \xi_1 y_t + \xi_2 z_{st} - b_0^1 - b_1^1 y_t) = 0 \quad (3)$$

$$E(1, z_{st})' (\xi_0 + \xi_1 y_t + \xi_2 z_{st} - b_0^2 - b_1^2 z_{st}) = 0 \quad (4)$$

or,

$$\begin{aligned}b_0^j &= \xi_0, \quad j = 1, 2 \\ b_1^1 &= \xi_1 + \xi_2 r \\ b_1^2 &= \xi_2 + \xi_1 \tilde{r}\end{aligned}$$

where $r = E y_t z_{st} / E y_t^2$, $\tilde{r} = E y_t z_{st} / E z_{st}^2$. Orthogonality conditions like (3) or (4) appear frequently in the macroeconomics literature. For example, Sargent (1999), Cho, Williams, and Sargent (2003) define a self-confirming equilibrium with respect to a very similar condition. Evans and Honkapohja (2001) show that under adaptive learning an underparameterized forecasting model may converge to a set of parameters that satisfy an orthogonality condition like (3). Many other applications that

employ (3) are discussed in Branch (2006) and Sargent (2008). The key feature of orthogonality conditions like (3),(4), are that within the context of their forecasting model, agents are unable to detect their misspecification.⁸

Underparameterized expectations affect not only agents' subjective expected return of stocks but also their subjective conditional variance of stock returns. Thus, b_0^j, b_1^j implies *perceived* $\sigma_{jt}^2 = E_t(p_{t+1} + \hat{y}_{t+1} - E_t^j(p_{t+1} + \hat{y}_{t+1}))^2$. Given the solutions for b_0^j, b_1^j the restricted perceptions values for perceived conditional variances are,

$$\sigma_1^2 = \left((1 + \xi_1)^2 + \frac{\xi_2^2 r^2 \rho^2}{1 - \rho^2} \right) \sigma_\varepsilon^2 - 2\xi_2^2 \rho \phi r Eyz + \frac{\xi_2^2}{1 - \phi^2} \sigma_\nu^2 \quad (5)$$

$$\sigma_2^2 = \left((1 + \xi_1)^2 + \frac{\xi_1^2 \rho^2}{1 - \rho^2} \right) \sigma_\varepsilon^2 - 2\xi_1^2 \rho \phi \tilde{r} Eyz + \left(\frac{\xi_1^2 \tilde{r}^2 \phi^2}{1 - \phi^2} + \xi_2^2 \right) \sigma_\nu^2 \quad (6)$$

Clearly, the RPE values for b_0^j, b_1^j depend jointly on σ_j^2 for $j = 1, 2$. The existence of an RPE is non-trivial.

Given exogenous processes y_t, z_{st} , and ξ_j, σ_j^2 $j = 1, 2$, and given the proportion n of agents using forecast model $j = 1$, a Restricted Perceptions Equilibrium (RPE) is then defined as a stochastic process $\{p_t\}$ of the form (2), where the coefficients satisfy

$$\xi_0 = \frac{\beta}{1 - \beta} \left[y_0 - \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}^{-1} s_0 \right],$$

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 1 - \frac{\beta \rho n}{a\sigma_1^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} & -\frac{\beta \rho nr}{a\sigma_1^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} \\ -\frac{\beta \phi (1-n) \tilde{r}}{a\sigma_2^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} & 1 - \frac{\beta (1-n) \phi}{a\sigma_2^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\beta \rho}{\beta} \\ -\frac{\beta \rho}{\left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} \end{bmatrix}$$

A general existence result is not available; ξ_1, ξ_2 depend non-linearly on σ_1^2, σ_2^2 , whose solutions also depend non-linearly on ξ_1, ξ_2 . However, the following result holds in the most empirically relevant case of very weakly correlated dividends and share supply and values of $n \in \{0, 1\}$.⁹

Proposition 1 *Let $r, \tilde{r} \rightarrow 0$. For sufficiently small values of a and $n \in \{0, 1\}$, there exists a unique RPE.*

Although agents in the model are underparameterizing their forecasting models, each agent's forecast does reflect the influence of that part of the omitted variable that

⁸Of course, if they step out of their model and run specification tests they could detect the misspecification. Below, we will see in real time simulations that they may still choose, for finite time, underparameterized models

⁹All proofs are in the Appendix.

is correlated with the variables used in their forecast. This property arises because of the orthogonality condition, which is satisfied in equilibrium: b^j depends on the two reduced-form parameters ξ_1, ξ_2 and also on the regression coefficient (r, \tilde{r}) . In addition, asset prices aggregate and reflect all available information – in this sense asset prices are partially revealing.

It is important to note that the model is self-referential: b^j and σ_j^2 , hence ξ_j are not free parameters but are equilibrium objects. For similar reasons, we do not want to treat n as a free parameter and now proceed to make it endogenous. In consequence, although agents use misspecified forecast models, there are still important cross-equation restrictions imposed on the dynamics that are analogous to the restrictions obtained under fully rational expectations.

In order to pin down n , we need a metric for evaluating forecast success. In order to stay in line with the assumption that agents are mean-variance maximizers, we also assume that agents adjust their trading profits for variance when deciding on forecast success. Thus, we assume that each agent ranks the two forecasting models according to,

$$U^j = E\pi_t^j - \frac{a}{2}\sigma_j^2 E z_{jt}^2$$

where $\pi_t^j = (p_{t+1} + \hat{y}_{t+1} - Rp_t) z_{jt}$ and E is the (unconditional) expectations operator. Note U^j measures the profits in certainty equivalence units of the good.

The Appendix computes π_t^j and $E z_{jt}^2$ for $j = 1, 2$. Predictor selection depends on the difference in fitness measures. Define $F(n) : [0, 1] \rightarrow \mathbb{R}$ as $F(n) = U^1 - U^2 = (E\pi_t^1 - E\pi_t^2) - (a/2)(\sigma_1^2 E z_{1t}^2 - \sigma_2^2 E z_{2t}^2)$. Then we can write this expression as

$$F(n) = \frac{1}{2a\beta^2\sigma_1^2\sigma_2^2} (B_0 + B_y E y_t^2 + B_{yz} E y_t z_{st} + B_z E z_{st}^2)$$

where B_0, B_y, B_z and B_{yz} are given by

$$\begin{aligned} B_0 &= \frac{-a\beta^2 s_0^2 (\sigma_1^2 - \sigma_2^2)}{n\sigma_2^2 + (1-n)\sigma_1^2} \\ B_y &= -(\sigma_1^2 - \sigma_2^2)(\xi_1 - \beta\rho) [\xi_1(1 - 2\beta\rho) - \beta\rho] + \beta^2 \rho^2 \sigma_2^2 (\xi_1^2 - r^2 \xi_2^2) \\ B_z &= -(\sigma_1^2 - \sigma_2^2) \xi_2^2 (1 - 2\beta\phi) + \beta^2 \phi^2 \sigma_1^2 (\tilde{r}^2 \xi_1^2 - \xi_2^2) \\ B_{yz} &= 2(1 - \beta\phi)(\sigma_1^2 - \sigma_2^2) \xi_2 [\beta\rho - \xi_1(1 - \beta\rho)] - 2\beta^2 \phi \rho (\tilde{r} \xi_1^2 \sigma_1^2 - r \xi_2^2 \sigma_2^2). \end{aligned}$$

Note that B_0, B_y, B_z and B_{yz} are functions of n because ξ_1, ξ_2, σ_1^2 , and σ_2^2 depend on n .

As in our earlier papers, we follow Brock and Hommes (1997) in assuming a multinomial logit (MNL) approach to predictor selection. The MNL approach has a venerable history in discrete decision making and is a natural way of introducing randomness in forecasting into the present environment. Young (2004) argues that

randomness in forecasting, much like mixed strategies in actions, provides robustness against model uncertainty and flexibility in environments with feedback. In this setting, agents are selecting their forecasting models from a discrete choice set, they are uncertain about the best forecast model specification, and so the MNL map is natural in this setting:

$$n = \frac{\exp(\alpha U^1)}{\exp(\alpha U^1) + \exp(\alpha U^2)}$$

which can be written,

$$n = \frac{1}{2} [\tanh \{\alpha F(n)\} + 1] \equiv T_\alpha(n) \quad (7)$$

In particular, $T : [0, 1] \rightarrow [0, 1]$ is a continuous and well-defined function provided that an RPE exists.

Definition. *A Misspecification Equilibrium n^* is a fixed point of the map $T: n^* = T(n^*)$.*

By Brouwer's theorem, a Misspecification Equilibrium (ME) exists in this model provided an RPE exists. The T-mapping is indexed by the parameter α which is typically called the 'intensity of choice' parameter. Since the MNL map derives from a random utility setting, finite values of α parameterize deviations from full utility maximization. The 'neoclassical' case is $\alpha \rightarrow \infty$. Our interest is mainly in equilibria where all agents choose only the best performing statistical model and so we will focus on the $\alpha \rightarrow \infty$ case.

3 Analytic Results

It is useful to re-write the function $F(n)$ as,

$$\frac{F(n)}{Ey_t^2} = \frac{1}{2\alpha\beta^2\sigma_1^2\sigma_2^2} \left(B_0 \frac{1}{Ey_t^2} + B_y + B_{yz}r + B_z Q \right)$$

where $Q = Ez_{st}^2/Ey_t^2$. The number and nature of Misspecification Equilibria depend on the properties of $F(n)$.

These are complicated expressions and general results are not available. However, using the argument in Branch and Evans (2007), the following result can be used to characterize possible equilibria:

Proposition 2 *Let $N_\alpha^* = \{n^* | n^* = T_\alpha(n^*)\}$ denote the set of Misspecification Equilibria. In the case of large α , N^* has one of the following properties:*

1. If $F(0) < 0$ and $F(1) < 0$ (Condition P0) then $n^* = 0 \in N^*$.
2. If $F(0) > 0$ and $F(1) > 0$ (Condition P1) then $n^* = 1 \in N^*$.
3. If $F(0) < 0$ and $F(1) > 0$ (Condition PM) then $n^* \in \{0, \hat{n}, 1\} \subseteq N^*$, where $\hat{n} \in (0, 1)$ is such that $F(\hat{n}) = 0$.
4. If $F(0) > 0$ and $F(1) < 0$ (Condition P) then $n^* = \hat{n} \in N^*$, where $\hat{n} \in (0, 1)$ is such that $F(\hat{n}) = 0$.

Because we do not know, in general, whether F is monotonic, we cannot rule out the existence of additional equilibria besides those listed. When Condition P0 or Condition P1 holds then either $n^* = 0$ or $n^* = 1$ is a Misspecification Equilibrium. If Condition PM holds then both $n^* = 0$ and $n^* = 1$ are Misspecification Equilibria. Thus, Condition PM is a sufficient condition for multiple equilibria and is the case that will receive further attention below. Condition P implies that there exists an interior Misspecification Equilibrium with heterogeneous expectations. In Branch and Evans (2006a) we said that when Condition P holds the model exhibits Intrinsic Heterogeneity. We explore the existence of Intrinsic Heterogeneity in a companion paper. Notice that under Condition PM there must also be an interior equilibrium \hat{n} for large α . However, because $F(n)$ is a continuous function, Condition PM implies that, whenever $F(n)$ is monotonic, this equilibrium satisfies $T'(F(\hat{n})) > 1$ and hence is unstable. Conversely, under Condition P, for large α there is an \hat{n} at which $F(n)$ crosses through zero from above and, as we showed in our earlier paper, this equilibrium is locally stable. Similarly, when $F(n)$ is non-monotonic it is possible for there to exist interior stable equilibria.

Proposition 2 does not state under which circumstances these conditions will arise. In fact, it does not even state whether all of the cases are possible. The signs of $F(0), F(1)$ depend in a complicated way on $\phi, \rho, a, \sigma_\varepsilon^2, \sigma_\nu^2, E\varepsilon\nu$.

Corollary 3 *Conditions P0, P1, PM and P can each be satisfied for appropriate choices of structural parameters.*

Numerical examples are given below.

Additional analytical results are available for certain limiting cases of interest. In particular, we have:

Corollary 4 *Assume $\rho, \phi > 0$ and $s_0 = 0$. For $|r|, |\tilde{r}|$ sufficiently small we have:*

$$(i) \text{ Condition P0 holds if } Q > \frac{-\beta^2 \rho^4}{a^2 \sigma_1^2 (-(\sigma_1^2 - \sigma_2^2)(1 - 2\beta\phi) - \sigma_1^2 \beta^2 \phi^2)(1 - \beta\rho)^2};$$

(ii) Condition P1 holds if $Q < \frac{-\beta^2 \rho^4 (1-\beta\phi)^2}{a^2 \sigma_2^2 ((2\beta\phi-1)(\sigma_1^2-\sigma_2^2)-\beta^2 \phi^2 \sigma_1^2)}$;

(iii) Condition PM holds if $\frac{-\beta^2 \rho^4 (1-\beta\phi)^2}{a^2 \sigma_2^2 ((2\beta\phi-1)(\sigma_1^2-\sigma_2^2)-\beta^2 \phi^2 \sigma_1^2)} < Q < \frac{-\beta^2 \rho^4}{a^2 \sigma_1^2 (-(\sigma_1^2-\sigma_2^2)(1-2\beta\phi)-\sigma_1^2 \beta^2 \phi^2)(1-\beta\rho)^2}$

This corollary shows the importance of risk aversion and the relative variance of supply shocks. For a given Q , values of perceived risk $a\sigma_j^2$ that are neither too high nor too low lead to multiple equilibria even in the case of low contemporaneous correlation between the exogenous shocks.

3.1 Some intuition

There are two exogenous processes driving asset prices: dividends and the supply of shares. Both stochastic processes though have two effects in (1): the direct effect and an indirect effect acting through expectations.¹⁰ The number and nature of equilibria depend on the balancing of these two effects. Notice that p_t depends positively on expectations. Thus, whether these direct effects are positively or negatively projected onto the asset price depends on the equilibrium belief parameters, which in turn depend on the equilibrium proportion of agents adopting the dividend forecasting model.

The feedback effects are:

$$\begin{aligned} E_t^1 p_{t+1} &= \xi_0 + (\xi_1 + \xi_2 r) \rho y_t \\ E_t^2 p_{t+1} &= \xi_0 + (\xi_2 + \xi_1 \tilde{r}) \phi z_{st} \end{aligned}$$

and $\sigma_{jt}^2 = Var_t^j(p_{t+1} + \hat{y}_{t+1})$. Notice in the expressions for ξ_1, ξ_2 in the special case above of $r, \tilde{r} \rightarrow 0$ that $a\sigma^2$ directly influences the size of ξ_2 and that ξ_2 is negative (because z_{st} has a negative direct effect):

$$\begin{aligned} \xi_1 &= \frac{\beta\rho}{1 - \beta\rho n} \\ \xi_2 &= -\frac{a\sigma^2\beta}{1 - \beta(1 - n)\phi} \end{aligned}$$

In this case where the shocks are uncorrelated, $r = \tilde{r} = 0$, beliefs reinforce the direct effect of dividends and supply of shares. Multiple equilibria arise naturally in this case for a range of perceived risk. The condition on a required for multiple equilibria puts bounds on the importance of the direct effect of z_{st} relative to dividends. If $a\sigma_2^2$

¹⁰Strictly speaking, the timing of the model is that dividends are paid at time $t + 1$, so that it is expected dividends that matter for price. Since agents have common beliefs on dividends, these have a common effect on stock prices, as specified in (1).

is large then the share supply forecast model always dominates, while when $a\sigma_1^2$ is sufficiently low the dividend model is necessarily superior. For intermediate values of perceived risk, either model can emerge as an equilibrium.

3.2 Numerical Examples

In this subsection we turn to numerical examples to illustrate our theoretical results. In each case we plot the T -map, $F(n)$, $\sigma_1^2(n)$, and $\sigma_2^2(n)$. We are interested in large α , so we set $\alpha = 10000$. Above we presented analytic results, for the special case of weakly correlated exogenous processes, and provided some more general intuition. We here choose particular parameter values to illustrate the rich theoretical properties of the model.

3.2.1 Multiple Equilibria

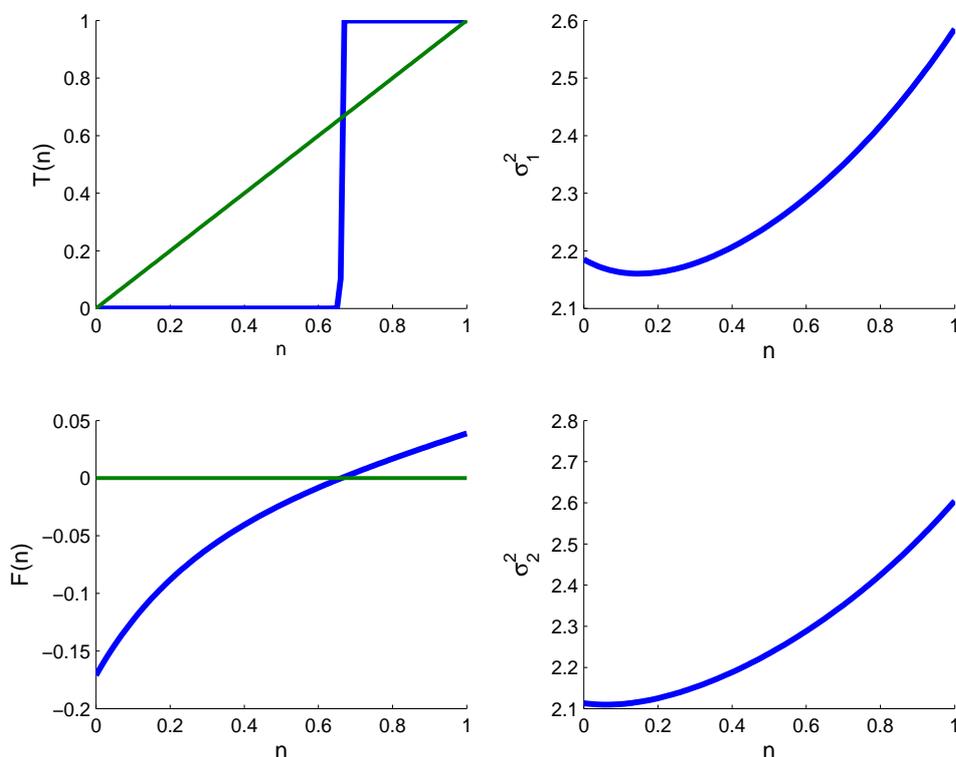
We adopt the parameter values $\rho = 0.4, \phi = 0.6, \beta = .95, \sigma_\nu = 1.25, \sigma_\varepsilon = 1, \sigma_{\nu\varepsilon} = .25, a = 0.1, y_0 = s_0 = 0$. For these parameter values, $r = .2763, \tilde{r} = .16842$. Figure 1 plots (clockwise starting from the northwest frontier): the T -map, the RPE values for σ_1^2 and σ_2^2 respectively and the risk-adjusted profit difference function $F(n)$. Each frontier plots these values against n . A Misspecification Equilibrium occurs when the T -map crosses the 45-degree line.

Notice first that $F(n)$ is monotonically increasing with $F(0) < 0, F(1) > 0$. As a result, the northwest panel demonstrates that there are multiple equilibria, in particular at $n = 0, n = 1$. This result is in line with our earlier intuition of the effect of positive feedback in self-referential models. In this case, there exist three equilibria. Notice, though, that the interior equilibrium occurs at \hat{n} where $F(\hat{n}) = 0$. This equilibrium is unstable in the sense that $T'(\hat{n}) > 1$ and so under a real-time learning and predictor selection dynamic, as considered below, we would not observe the interior equilibrium as an outcome.

In the right-most panels of the figures $\sigma_1^2(n), \sigma_2^2(n)$ are plotted. These panels illustrate the manner in which perceived risk depends on the distribution of agents across misspecified forecasting models.

The propositions and these numerical examples suggest that a , the degree of risk aversion, plays a significant role in the nature of the equilibria. To study this further Figure 2 plots the comparative static effects of changes in a , on the value of n^* , for the parameterization used to generate Figure 1. In particular, Figure 2 is a bifurcation diagram with a as the bifurcating parameter. To generate the figure we consider all

Figure 1: Multiple Misspecification Equilibria.



values of a in the interval $[0, 0.15]$ and plot all corresponding fixed points to T .¹¹

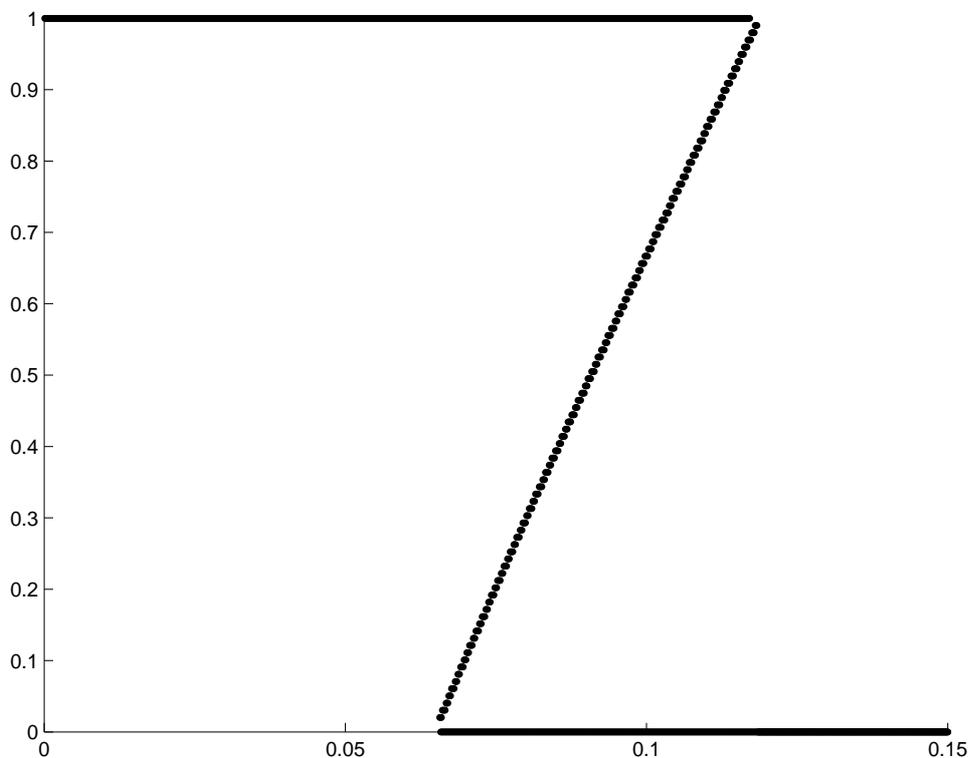
Figure 2 plots the bifurcation diagram. For low values of a there is a unique equilibrium at $n = 1$, for medium a there are multiple equilibria, and for large enough a there is a unique equilibrium at $n = 0$ thereafter. These results are in line with Corollary 4. A similar diagram, of course, exists for Q , the relative variance of dividends.

4 Empirical Implications: Markov Switching Returns and Variances

As a means of highlighting the model's empirical implications we focus on two dynamic properties of asset markets that have received significant attention in the

¹¹In Brock and Hommes (1997) the 'intensity of choice' parameter α was treated as a bifurcation parameter. In this paper, we are primarily interested in $\alpha \rightarrow \infty$ to concentrate on equilibria where all agents only choose the best performing models. Thus, in this stochastic setting the degree of risk-aversion is a more interesting and relevant bifurcation parameter.

Figure 2: Bifurcation Diagram.



finance literature: regime switching means and volatilities in excess returns (c.f. Guidolin and Timmermann (2007)). In the next Section, we calibrate the model and demonstrate its ability to match these empirical features. The current Section aims to illustrate the channels through which the model is capable of matching the empirical regularities.

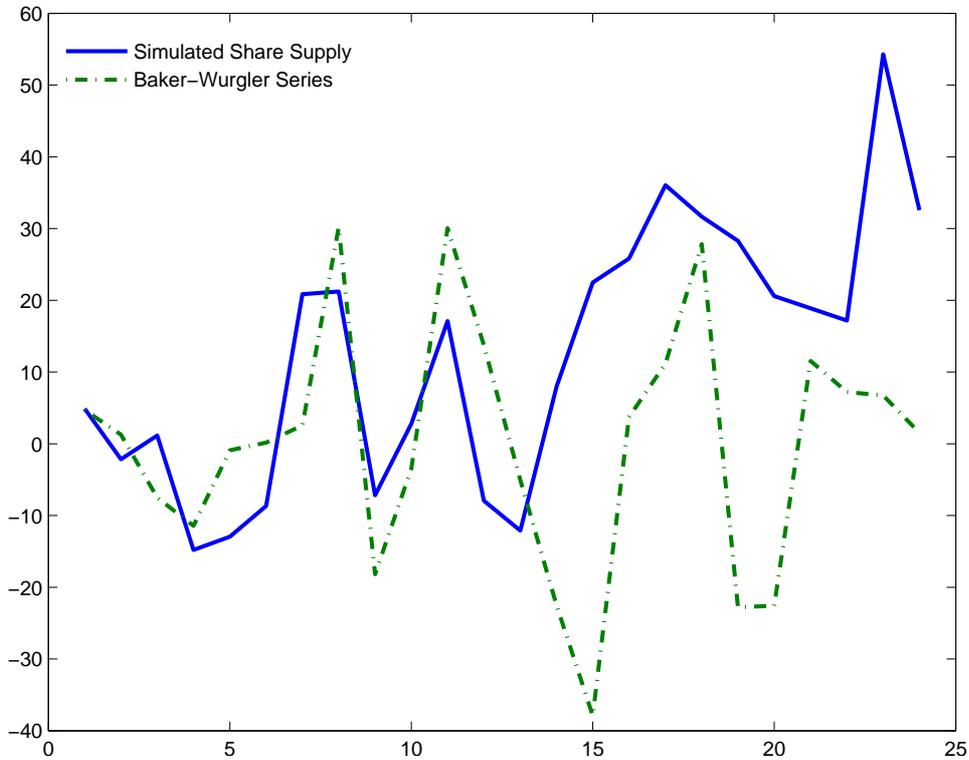
4.1 Calibration

In this Section we assess to what degree the simple model presented here can account for some empirical regularities in excess return dynamics. In order to make a meaningful comparison we need to choose parameter values for the model. This subsection discusses our choice of parameter values.

For the purposes of the model at hand, the most crucial parameters for calibration are the autoregressive parameters and covariances for the dividend and share supply processes. Data on U.S. dividends are widely available. Data on share supply are more limited. For share supply we adopt the series constructed by Baker and Wurgler

(2000).¹² Baker and Wurgler (2000) calculate total new annual (nominal) equity issues in the U.S. Ideally, one would have a data series on all new issues, repurchases, bankruptcies, etc. Such data is not readily available and so the Baker-Wurgler data is the most comprehensive accounting of the U.S. time series of share supply. Figure 3 plots the Baker-Wurgler data against simulated data drawn from the calibrated process for share supply. We calibrate the dividend process from data on corporate profits (after tax) and net dividends from the Economic Report of the President.¹³ The data are reported in nominal terms and we adjust them to 1995 dollars using the consumer price index obtained from the Economic Report of the President.

Figure 3: Simulated and Actual (de-trended) Share Supply Series. The simulated series is a representative sequence from the calibrated AR(1) process. The U.S. share supply series is from Baker and Wurgler (2000), detrended using a linear time trend.



Both dividends (or corporate profits) and share supply exhibit a trend. We detrend the data and estimate an AR(1) for the resulting series. We then calculate the associated AR(1) parameter and standard deviation implied by this regression's

¹²Data obtained from Wurgler's website: <http://pages.stern.nyu.edu/~jwurgler>.

¹³Specifically, we look at corporate profits with inventory evaluation and capital adjustments. The data are obtained from Table B90.

residuals. These are then used as the calibrated values for $\rho, \phi, \sigma_\varepsilon, \sigma_\nu, \sigma_{\nu\varepsilon}$. Table 1 reports the results.

Parameter	Calibration
ϕ	0.6771
ρ	0.8837
σ_ε	0.2235
σ_ν	1.7
$\sigma_{\nu\varepsilon}$	0
β	.9975
a	.145
λ	.01
γ	.01
κ	.50
α	2

Table 1: Calibrated parameter values. Share data comes from Baker and Wurgler (2000). Dividend and C.P.I. data are from the 2005 Economic Report of the President, Table B90.

The remaining parameters are $\beta, \alpha, \lambda, \gamma, \kappa, a$. Following a large literature, we set $\beta = .9975$, based on the one-month risk free rate. The parameters λ, γ, κ govern the rate at which agents adjust their econometric estimates in a real-time learning algorithm, to be introduced below in Section 4.2. Essentially, λ, κ, γ are geometrically declining weights on recent forecast errors. $\lambda = \gamma$ is calibrated at .01, the value reported in a VAR forecasting exercise in Branch and Evans (2006b). We choose a value of $\kappa = .5$. Smaller values of κ imply more smoothing in estimating risk-adjusted profit differences and implies fewer switching between equilibria. We found that the excess returns results reported below were robust to different values of κ , though different values had small effects on the estimated transition probabilities. Picking the value for a is difficult. We choose a value that is empirically not implausible and that leads to the kind of dynamics described above. The parameter a can be thought of as the coefficient of absolute risk aversion. Most experimental studies tend to favor CRRA over CARA, though Holt and Laury report values in the range of 0.1-0.2. In our model, given all of the other parameters, a controls the basin of attraction between equilibria and so will have implications for the frequency of switching and the size of shocks that will induce switching. We set $a = .145$. Smaller values of a tend to increase the proportion of time spent near the $n = 1$ equilibrium and larger values of a to increase the proportion of time spent near the $n = 0$ equilibrium. Finally, we fix $\alpha = 2$, in line with the value considered large but finite in Brock and Hommes (1997), though our results are robust to other values of α . Finally, we treat the constants y_0, s_0 as free parameters by setting $y_0 = 0.5, s_0 = .70$. We take this to

be an approximation to a time trend in dividends and share supply, a consideration we abstract from for analytic convenience. For these parameter values there exist multiple Misspecification Equilibria.

4.2 Regime-Switching Excess Returns

In the real-time learning and dynamic predictor selection version of the model agents do not have fixed beliefs. Beliefs are generated using least-squares in real-time. Time varying parameter estimates make it possible that a sequence of shocks could move the economy from one equilibrium to another (in the case of multiple ME). For this reason, agents will want to remain guarded against the possibility of a regime change and choose their forecasting strategy in real time as well.

Price is now given by the law of motion,

$$p_t = \xi_0(b_0^1, b_0^2, \sigma_{1,t-1}^2, \sigma_{2,t-1}^2, n_{t-1}) + \xi_1(b_{t-1}^1, \sigma_{1,t-1}^2, \sigma_{2,t-1}^2, n_{t-1})y_t \\ + \xi_2(b_{t-1}^2, \sigma_{1,t-1}^2, \sigma_{2,t-1}^2, n_{t-1})z_{st}$$

The timing of the model is that at the end of each period agents update their beliefs of $b^1, b^2, \sigma_1^2, \sigma_2^2$, their risk-adjusted expected profits, and their model choice n . At time t then price depends on the real-time learning and dynamic predictor selection from the end of period $t-1$. We make this timing assumption to avoid the simultaneity between prices and beliefs.¹⁴ Using recursive least-squares (RLS), the belief parameters are calculated as

$$\theta_t^j = \theta_{t-1}^j + \lambda_t R_{j,t}^{-1} x_{j,t-1} \left(p_t - \theta_{t-1}^{j'} x_{j,t-1} \right), \quad j = 1, 2,$$

where

$$R_{j,t} = R_{j,t-1} + \lambda_t (x_{j,t-1}^2 - R_{j,t-1}), \quad j = 1, 2,$$

is the estimated state covariance matrix, $\theta = (b_0^j, b_1^j)'$, and $x_{1t} = (1, y_t)'$, $x_{2t} = (1, z_{st})'$. Using a similar recursive algorithm for the conditional variances of excess returns produces estimates

$$\sigma_{jt}^2 = \sigma_{jt-1}^2 + \gamma_t \left((p_t + \hat{y}_t - E_{t-1}^j(p_t + \hat{y}_t))^2 - \sigma_{jt-1}^2 \right)$$

The terms λ_t, γ_t are typically referred to as gain sequences. Two cases are assumed in the literature: a decreasing gain, $\lambda_t = \gamma_t = t^{-1}$ so that $\lambda_t, \gamma_t \rightarrow 0$; and a constant gain, $\lambda_t = \lambda \in (0, 1), \gamma_t = \gamma \in (0, 1)$. With a decreasing gain, convergence to the restricted perceptions values of $b^1, b^2, \sigma^1, \sigma^2$ is possible. Our interest, though, is in

¹⁴See Evans and Honkapohja (2001) and Brock and Hommes (1997) for further discussion of these issues.

demonstrating the model's implications for its asymptotic dynamics, which will be the central interest in the calibrated version of the model. Thus, we focus on the constant gain case where agents respond to past forecast errors with time-invariant weights λ, γ .

In order to choose their predictors, agents also estimate in real-time the (risk adjusted) expected profits:

$$\hat{EU}_t^j = \hat{E}\pi_t^j - \frac{a}{2}\sigma_{jt}^2\hat{E}z_{jt}^2$$

where

$$\begin{aligned}\hat{E}\pi_t^j &= \hat{E}\pi_{t-1}^j + \kappa \left((1/a\sigma_{jt}^2) (p_t + \hat{y}_t - (1/\beta)p_{t-1}) (E_{t-1}^j p_t + y_0 + \rho y_{t-1} - (1/\beta)p_{t-1}) - \hat{E}\pi_{t-1}^j \right) \\ \hat{E}z_{jt}^2 &= \hat{E}z_{jt-1}^2 + \kappa \left(((1/a\sigma_{jt}^2)^2 (E_{t-1}^j p_t + y_0 + \rho y_{t-1} - (1/\beta)p_{t-1})^2 - \hat{E}z_{jt-1}^2) \right)\end{aligned}$$

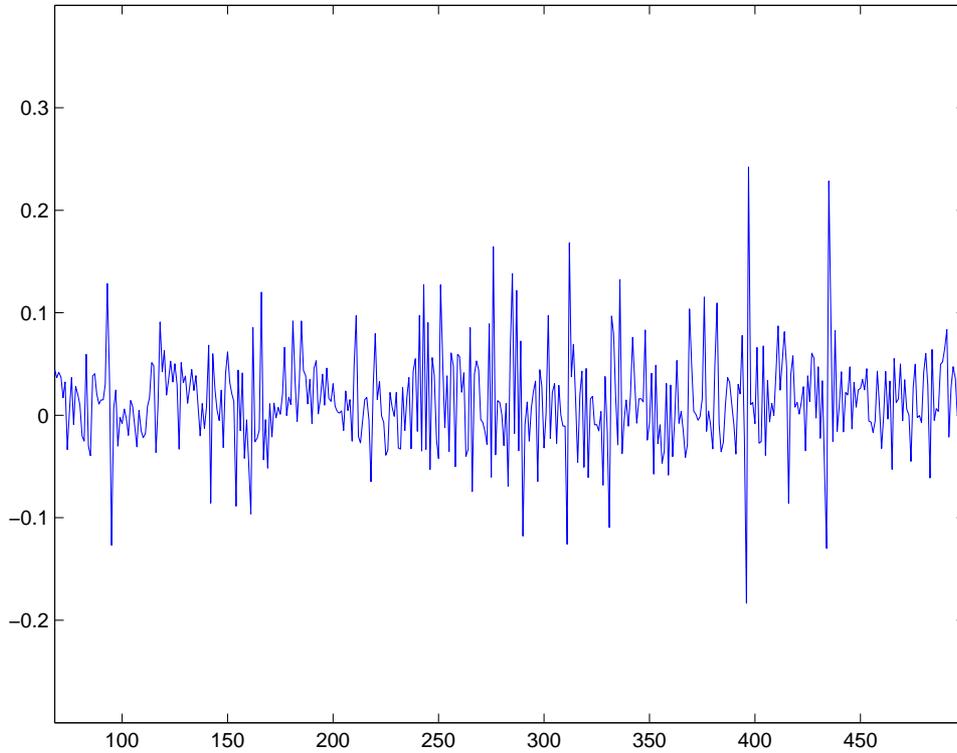
We will also assume a constant value for κ . Using this recursive estimate of expected trading profits, the law of motion for predictor proportions now follows,

$$n_t = \frac{1}{2} \left(\tanh \left[\frac{\alpha}{2} \left(\hat{EU}_t^1 - \hat{EU}_t^2 \right) \right] + 1 \right)$$

We turn to simulations of the real-time version of the model to illustrate the sense in which the model generates regime switching excess rates of return. We assume the calibrated parameter values from Table 1. Larger values for κ than λ implies that agents are more concerned with the possibility of regime change in equilibrium trading strategies than belief parameters. We draw initial values for $n, b^j, \sigma_j^2, R^j, j = 1, 2$ randomly and then simulate the model for a transient period of length 10,000 assuming a decreasing gain for λ, κ . The assumption of a decreasing gain during the transient period ensures that at the beginning of the simulation period the model will be near their equilibrium values. We then simulate the model using the calibrated constant gains for λ, κ . Figure 4 plots the results from a typical simulation.

In Figure 4 the solid line represents the simulated values for the excess rate of return at monthly rates. Two interesting features arise. First, the rates of return switch between periods of higher average rates and lower average rates. Second, the periods of high rates of return coincide with periods of high volatility, and low average rates of return coincide with lower volatility. This leads to, in real-time, the economy switching between high return/high volatility periods and low return/low volatility periods. The switches between these two regimes occur frequently and persistently. By way of comparison, Figure 5 plots actual S& P 500 monthly excess rates of return over the period 1950:1986. Kim, Nelson, and Startz (1998) find that the returns in Figure 5 are consistent with a Markov-switching model in means and variances. The data exhibit patterns very similar to the simulation in Figure 4.

Figure 4: Regime-switching Excess Returns.

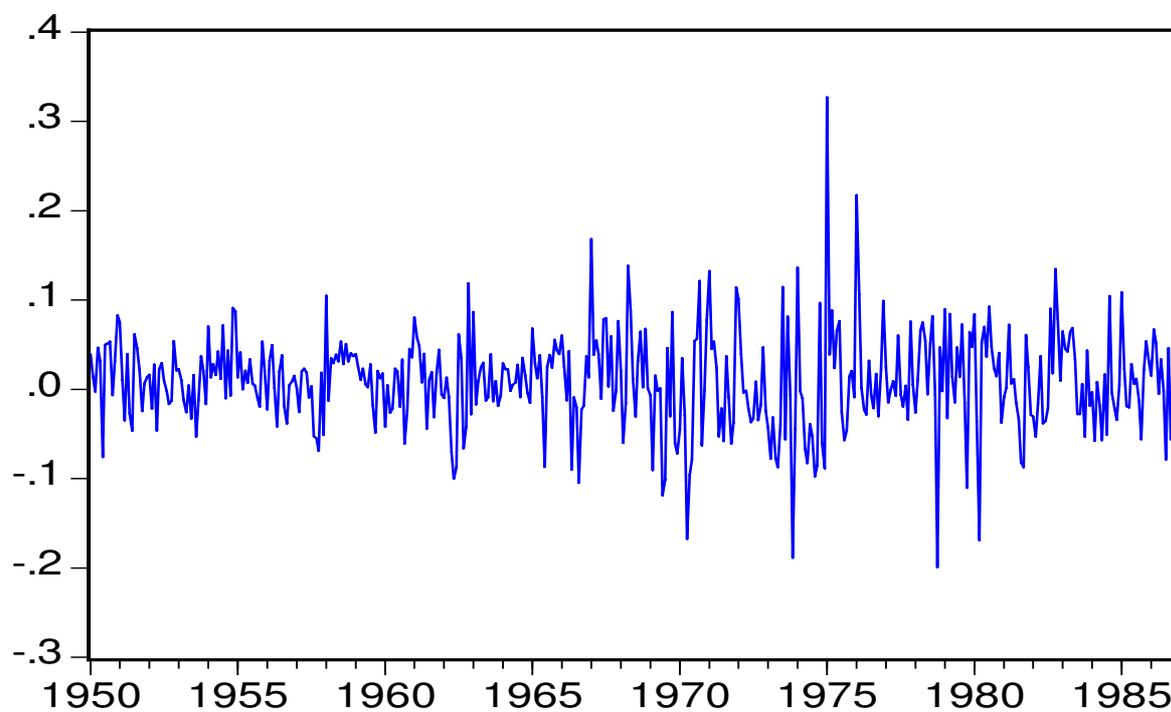


The intuition for why the economy may switch from one equilibrium to another revolves around the interaction between the exogenous shocks and the gain parameters λ, γ, κ . A particularly large shock, mediated through beliefs via λ, γ, κ , may induce agents to switch forecasting models – thus, jumping the economy from one basin of attraction to another. Because λ, γ, κ are positive constants, there are repeated realizations of shocks sufficiently large to switch the economy between equilibria. The persistence in a particular shock, and the frequency with which these regime switches occur, are governed by a complicated interaction between the gain parameters λ, γ, κ , the intensity of choice parameter α , and the stochastic shocks y_t, z_{st} .

This insight of real-time learning and dynamic predictor selection leading to interesting and complicated long-run dynamics in excess returns is the motivation for the next Section which turns to a calibrated version of the real-time model.

One might wonder whether the switching between equilibria evident in Figure 5 might present an exploitable trading opportunity for agents who incorporate into their forecasting model both dividends and share supply. In the next Section, which gives further analysis of our calibrated version, we consider several extensions in which an agent, given more information, will still select parsimonious trading strategies.

Figure 5: Regime-switching Excess Returns in S&P 500 1926-1986.



4.3 Matching the Data

There is a large literature on excess volatility and volatility clustering in returns. For example, Turner, Startz, and Nelson (1989) find evidence for regime switching conditional heteroskedasticity in stock market returns. Bollerslev, Chou, and Kroner (1992) find ARCH effects in stock returns. The discussion of under/overreaction above is a subset of a much broader literature on long-run predictability of stock returns.

Recently, one way this predictability has arisen is through Markov-switching in mean returns for financial variables (Ang and Bekaert (2002)). There is also a literature that makes the link, at the individual stock level, between idiosyncratic volatility and average returns (Merton (1987)). Guidolin and Timmermann (2005, 2007b, 2008) provide evidence for aggregate U.S. stock data that suggests that average returns and volatility follow a finite-state Markov switching process.

Section 4.2 showed that an empirical implication of the present model is that excess returns may follow a regime switching process, thereby exhibiting both persistence and volatility clustering. We now study this issue more systematically by generating time-series data on excess returns from the calibrated version of the model. Our methodology is to take the real-time learning and dynamic predictor selection version of the model, as developed above, parameterize the model according to Table 1, and generate estimates of the means, variances, and regime durations. We then compare our simulated results with the estimates reported by Guidolin and Timmermann (2007). To generate these estimates we simulate the model for a transient period of length 10000, we then store as data the next 5000 periods. We identify the data in each simulation according to its “regime,” i.e. whether $n = 0$ or $n = 1$ and within each regime we calculate the average excess rate of return $\bar{R}_j, j = 0, 1$ and variance $\sigma_j^2, j = 0, 1$. We repeat this 5000 times and report the mean estimates.¹⁵ Table 2 reports the mean value of these calculations across all simulations, and compares them to Guidolin and Timmermann.

Regime-Switching Returns and Volatility							
	\bar{R}_0	\bar{R}_1	$\bar{\sigma}_0^2$	$\bar{\sigma}_1^2$	π_0	π_1	\bar{R}
Simulated Data	.352	.144	.1728	.048	.3998	.6002	.18
Guidolin-Timmermann Data	.358	.0156	.1291	.1167	.5327	.4673	.0792

Table 2: Summary statistics for simulated model and Guidolin and Timmermann (2007).

Table 2 shows that the model yields volatility clustering and persistence in returns. The $n = 0$ state has higher average excess returns than the $n = 1$ state and higher volatility. Because the model switches between states these patterns are persistent across time. However, unlike in Guidolin and Timmermann, the switches are not

¹⁵Numerical explorations suggest that 5000 simulations of 15000 periods each produced stable results, suggesting that the model has converged to a unique invariant distribution.

governed by a Markov chain but occur as unanticipated shocks push the stock price from one basin of attraction to another. Estimates are monthly at annualized rates.

Table 2 also shows that the calibrated model delivers estimates that are similar to the data of Guidolin and Timmermann (2007). The table reports the relative average returns and variances across the high and low states, calculated from the data in Guidolin and Timmermann (2007).¹⁶ The high state has a return of the same magnitude as the simulated data, however, the variance is slightly lower in the data. For the low state, the simulated data have a higher mean excess return and lower variance than reported in Guidolin and Timmermann, though the magnitudes are reasonably close. The simple model presented here delivers simulated data surprisingly close to U.S. stock market data. A larger scale model, with more realistic shock processes, would be likely to fit the data even better. We note that for other parameter choices, such as a smaller value for a would lower the within regime average excess return but deliver an overall average return in line with the data.

The results presented illustrate that in economic environments where traders learn and fine-tune their models in real-time, excess returns will exhibit regime-switching means and variances. It is reasonable to wonder whether this conclusion follows from the underparameterization restrictions we impose on agents, or whether regime-switching dynamics may persist even when agents may select from a fully specified model. In particular, would agents continue to adhere to switching between underparameterized models if they observed data like in Figure 4? To address this question we report on the following experiment. We augment the baseline model to include a third predictor: a bivariate regime-switching model. Agents may choose from the two parsimonious models or may choose a bivariate model able to capture well the regime-switching behavior exhibited above. We assume that the regime-switching model is a threshold, or self-exciting, model as in Potter (1995), where agents who adopt the bivariate model are assumed to believe the regime is triggered by the average excess rate of return over the previous three months rather than an exogenous Markov chain. We assume that whenever the excess rate of return is greater than \bar{R} , then the bivariate model identifies the market as being in the high return period. In the results presented below we set $\bar{R} = 0.15$. The question we ask is whether the underparameterized models would persist when a bivariate regime-switching model may also be adopted by agents. Table 3 reports on the results from this experiment, where the model is simulated for 40,000 periods, the final values for n_0, n_1, n_2 are recorded, and then averaged across 1000 simulations.

When the gain on predictor fitness is small ($\kappa = .001$), then, on average, 86% of the time agents will adopt the bivariate model. In this case we also find that the

¹⁶Guidolin and Timmermann estimate a regime-switching model where the means and variances depend on a 4-state Markov chain. We use their estimates to construct a corresponding two state Markov chain.

Results with a bivariate model as a choice			
κ	n_0	n_1	n_2
0.001	0.0247	0.1065	0.8687
0.50	0.222	0.2678	0.5122

Table 3: Average predictor proportions in simulations with a bivariate regime-switching model choice.

parameters of the bivariate model become close to those of the REE.¹⁷ However, when the gain is set at a higher level ($\kappa = 0.5$), agents will use one of the two univariate models almost half the time. Furthermore, the bivariate model itself exhibits regime-switching behavior, with substantially different coefficients in its two states.¹⁸ Thus with higher gains for predictor fitness our model continues to exhibit regime-switching behavior even when a bivariate model is available that could in principle converge to the REE. Adding a small cost to using the more complex model, either to reflect the costs of using more sophisticated models, as argued in Brock and Hommes (1997), or to incorporate a psychological preference for simple models, would lead agents to rely on simple models in most periods.

Thus the findings of Table 3, for the case $\kappa = 0.5$, strengthens our findings of endogenous regime-switching returns and volatilities. When agents place a sufficiently high weight on recent performance, they will frequently choose one of the parsimonious forecasting strategies, even when a bivariate model is available, and the bivariate model will itself have two distinct states. Thus endogenous regime-switching can arise in a real time setting even when traders are not restricted to underparameterized models.

5 Conclusion

This paper has developed a theory of underparameterization and learning in a simple asset pricing model. Asset price is driven by expectations of future price and exogenous processes for dividends and the supply of asset shares, where the latter is viewed as a proxy for asset float. Agents forecast price by projecting it onto either dividends or share supply. Although agents are forced to underparameterize, we assume that they attempt to do so in an optimal way, through our twin assumptions that the fore-

¹⁷In our experiments, we also confirmed an analogous result for decreasing gain sequences for κ , except that during the learning transition the stochastic nature of the market leads to non-trivial fractions of agents to adopt one of the more parsimonious model.

¹⁸In essence we then obtain a model with regime-switching between four states.

cast models impose the relevant orthogonality conditions and that agents choose only models that maximize, or almost maximize, risk-adjusted expected trading profits. In our framework, model parameters and the distribution of agents across forecasting models are jointly determined in equilibrium. The approach advocated in this paper can be seen as a generalization of Hong, Stein, and Yu (2005) and Barberis, Shleifer, and Vishny (1998) to a framework in which parameters and models are determined endogenously in equilibrium.

We demonstrate that underparameterization and misspecification equilibria can arise in this simple asset pricing model. Depending on the complicated interaction between the exogenous processes and the degree of risk-aversion of agents, multiple Misspecification Equilibria can arise as an equilibrium outcome. Adding real-time learning and dynamic predictor selection generates regime-switching dynamics in excess returns.

When the model is calibrated to U.S. stock data we find that the model is capable of capturing the regime-switching empirical features that have been extensively documented in U.S. stock data. Because of the richness of the theoretical results, and the broad empirical implications for excess returns, the approach in this paper seems to provide a suitable balance between rational expectations and fully behavioral approaches.

A Appendix

Detailed Computations for Section 1: It is straightforward to compute that

$$\begin{aligned}
E\pi_t^1 &= \frac{1}{a\sigma_1^2} \left\{ [\rho(1 + \xi_1 + \xi_2 r) - \beta^{-1}\xi_1] [(1 + \xi_1\rho - \beta^{-1}\xi_1) Ey_t^2 \right. \\
&\quad + \xi_2(\phi - \beta^{-1}) ([\rho(1 + \xi_1 + \xi_2 r) - \beta^{-1}\xi_1] - \beta^{-1} [(1 + \xi_1)\rho - \beta^{-1}\xi_1]) Ey_t z_{st} \\
&\quad \left. - \beta^{-1}\xi_2^2(\phi - \beta^{-1})Ez_{st}^2 + s_0^2 / (n/a\sigma_1^2 + (1 - n)/a\sigma_2^2)^2 \right\} \\
E\pi_t^2 &= \frac{1}{a\sigma_2^2} \left\{ (\rho - \beta^{-1}\xi_1) [(1 + \xi_1)\rho - \beta^{-1}\xi_1] Ey_t^2 \right. \\
&\quad + (\xi_2(\phi - \beta^{-1})(\rho - \beta^{-1}\xi_1) + [(1 + \xi_1)\rho - \beta^{-1}\xi_1] [\xi_2(\phi - \beta^{-1}) + \xi_1\tilde{r}\phi] Ey_t z_{st}) \\
&\quad \left. + \xi_2(\phi - \beta^{-1}) [\xi_2(\phi - \beta^{-1}) + \xi_1\tilde{r}\phi] Ez_{st}^2 + s_0^2 / (n/a\sigma_1^2 + (1 - n)/a\sigma_2^2)^2 \right\} \\
Ez_{1t}^2 &= \left(\frac{1}{a\sigma_1^2} \right)^2 \left\{ [(1 + \xi_1 + \xi_2 r)\rho - \beta^{-1}\xi_1]^2 Ey_t^2 \right. \\
&\quad \left. - 2\beta^{-1}\xi_2 [(1 + \xi_1 + \xi_2 r)\rho - \beta^{-1}\xi_1] Ey_t z_{st} + \beta^{-2}\xi_2^2 Ez_{st}^2 + s_0^2 / (n/a\sigma_1^2 + (1 - n)/a\sigma_2^2)^2 \right\} \\
Ez_{2t}^2 &= \left(\frac{1}{a\sigma_2^2} \right)^2 \left\{ (\rho - \beta^{-1}\xi_1)^2 Ey_t^2 + 2(\rho - \beta^{-1}\xi_1) [(\xi_2 + \xi_1\tilde{r})\phi - \beta^{-1}\xi_2] Ey_t z_{st} \right. \\
&\quad \left. + [(\xi_2 + \xi_1\tilde{r})\phi - \beta^{-1}\xi_2]^2 Ez_{st}^2 + s_0^2 / (n/a\sigma_1^2 + (1 - n)/a\sigma_2^2)^2 \right\}
\end{aligned}$$

Proof of Proposition 1. A unique RPE exists provided

$$\begin{bmatrix} 1 - \frac{\beta\rho n}{a\sigma_1^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} & -\frac{\beta\rho n r}{a\sigma_1^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} \\ -\frac{\beta\phi(1-n)\tilde{r}}{a\sigma_2^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} & 1 - \frac{\beta(1-n)\phi}{a\sigma_2^2 \left\{ \frac{n}{a\sigma_1^2} + \frac{1-n}{a\sigma_2^2} \right\}} \end{bmatrix}^{-1}$$

exists and real-valued solutions exist for $\sigma_j^2, j = 1, 2$ in (5)-(6). Let n take values in $\{0, 1\}$. For a sufficiently small, σ_j^2 has a real solution. It is straightforward to verify that the inverse exists under these conditions. ■

Proof of Corollary 4. Let $s_0 = 0$. As $r, \tilde{r} \rightarrow 0$ we have

$$\begin{aligned}
B_y(0) &\rightarrow (\beta^2 \rho^2)^2 \sigma_2^2 \\
B_y(1) &\rightarrow \frac{(\beta^2 \rho^2)^2 \sigma_1^2}{(1 - \beta\rho)^2} \\
B_z(0) &\rightarrow \left(\frac{\beta a \sigma_2^2}{1 - \beta\phi} \right)^2 [(2\beta\phi - 1) (\sigma_1^2 - \sigma_2^2) - \beta^2 \phi^2 \sigma_1^2] \\
B_z(1) &\rightarrow \beta^2 a^2 (\sigma_1^2)^2 (-\sigma_1^2 (1 - \beta\phi)^2 + (1 - 2\beta\phi)^2 \sigma_2^2)
\end{aligned}$$

Then straightforward algebra leads to the conditions in Corollary 4. ■

References

- [1] Ang, A., and G. Bekaert, 2002, "International Asset Allocation with Regime Switches," *Review of Financial Studies*, 15, 4, 1137-1187.
- [2] Anderson, Evan, Eric Ghysels, and Jennifer L. Juergens, 2005, "Do Heterogeneous Beliefs Matter for Asset Pricing?" *Review of Financial Studies*, 18, 3, 875-924.
- [3] Baker, Malcolm, and Jeffrey Wurgler, 2000, "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance*, 55, 5, 2219-2257.
- [4] Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 3, 307-343.
- [5] Barsky, Robert B., and Brad DeLong, 1993, "Why does the Stock Market Fluctuate?" *Quarterly Journal of Economics*, 108, 1, 291-311.
- [6] Bekaert, Geert, and Robert J. Hodrick, 1992, "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets," *Journal of Finance*, 47, 2, 467-509.
- [7] Bernard, Victor L., 1992, "Stock price reactions to earnings announcements," in R. Thaler, ed.: *Advances in Behavioral Finance*, Russell Sage Foundation, New York.
- [8] Brennan, Michael J. and Yihong Xia, 2001, "Stock Price Volatility and Equity Premium," *Journal of Monetary Economics*, 47, 2, 249-283.
- [9] Bollerslev, T., Chou, R., and K. Kroner, 1992, "ARCH Modeling in Finance: a Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 1, 5-59.
- [10] Branch, William A., 2004, "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations," *Economic Journal*, 114, 6, 592-621.
- [11] Branch, William A., 2006, "Restricted Perceptions Equilibria and Learning in Macroeconomics," in D. Colander, ed.: *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge University Press, New York.
- [12] Branch, William A., and George W. Evans, 2006a, "Intrinsic Heterogeneity in Expectation Formation," *Journal of Economic Theory*, 127, 1, 264-295.

- [13] Branch, William A., and George W. Evans, 2006b, "A Simple Recursive Forecasting Model," *Economics Letters*, 91, 2, 158-166.
- [14] Branch, William A., and George W. Evans, 2007, "Model Uncertainty and Endogenous Volatility," *Review of Economic Dynamics*, 10, 207-237.
- [15] Brock, W.A., and Cars H. Hommes, 1997, "A Rational Route to Randomness," *Econometrica*, 65, 5, 1059-1160.
- [16] Brock, W.A., and Cars H. Hommes, 1998, "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model", *Journal of Economic Dynamics and Control*, 22, 8, 1235-74.
- [17] Brock, W.A. and Cars H. Hommes, "Heterogeneous Beliefs and Routes to Complex Dynamics in Asset Pricing Models with Price Contingent Contracts," mimeo.
- [18] Cagetti, Marco, Lars Peter Hansen, Thomas J. Sargent, and Noah Williams, 2002, "Robustness and Pricing with Uncertain Growth," *Review of Financial Studies*, 15, 2, 363-404.
- [19] Campbell, John Y., and John Ammer, 1993, "What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns," *Journal of Finance*, 48, 1, 3-37.
- [20] Cho, In Koo, Noah Williams, and Thomas J. Sargent, 2003, "Escaping Nash Inflation," *Review of Economic Studies*, 69, 1, 1-40.
- [21] Cochrane, John H., 2001, *Asset Pricing*, Princeton University Press, Princeton, NJ.
- [22] Cochrane, John H., 2005, "Asset Pricing Program Review: Liquidity, Trading and Asset Prices," mimeo.
- [23] Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1991, "Speculative Dynamics," *Review of Economic Studies*, 58, 3, 529-546.
- [24] Debondt, W., and R. Thaler, 1985, "Does the Stock Market Overreact?" *Journal of Finance*, 40, 3, 793-808.
- [25] Evans, George W., and Seppo Honkapohja, 2001, *Learning and Expectations in Macroeconomics*, Princeton University Press, Princeton, NJ.
- [26] Evans, George W., and Garey Ramey, 2006, "Adaptive Expectations, Underparameterization and the Lucas Critique," *Journal of Monetary Economics*, 53, 249-264.

- [27] Fama, Eugene F., and Kenneth R. French, 1996, "Multifactor Explanations of Asset-Pricing Anomalies," *Journal of Finance*, 51, 1 426-465.
- [28] Garcia, Rene, and Pierre Perron, 1996, "An Analysis of the Real Interest Rate under Regime Shifts," *Review of Economics and Statistics*, 78, 1, 111-125.
- [29] Gaunersdorfer, Andrea, 2001, "Endogenous Fluctuations in a Simple Asset Pricing Model with Heterogeneous Agents," *Journal of Economic Dynamics and Control*, 24, 5, 799-831.
- [30] Grossman, Sanford J., and Joseph E. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70, 3, 383-408.
- [31] Guidolin, Massimo, and Allan Timmermann, 2005, "Economic Implications of Bull and Bear Regimes in UK Stock and Bond Returns," *Economic Journal*, 115, 5, 111-143.
- [32] Guidolin, Massimo, and Allan Timmermann, 2007a, "Properties of Equilibrium Asset Prices under Alternative Learning Schemes," *Journal of Economic Dynamics and Control*, 31, 1, 161-217.
- [33] Guidolin, Massimo, and Allan Timmermann, 2007b, "Asset Allocation under Multivariate Regime Switching," *Journal of Economic Dynamics and Control*, 31, 11, 3503-3544.
- [34] Guidolin, Massimo, and Allan Timmermann, 2008, "International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences," *Review of Financial Studies*, 21, 2, 889-935.
- [35] Hirshleifer, David A., and Siew Hong Teoh, 2004, "Limited Investor Attention and Stock Market Misreaction to Accounting Information," mimeo.
- [36] Harvey, Andrew C., 1990, *The Econometric Analysis of Time Series*, second edition, MIT Press, Cambridge, Mass.
- [37] Holt, Charles, and Susan Laury, 2002, "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 5, 1644-1655.
- [38] Hong, Harrison, and Jeremy C. Stein, 1999, "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *Journal of Finance*, 54,6, 2143-2184.
- [39] Hong, Harrison, Jeremy C. Stein. and Jialin Yu, 2005, "Simple Forecasts and Paradigm Shifts," forthcoming *Journal of Finance*, 62,3, 1207-1242.
- [40] Hong, Harrison, Jose Scheinkman, and Wei Xiong, 2005, "Asset Float and Speculative Bubbles," forthcoming *Journal of Finance*, 61,3, 1073-1117.

- [41] Kim, Chang-Jin, Charles R. Nelson, and Richard Startz, 1998, "Testing for Mean Reversion in Heteroskedastic Data Based on Gibbs-Sampling-Augmented Randomization," *Journal of Empirical Finance*, 5, 1, 131-154.
- [42] Lamont, Owen, and Richard Thaler, 2003, "Can the Market Add and Subtract? Mispricing in Tech Stock Carve-outs," *Journal of Political Economy*, 111, 1, 227-268.
- [43] Lansing, Kevin J., 2006, "Lock-in of Extrapolative Expectations in an Asset Pricing Model," *Macroeconomic Dynamics*, 10, 2, 317-348.
- [44] Lettau, Martin, and Sydney C. Ludvigson, 2005, "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76, 3, 583-626.
- [45] Lewellen, Jonathan, and Jay Shanken, 2002, "Learning, Asset Pricing Tests, and Market Efficiency," *Journal of Finance*, 57, 3, 1113-1145.
- [46] Manski, C.F., and D. McFadden, 1981, *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, MA.
- [47] Marcet, Albert, and Thomas J. Sargent, 1989, "Convergence of Least-Squares Learning Mechanisms in Self-Referential Linear Stochastic Models," *Journal of Economic Theory*, 48, 2, 337-368.
- [48] Merton, Robert C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, 42, 2, 483-510.
- [49] Nelson, Charles, 1972, "The Prediction Performance of the FRB-MIT-PENN Model of the US Economy," *American Economic Review*, 62, 902-917.
- [50] Ofek, Eli, and Matthew Richardson, 2003, "Dotcom Mania: The Rise and Fall of Internet Stock Prices," *Journal of Finance*, 58, 3, 1113-1137.
- [51] Pastor, Lubos, and Pietro Veronesi, 2003, "Stock Valuation and Learning about Profitability," *Journal of Finance*, 58, 4, 1749-1789.
- [52] Peng, Lin, and Wei Xiong, 2006, "Investor Attention, Overconfidence and Category Learning," *Journal of Financial Economics*, 80, 3, 563-601.
- [53] Potter, Simon, 1995, "A Nonlinear Approach to U.S. GNP," *Journal of Applied Econometrics*, 10, 2, 109-125.
- [54] Perez-Quiros, Gabriela, and Allan Timmermann, 2000, "Firm Size and Cyclical Variation in Stock Returns," *Journal of Finance*, 55, 3, 1229-1262.
- [55] Sargent, Thomas J., 1999, *The Conquest of American Inflation*, Princeton University Press.

- [56] Timmermann, Allan, 1994, "Can Agents Learn to Form Rational Expectations? Some Results on Convergence and Stability of Learning in the UK Stock Market," *Economic Journal*, 104, 4, 777-798.
- [57] Timmermann, Allan, 1996, "Excess Volatility and Predictability of Stock Returns in Autoregressive Dividend Models with Learning," *Review of Economic Studies*, 523-557.
- [58] Turner, C., Startz, C., and C. Nelson, 1989, "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market," *Journal of Financial Economics*, 25, 1, 3-22.
- [59] Williams, Noah, 2004, "Escape Dynamics in Learning Models," mimeo., Princeton University.
- [60] Young, Peyton H., 2004, *Strategic Learning and its Limits*, Oxford University Press.