Natural Expectations, Macroeconomic Dynamics, and Asset Pricing

Andreas Fuster, Benjamin Hebert, David Laibson*

April 3, 2011

Abstract

How does an economy behave if (1) fundamentals are truly hump-shaped, exhibiting momentum in the short run and partial mean reversion in the long run and (2) agents do not know that fundamentals are hump-shaped and base their beliefs on parsimonious models that they fit to the data? A class of parsimonious models engenders qualitatively similar biases and generate empirically observed patterns in asset prices and macroeconomic dynamics. First, parsimonious models will robustly pick up the short-term momentum in fundamentals but will generally fail to fully capture the long-run mean reversion. Beliefs will therefore be characterized by endogenous extrapolation bias and pro-cyclical excess optimism. Second, asset prices will be highly volatile and exhibit partial mean reversion—i.e., overreaction. Excess returns will be negatively predicted by lagged excess returns, P/E ratios, and consumption growth. Third, real economic activity will have amplified cycles. For example, consumption growth will be negatively auto-correlated in the medium run. Fourth, the equity premium will be large. Agents will perceive that equities are very risky when in fact long-run equity returns will covary only weakly with long-run consumption growth. If agents had rational expectations, the equity premium would be close to zero. Fifth, sophisticated agents—i.e., those who are assumed to know the true model—will hold far more equity than investors who use parsimonious models. Moreover, sophisticated agents will follow a counter-cyclical asset allocation policy. These predicted effects are qualitatively confirmed in U.S. data.

*Harvard University, Department of Economics. We are grateful to Nicholas Barberis, John Beshears, Markus Brunnermeier, John Campbell, James Choi, Emmanuel Farhi, Kenneth French, Xavier Gabaix, Stefano Giglio, Blake LeBaron, Greg Mankiw, Joshua Schwartzstein, Andrei Shleifer, Jeremy Stein, Jim Stock, Michael Woodford, and seminar/conference participants for helpful comments and discussions. We are indebted to Brendan Price for excellent research assistance.

1 Introduction

Since the 1970's, most macroeconomic models have incorporated the assumption that people have rational expectations; i.e., people know the true model of the economy.

In the current paper, we follow a different tradition and assume that agents use simple prediction models, which are estimated using historical data.[1] In other words, *agents adopt a parsimonious model that fits the available data*. In general, this parsimonious model will not nest the true model (though the true model may nest the parsimonious model). We call the resulting beliefs *natural expectations*. We assume that agents use simplified models because this is how, in practice, economists and non-economists—statisticians, professional forecasters, and firms—actually do build models.

People use simple models for a wide range of good reasons. Simple models are easier to understand, easier to explain, and easier to employ. Simplicity also reduces the risks of overfitting, which is the reasoning that underlies many formal model selection criteria. Whatever the mix of reasons—pragmatic, psychological/suboptimal, and statistical—economic agents usually *do* use simple models to understand economic dynamics. We show that a class of parsimonious models engenders qualitatively similar biases and generate empirically observed patterns in asset prices and macroeconomic dynamics.

To illustrate this claim, we study an economy in which fundamentals are hump-shaped, exhibiting short-run momentum and long-run partial mean reversion. Hump-shaped dynamics are controversial in the sense that economists continue to debate whether such cyclical dynamics are present in aggregate fluctuations. This debate is consistent with our claims, since we only want to argue that hump-shaped dynamics are plausible. In fact, hump-shaped dynamics match the point estimates from ARIMA models of various economic time series[2], though standard errors are large enough that the data do not rule out dynamics that are far

[1] For example, Barberis, Shleifer, and Vishny (1998), Hong, Stein and Yu (2007), and Branch and Evans (2010) study settings in which agents estimate a misspecified model and optimize against that model.

[2] Fuster, Laibson and Mendel (2010) provide some empirical evidence for hump-shapes in different economic time series. That paper also presents a simplified version of the model in this paper.

more persistent.[3]

We will ask, "How would an economy behave if (1) fundamentals were truly hump-shaped, and (2) agents adopted a parsimonious model of the fundamental process, fit to the available data?"[4] We embed these two assumptions—hump-shaped fundamentals and natural expectations—in a consumption/asset-pricing model. We use a habit model to generate slow adjustment in consumption; however, as we combine the habit with CARA preferences (Alessie and Lusardi 1997), these habits do not produce counter-cyclical variation in risk premia.[5]

The following five sets of results emerge from our analysis. Some of these results are comparative: how does equilibrium behavior in an economy with natural expectations compare to equilibrium behavior that would have arisen if agents knew the true model—i.e., if agents had rational expectations?

First, simple models robustly pick up the short-term momentum in fundamentals but often fail to capture the full extent of long-run mean reversion. Consequently, under natural expectations beliefs will often be characterized by endogenous extrapolation bias in levels: forecasts about fundamentals will be too persistent since simple models tend to miss some of the long-run mean reversion. This implies that beliefs will be too optimistic in good times and too pessimistic in bad times relative to the rational expectations benchmark.

Second, under natural expectations asset prices will be highly volatile (LeRoy and Porter 1981, Shiller 1981) and exhibit partial mean reversion—i.e., overreaction. Excess returns will be negatively predicted by lagged excess returns (Fama and French 1988a, Poterba and Summers 1988), price/earnings ratios (Campbell and Shiller 1988a, 2005), and consumption growth. Excess returns will be positively predicted by lagged 'cay', a measure of transitory

[3]There is considerable debate about the cyclicality of economic activity. See Campbell and Mankiw (1987), Cochrane (1988), Morley, Nelson, and Zivot (2003) and Perron and Wada (2009) for a range of opinions.

[4]To keep the analysis in our paper maximally tractable, we abstract away from learning and give agents a fixed simple model that is estimated from data available in 2011. Adding learning would complicate the analysis, but not change the qualitative results.

[5]In contrast, Campbell and Cochrane (1999) use a habit model that generates counter-cyclical risk premia, thereby explaining numerous asset pricing regularities.

deviations of consumption from wealth (Lettau and Ludvigson 2001).[6]

Third, real economic activity will have amplified cycles. For example, consumption growth will be positively auto-correlated in the short run and *negatively* auto-correlated in the medium run. Consumption growth will be weakly negatively predicted by lagged excess returns and P/E ratios.

Fourth, the equity premium will be large. Agents will perceive that equities are very risky when in fact long-run equity returns will covary only weakly with long-run consumption growth. The covariance of consumption growth and asset returns will be close to zero over short-run horizons (because of slow adjustment in consumption), it will be higher over medium-run horizons (as consumption catches up with asset prices), and it will fall again over long-run horizons (as asset prices and consumption both mean revert). If agents had rational expectations, the equity premium in our economy would be close to zero.

Fifth, sophisticated agents—i.e., those who are assumed to know the true model—will hold far more equity than investors with natural expectations. Moreover, sophisticated agents will be "value" investors, following a counter-cyclical investment policy.

These five sets of predicted effects are qualitatively confirmed in U.S. data. An economy in which agents estimate simple models—e.g., an AR(10) in earnings growth, instead of an AR(40)—generates simulated behavior that quantitatively matches the point estimates observed in U.S. data. However, the moments that we study have large standard errors because of the limited span of available data. It is therefore not possible to reject classical models in which excess returns are unpredictable and consumption is consistent with rational expectations.

The body of the paper is divided into six sections; appendices are used for derivations. Section 2 discusses the econometric and psychological motivations for natural expectations, as well as the related literature. Section 3 solves and calibrates a consumption-based asset pricing model. Section 4 reports model simulations and compares these simulations to the

[6]When 'cay' is low, wealth is transitorily high. In our economy, this predicts low future excess returns.

empirical evidence from U.S. We focus on eight moments that summarize the key properties of the model. We show that a parsimonious version of the model matches these moments. Section 5 discusses the behavior of sophisticated agents. Section 6 concludes and identifies directions for future work.

2 The Appeal of Simple Models

The premise of our approach to understanding macroeconomic and financial dynamics is that economic agents tend to make forecasts based on statistical or mental models that are reasonable given the data available to them, but "too simple" to fully capture the long-term dynamics of many economic time series.

In this section, we motivate our assumptions both on statistical and psychological grounds, though we believe that the psychological motivations are more important.[7] We will also discuss how simple models lead agents to overestimate the persistence of shocks when the true process is hump-shaped. Finally, we summarize evidence from different settings that are related to our approach.

2.1 Statistical Perspective

Choosing the right model to forecast an economic time series is by no means a trivial task, and there is no single generally accepted way of doing so. When choosing how many parameters to include, a modeler faces a trade-off between improving the fit of the model in-sample and the risk of overfitting the available data, which may result in poor out-of-sample forecasts. A number of formal statistical criteria have been proposed to formalize the trade-off between fit and parsimony. The best-known and most popular ones are the Akaike Information

[7]Brav and Heaton (2002) emphasize that when explaining "financial anomalies," behavioral theories and rational theories with structural uncertainty are often very similar mathematically and also in terms of predictions. We remain relatively agnostic as to whether the reliance on models that fail to capture long-term dynamics is "behavioral" or "rational" in our setting, and focus on the implications.

Criterion (AIC) and the Schwarz or Bayesian Information Criterion (BIC).[8] These criteria are both asymptotically optimal, but in different ways. If the set of candidate models is not presumed to contain the true model (for instance because the true model is of infinite dimension), the AIC is *efficient* in that it will select the candidate model with minimum mean squared error distribution with probability 1 as the sample size grows to infinity. If instead the true model is among the set of candidate models, the BIC will select it with probability 1 asymptotically (the BIC is *consistent*), while the AIC will tend to result in an overparametrized model (Hannan 1980).

When the likelihood function is Gaussian, the two criteria can be written as follows:

$$\text{AIC} = \ln(\hat{\sigma}^2) + \frac{2k}{T}$$
$$\text{BIC} = \ln(\hat{\sigma}^2) + \frac{\ln(T) \cdot k}{T}$$

where $\hat{\sigma}^2$ is the error variance estimated by maximum likelihood, k the number of parameters (including the constant), and T the number of observations. One is supposed to pick the model with the smallest value of the preferred criterion. Both criteria trade off fit and parsimony: as k increases, $\hat{\sigma}^2$ decreases, but the second ("penalty") term increases. Generally, the BIC imposes a larger penalty for increasing the number of parameters, and thus will tend to select models with fewer parameters than the AIC. A relatively recent literature based on extensive simulation studies has argued that a version of the AIC corrected for small sample sizes, the AICc, tends to perform well (in various senses) and, relative to the AIC, reduces the likelihood of overfitting.[9]

It is not clear which criterion should be preferred. If one believes that the true model is among the candidate models, and if $T \to \infty$, there seems to be a clear case for the BIC. Yet, in practice, these conditions are rarely met, and it appears that which criterion is preferred

[8] McQuarrie and Tsai (1998) provide an overview of these and other selection criteria.
[9] The AICc can be written as $\ln(\hat{\sigma}^2) + \frac{T+k-1}{T-k-1}$.

seems to be to some extent a matter of taste.[10] While strictly speaking, each criteria picks one model, a reasonable pragmatic strategy would be to choose a model that is close to the best model according to different criteria, or perhaps to make predictions based on averaging the predictions of different models.

For the purpose of our work, what matters is whether a modeler who is presented with a time series of length equal to the typical macroeconomic time series (e.g., 250 quarters) and who relies on (one of) these information criteria will generally pick a model that correctly captures the properties of the data-generating process. In particular, we are interested in time series that feature hump-shaped dynamics, in the sense that they are characterized by momentum (positive autocorrelation in growth rates) in the short run but (partially) revert back to the mean in the long run. To our knowledge, this is not a question that has been formally analyzed in the existing literature, so we now present some suggestive evidence from Monte Carlo simulations to shed light on this question.

Assume that the true data generating process is an ARIMA(0,1,16), and for simplicity assume that the modeler considers univariate models of the ARIMA$(p, 1, 0)$ and ARIMA$(0, 1, q)$ classes. The process we are interested in has a hump-shaped impulse response function with a long-term persistence of 0.5.[11] We draw 100 samples of length 255 (which is the number of observations we will use in the estimation of the aggregate earnings process) and estimate AR$(p, 1, 0)$ models for $p = 1, ..., 25$ and MA$(0, 1, q)$ models for $q = 4, 8, 12, 16, 20, 24$.[12]

[10] McQuarrie and Tsai (1998) indicate that "AIC is probably the most commonly used model selection criterion for time series data" (p. 2). On the other hand, Neath and Cavanaugh (1997) note that BIC "is often preferred over AIC by practitioners who find appeal in either its Bayesian justification or its tendency to choose more parsimonious models than AIC" (p. 559) For an interesting discussion on this topic from an online forum for "statisticians, data analysts, data miners and data visualization experts," see http://stats.stackexchange.com/questions/577/is-there-any-reason-to-prefer-the-aic-or-bic-over-the-other.

[11] The MA coefficients are: +0.15, +0.1, -0.02, -0.05, -0.05, -0.05, -0.05, -0.05, -0.06, -0.06, -0.06, -0.06, -0.06, -0.06, -0.06, -0.06. The long-term persistence of an ARIMA$(p, 1, q)$ process is given by (1+sum of MA coefficients)/(1-sum of AR coefficients).

[12] Estimating MA models with a large number of terms is computationally demanding, which is why we only estimate a subset of the possible models one might consider. In our simulations the maximum likelihood estimation of MA models sometimes failed to converge, or the maximum likelihood estimator of the root (i.e., the negative of the sum of the lagged MA coefficients) "piled up" at a value of 1, which means that the predicted long-term persistence of a shock equaled 0. This pileup occurs because the sample likelihood function is locally flat at an MA root of 1, so that it is a local maximum of the likelihood function and may be the global maximum in finite samples, even if the true MA root is less than unity (see Campbell and

In our 100 simulations, the AICc selects an AR(1), AR(2), or AR(3) in 60 cases; the BIC does so in all cases. The average estimated long-term persistence of a shock implied by the AICc-selected model equals 1.02, and is below 0.8 only in 25% of cases. For the BIC-selected model, it is never below 1. Thus, in the vast majority of cases, a modeler who considers this set of candidate models and selects based on AICc or BIC would grossly overestimate the persistence of shocks to the process. Note that while it appears difficult to capture the mean-reversion in a sample of this length, it is not impossible: for instance, the implied persistence from an estimated MA(16) (the correct model) averages 0.56 and is below 0.8 in 88% of cases.[13] Yet the improvement that the MA(16) brings in terms of fit relative to an AR(2), say, is only rarely worth its increased complexity here (an MA(q) with $q \geq 16$ is selected nine times by the AICc).

In large part, the difficulty that these models have in detecting mean reversion seems due to the relatively short sample length. If we repeat the same exercise but with 1255 observations instead of 255 (the equivalent of an additional 250 years of data, if this series were quarterly), the AICc selects the MA(16) in 75% of cases, and the mean long-term persistence of a shock estimated by the preferred model decreases to 0.53, i.e. very close to the actual persistence of the process.

We have conducted similar simulations with other data generating processes. While with somewhat less complicated hump-shape patterns, the AICc often does well in terms of selecting a model that gets the long-term persistence approximately right, the BIC with its greater emphasis on parsimony tends to select low-order models (with 5 coefficients or less) in the vast majority of cases. Thus, overall our simulations suggest that even a sophisticated modeler who goes through the trouble of estimating and comparing a wide selection of candidate models may well end up with a model that vastly overestimates the long-term

Mankiw 1987 or Stock 1994 for discussions). It seems likely that a modeler would be highly skeptical of such a result. Also, our modeler does not consider ARIMA($p, 1, q$) models with both p and $q > 0$. While such models are often estimated in practice, based on our results, it seems very unlikely that estimating more models would change our qualitative conclusions.

[13] For these calculations, we drop simulations in which the sum of MA coefficients equalled exactly -1 (cf. previous footnote).

persistence of shocks. To be clear, we are not claiming that it is ever impossible to find a model that gets the long-term dynamics right, in particular if one uses more sophisticated methods than ARIMAs (e.g., multivariate models) or adds more ex-ante imposed structure to the model—we are saying that it is not straightforward.

So far, we have discussed the model selection problem under the assumption that the stochastic process of interest is stable over time and that economic agents should normatively use all the historical data that is available to them in estimating a model, and weigh all observations equally. However, it is possible that the properties of the data generating process change over time, and so it may well be optimal for an agent to downweight or discard old data.[14] While we will not attempt a normative analysis of optimal model selection under the possibility that parameters of the data generating process (such as the mean growth rate of earnings) change over time, it seems intuitive that in such a world, agents would tend to use shorter effective samples to estimate their models, and this would increase the likelihood of picking a low-order model.[15]

2.2 Psychological Perspective

Psychological factors also lead agents to adopt simple models. In fact, we believe that psychological considerations are far more important in driving preferences for simplicity than the statistical considerations reviewed above. Even *if* non-parsimonious, complicated models were statistically optimal, real people would probably not adopt such models.

Constraints on memory and cognition make it difficult for agents to work with complicated models, leading decision-makers to adopt simplifications and heuristics (e.g., Tversky and Kahneman 1973, 1974; Gigerenzer and Goldstein 1996; Gabaix, Laibson, Moloche and Weinberg 2006; Gabaix 2011). Simple models are relatively easy to comprehend, use, revise, and explain. Complicated "black-box" models are viewed with a degree of suspicion;

[14] A large literature in statistics and economics studies methods to detect regime switching or structural breaks.

[15] Model selection criteria like the AIC, AICc, and BIC all *increase* the penalty for model complexity as the sample size falls.

programmers themselves are boundedly rational so high-dimensional computational models often have subtle programming errors and other unintended or poorly understood features that make them unreliable. Even when people do use complicated computational models, the conclusions are rarely taken at face value and usually tempered with "common sense" about how the world works.

The 'psychological' motives for simplicity can also be interpreted through a rational lens. If complicated models tend to induce confusion or mistakes, then simple models may be the lesser of two evils. Complicated models also engender high costs for agents with costly cognition[16], providing another rational explanation for complexity avoidance.

There are also some specific psychological biases that reinforce our approach. The heuristic of representativeness (Kahneman and Tversky 1973; Tverksy and Kahneman 1974) leads people to believe that small samples are representative of the world at large. Representativeness has two implications for our analysis.

First, recent observations are viewed as representative of the future. Thus representativeness leads agents to underestimate the degree of mean reversion (Kahneman and Tversky 1973).

Second, representativeness leads agents to mistakenly believe that the properties of population samples will be reliably observed in small samples (e.g., Rabin 2002; Rabin and Vayanos 2010). This is sometimes referred to as the (psychological) 'law of small numbers.' This error increases agents' confidence in estimates derived from small samples. Agents mistakenly believe that the ergodic properties of a time series can be inferred by studying only a short sub-sample (e.g., 20 years of data). In addition, a willingness to rely on short sub-samples implicitly reinforces the tendency to rely on models without long lag effects.

The availability heuristic (Tversky and Kahneman 1973) also reinforces our modeling approach. Availability leads people to overweight information that is easily accessible and salient. Hence, availability bias also implies that people will overweight recent data and

[16]For example, see Sims (1998, 2003), Gabaix et al. (2006), Woodford (2009).

underweight data from the distant past.

Some observors have argued that related biases play an important role in driving aggregate dynamics. For instance, Reinhart and Rogoff (2009) document how investors time and time again fall prey to the belief that "this time is different" and that this causes recurrent financial crises. Relatedly, Shiller (2005) points out the lure of "new era" stories and how they are associated with episodes of bubbles in asset markets. Barberis (2010) notes that over-extrapolation of past prices changes may have been an important psychological driving force during the run-up to the Great Recession. We now turn to a large body of related economic research.

2.3 Additional Related Literature

There is a small but growing body of evidence on deviations from rational expectations. For instance, lab experiments in which subjects are asked to forecast financial or other time series find that extrapolative expectations or "trend following" provide a good description of observed beliefs (De Bondt 1993; Hey 1994) and may be a driving force behind the bubbles that are observed in asset-market experiments (Haruvy, Lahav, and Noussair 2007; Hommes et al. 2008). On the other hand, Dwyer et al. (1993) finds that subjects' forecasts of a random walk (in which growth has no persistence) do not deviate systematically from the rational expectations forecast. This is consistent with our model, where extrapolation is not "baked in" but depends on the predictions generated by estimating simple models on the available data.

In field data, a number of papers have argued that asset allocation choices are affected by extrapolation of recent price appreciation (Chevalier and Ellison 1997; Sirri and Tufano 1998; Benartzi 2001; Choi et al 2004, 2009; Benartzi and Thaler 2007; Chalmers and Reuter 2009; Previtero 2010; Malmendier and Nagel 2011). One could argue that biases in expectations do not matter for asset pricing because investors with biased beliefs hold relatively little wealth. However, Vissing-Jorgensen (2003) shows that at the peak of the market in 2000-01,

even wealthy investors expected stock returns to remain high. Bacchetta, Mertens, and van Wincoop (2009) conduct a similar exercise and find that, in several asset markets, investors' expectational errors about future returns are predicted by the same variables that predict excess returns.[17]

One might alternatively think that the expectations held (and made public) by financial analysts are not biased. However, De Bondt and Thaler (1990) argue that security analysts overreact and make earnings-per-share forecasts that are too extreme.[18] Most of the studies in this literature look at relatively short-run forecasts, while our model mostly has implications for long-run forecasts. Bulkley and Harris (1997) study five-years earnings forecasts for about 500 U.S. companies and report results which are consistent with our model: i) analysts appear to extrapolate past growth in earnings when forecasting future growth, even though there is pronounced negative serial correlation in earnings growth over five-year periods, and ii) analysts' forecasts and excess returns over the subsequent five years are significantly negatively correlated.[19] Chan, Karceski, and Lakonishok (2003) provide further evidence that there is little predictability of long-term earnings growth rates, but that investors and analysts behave as if recent growth rates were positive predictors of future growth.

A significant literature in behavioral finance has accumulated evidence on cross-sectional stock return patterns that are consistent with such biases in expectations having strong effects on prices: De Bondt and Thaler (1985, 1989) and Lakonishok, Shleifer, and Vishny (1994) are among the best-known examples of such work.[20] Baker and Wurgler (2007) doc-

[17]For instance, a high dividend/price ratio is a strong predictor of high subsequent excess returns. Similarly, when 'cay' (Lettau and Ludvigson 2001) is high, excess returns also tend to be higher than anticipated. However, surveyed investors tend to expect low excess returns when 'cay' is high.

[18]Other studies instead find that analysts underreact. Easterwood and Nutt (1999) argue that analysts overreact to positive information but underreact to negative information. Lim (2001) argues that considering analysts' objective function can "rationalize" their biases.

[19]La Porta (1996) finds a negative relation between analysts' long-term growth estimates and future one-year risk-adjusted returns. Bergman and Roychowdhury (2008) document a positive relation between the consumer confidence index (a proxy for market sentiment) and the error in long-horizon earnings estimates of financial analysts, consistent with the idea that when times are good, market participants may insufficiently adjust for subsequent mean reversion.

[20]More recently, Chen, Moise and Zhao (2009) argue that myopic extrapolation can also explain momentum, if investors completely miss the hump-shaped dynamics of firm-specific earnings shocks and simply treat current earnings shocks as permanent. They point out that apart from cognitive biases, the practice

ument that empirical measures of investor sentiment predict cross-sectional return patterns and also aggregate returns.

Apart from stock markets, other asset markets may also be influenced by biased beliefs. For instance, Greenwood and Hanson (2010) document patterns in bond risk premia that can be explained by investors extrapolating recent returns or default rates. Periods of high returns on corporate bonds are followed by a decline in issuer quality and low or negative excess returns on corporate debt in a highly predictable manner. Also, biased (extrapolative) beliefs have been advanced as a key explanation behind the recent housing bubble as well as earlier boom-bust cycles (Abraham and Hendershott 1996, Muellbauer and Murphy 1997, Case and Shiller 2003, Gerardi et al. 2008, Goetzman, Peng and Yen 2009, Piazzesi and Schneider 2009, Glaeser, Gottlieb and Gyourko 2010).

A variety of "behavioral" models have been proposed to explain stock return patterns, including DeLong et al. (1990), Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer and Subrahmanyam (1998) and Hong and Stein (1999). Closely related are models in which investors continuously update their belief about future dividend growth or other parameters. This learning, which can be interpreted as behavioral or fully rational (similar to our model), generates predictability in returns as well as excess volatiliy. Among the best known papers in this literature are Barsky and DeLong (1993) and Timmermann (1993). While our approach is closely related to these earlier papers, most previous authors consider simpler setups (often partial equiliibrium valuation models without consumption) to illustrate the consequences of biased beliefs or learning, and do not study the interrelation between asset prices and other macroeconomic variables. An exception to this, and similar in spirit to our exercise, is for instance a paper by Cecchetti, Lam, and Mark (2000). They show that distorted beliefs about the growth rate of the aggregate endowment can generate high, volatile, and predictable excess returns on equity, as in the data.[21] Similarly, Lansing (2006), Choi (2006) and Adam

to price securities using earnings multiples can also contribute to this phenomenon.

[21]The type of distortion that Cecchetti et al. focus on has agents *underestimate* the persistence of good and bad shocks to endowment growth (see Gourinchas and Tornell 2004 for a related model in an international finance context), while our "natural expectations" agents will *overestimate* the persistence of shocks to

and Marcet (2010) study economies where (all or a subset of) investors hold extrapolative beliefs, or learn about the return process, in otherwise standard Lucas tree economies with CRRA utility, and show that this can generate realistic asset pricing dynamics.

More generally, a large literature in macroeconomics and finance, surveyed in Sargent (1993), Evans and Honkapohja (2001, 2009) and Pastor and Veronesi (2009) assumes that agents are rational (in the sense that they think like highly sophisticated statisticians) but have to learn the relevant parameters of the reduced form equations governing the economy over time.[22] While the early papers in this literature mostly focused on whether expectations would ultimately converge to the rational expectations equilibrium, more recent work has considered what happens if agents have misspecified models and/or downweight older data, and finds that this can generate additional volatility and persistence of shocks in asset prices and/or the economy (e.g. Branch and Evans 2007, 2010; Hong, Stein, and Yu 2007; Huang, Liu and Zha 2009; Eusepi and Preston forthcoming). In these models, "misspecification" means that agents omit a relevant variable from their forecasting equation, while in our model, it means that they may not include enough lags of the variable they are trying to forecast.[23]

An alternative modeling approach, exemplified by the agent-based literature, assumes that agents (probabilistically) select among different forecasting models based on past performance of the models (for example, LeBaron et al. 1999; Tesfatsion and Judd 2006; De Grauwe 2010; LeBaron 2010). An advantage of such models, which are usually analyzed computationally, is that they generate heterogeneity in beliefs, which our model omits. Such heterogeneity allows for instance the study of wealth dynamics and trading volume.[24]

Finally, there is of course a very large literature that studies asset pricing and macroeconomic dynamics under the assumption of perfect rationality and knowledge of the economy's

earnings growth.

[22]For an early example, see Friedman (1979).

[23]The downweighting of old data is often captured by assuming "constant gain" rather than "decreasing gain" (least squares) learning. Some papers, such as Marcet and Nicolini (2003), endogenize agents' choice between constant and decreasing gain based on recent prediction errors.

[24]See Hong and Stein (2007) for a discussion of models of disagreement in a finance context.

structure.[25] While basic rational models are rather unsuccessful in generating realistic asset pricing patterns, more elaborate versions can generate a high equity premium, volatile asset prices, and predictability of excess return. Perhaps the most successful and influential consumption-based models over the past 15 years are a) Campbell and Cochrane's (1999) habit model, which matches many of the main observed empirical asset pricing phenomena through counter-cyclical and (on average) high effective risk aversion, and b) the "long-run risks" model by Bansal and Yaron (2004), who engage in an exercise somewhat similar to ours: they assume that the world is characterized by a driving process for which evidence is statistically rather weak (namely, a consumption growth process with a small predictable component as well as time-varying volatility) and study the implications for asset pricing (in their case, using Epstein-Zin-Weil preferences with an elasticity of intertemporal substitution above 1).

3 Consumption Model with Asset Pricing

We now illustrate our approach by characterizing equilibrium in an economy in which agents hold natural expectations.

We study an open endowment economy with two assets. The first asset is foreign debt, b_t, which is borrowed at a fixed international (gross) interest rate R. We introduce foreign debt since we do not want a mechanistic linkage between production and consumption. The second asset is a domestic equity tree, which generates a dividend, d_t, that is stationary in first differences. In other words, Δd_t is an AR(n),

$$\Delta d_t = \rho_1 \Delta d_{t-1} + ... + \rho_n \Delta d_{t-n} + \sigma \varepsilon_t$$

where ε_t is an iid Gaussian shock with unit variance.[26] We assume that the equity tree

[25] See Campbell (2003) and Cochrane (2007) for surveys of the rational asset pricing literature.

[26] Without loss of generality, we assume that the process has no deterministic drift. Adding deterministic would not change any of our results on comovement.

must be held by domestic agents.[27] We will distinguish between the "true" data generating process for dividends, and the perceived data generating process for dividends. When the true data generating process matches the perceived data generating process, agents hold rational expectations. Our focus, however, is on cases in which the perceived data generating process has fewer lags than the true data generating processes. As discussed in Section 2, this can occur for various reasons, and we will refer to this as "natural expectations."

Our timing conventions and wealth definitions follow. Period t is divided into the following sequential subperiods. We describe the model using a "quarterly" frequency—since this is the calibration that we will use—but the model can be calibrated for any period of observation.

1. On the first day of the quarter – "January 1" – households start with debt, b_t, and θ_{t-1} units of claims to the risky asset.

2. Time passes and production occurs from January 1 to March 31. Households make no additional choices until the end of the quarter: March 31.

3. At the end of the quarter – March 31 – the dividend of the equity tree is realized and paid to households: d_t per unit of claim. So households receive $\theta_{t-1}d_t$ units of output.

4. The tree is priced (ex-dividend): p_t per unit of claim (where output units are the numeraire). So households can sell the tree for $\theta_{t-1}p_t$ units of output.

5. Wealth is measured at the end of the quarter,

$$\begin{aligned} w_t &= -Rb_t + \theta_{t-1}d_t + \theta_{t-1}p_t \\ &= (w_{t-1} - c_{t-1} - \theta_{t-1}p_{t-1})\,R + \theta_{t-1}\,(d_t + p_t)\,. \end{aligned}$$

So wealth, w_t, is measured after production has occured in the period but **before**

[27] If the equity tree were held by foreign investors, this would drastically dilute the associated risk. Domestic ownership can be motivated by home bias.

consumption is chosen.

6. Consumption is chosen: c_t.

7. Asset allocation takes place: agent buys θ_t units of equity at price p_t.

8. End-of-quarter debt – after these transactions – is

$$b_{t+1} = w_t - c_t - \theta_t p_t.$$

9. Households start the next quarter (beginning on April 1, which is period $t+1$), with this level of debt.

Period $t+1$ continues and the cycle of subperiods restarts.

3.1 Preferences and the Bellman Equation

We have two desiderata for preferences. We want preferences that generate a closed form solution. We also want preferences—or technology—that will generate slow aggregate adjustment in consumption *without* assuming procyclical risk tolerance. For evidence on the slow adjustment of consumption when responding to wealth shocks, see Dynan and Maki (2001), Gabaix and Laibson (2001), Parker (2001), and Carroll, Sommer and Slacalek (forthcoming).

Motivated by these goals, we adopt exponential preferences (e.g., Caballero 1990) with habits, as studied by Alessie and Lusardi (1997):

$$u(c_t, c_{t-1}) = -\frac{1}{\alpha} \exp\left(-\alpha \left[c_t - \gamma c_{t-1}\right]\right)$$

The parameter $\gamma \in [0,1]$, reflects the strength of the habit. As we show below, in this formulation habits only serve to slow down consumption adjustments. If we did not include habits, our model would imply counterfactually rapid adjustment in consumption. Our

habits are not operating as in Campbell and Cochrane (1999), where the habit is constructed in a way that generates counter-cyclical relative risk aversion.

Utility flows are weighted with discount factor δ, so that lifetime utility is given by

$$\sum_{s=0}^{\infty} \delta^s u(c_{t+s}, c_{t+s-1}).$$

Lifetime utility is maximized with respect to the dynamic budgets constraints summarized above. Here is the decentralized Bellman Equation, which includes the state variables that we have already introduced as well as the vector of historical dividends, $\overrightarrow{d}_t$.

$$V\left(c_{t-1}, w_t, p_t, \overrightarrow{d}_t\right) = \sup_{\theta_t, c_t} u(c_t, c_{t-1}) + E_t \delta V\left(c_t,\ (w_t - c_t - \theta_t p_t) R + \theta_t (d_{t+1} + p_{t+1}),\ p_{t+1},\ \overrightarrow{d}_{t+1}\right)$$

since

$$w_{t+1} = (w_t - c_t - \theta_t p_t) R + \theta_t (d_{t+1} + p_{t+1}).$$

For now, we will not pin down the conditional expectation operator, E_t. It depends only on the perceived data generating process for dividends. We will study the predictions of the model under rational expectations and a large set of specifications for natural expectations.

The planner's Bellman Equation (which eliminates the asset allocation issue and eliminates price-based wealth measurement) is the following:

$$V\left(c_{t-1}, b_t, \overrightarrow{d}_t\right) = \sup_{c_t} u(c_t, c_{t-1}) + E_t \delta V\left(c_t,\ (c_t + R b_t - d_t),\ \overrightarrow{d}_{t+1}\right).$$

since

$$b_{t+1} = c_t + R b_t - d_t.$$

3.2 Value Functions, Policy Functions, and Asset Pricing

We first study the representative household's problem, since we here assume that the economy is populated by homogeneous households. In Section 5, we introduce a zero measure of rational expectations agents and study their policy functions.

For the representative agent's problem, we can ignore the asset allocation decision as well as asset pricing. Appendix A shows that

$$c_t = \frac{\gamma}{R} c_{t-1} + \left(1 - \frac{\gamma}{R}\right) x_t - \psi. \tag{1}$$

This implies that consumption is a weighted average of lagged consumption and the (risk-neutral) annuity value of perceived future dividends, x_t, down-shifted by an additive constant ψ (a precautionary savings effect).

$$x_t = \frac{R-1}{R}\left[-Rb_t + \sum_{s=0}^{\infty} \frac{E_t d_{t+s}}{R^s}\right] \tag{2}$$

$$\psi = \frac{1}{R-1}\left[\frac{1}{\alpha}\ln(R\delta) + \frac{\alpha}{2} Var_t(\Delta c_{t+1})\right]. \tag{3}$$

For the planner, the value function takes a simple form:

$$V\left(c_{t-1}, b_t, \overrightarrow{d}_t\right) = \frac{-1}{\alpha(1-\delta)} \exp\left(-\alpha\left[c_t - \gamma c_{t-1}\right]\right) \tag{4}$$

where c_t is given by equation (1).

To calculate the equilibrium price of the Lucas tree, we consider the asset allocation problem (as opposed to the planner's problem). We then solve for the asset price that leads the representative agent to hold one unit of the equity tree. These calculations are provided in Appendix B, where we show that the equilibrium price of the Lucas tree is given by:

$$p_t = \sum_{s=1}^{\infty} \frac{E_t d_{t+s}}{R^s} - \frac{R\alpha \times Var_t(\Delta c_{t+1})}{\left(1 - \frac{\gamma}{R}\right)(R-1)^2}.$$

Appendix B also provides a closed form expression as a function of the dividend history. Note that this is not bounded below, since earnings are not bounded below in this arithmetic (exponential utility) model.

Using the asset pricing relationship it is possible to re-express the consumption function in terms of total wealth, $w_t = -Rb_t + d_t + p_t$. Here we will are studying a representative agent economy so that $\theta_t = 1$ for all t.

$$\begin{aligned} c_t &= \frac{\gamma}{R} c_{t-1} + \left(1 - \frac{\gamma}{R}\right) x_t - \psi \\ &= \frac{\gamma}{R} c_{t-1} + \left(1 - \frac{\gamma}{R}\right) \left(\frac{R-1}{R} \left[-Rb_t + d_t + p_t + \frac{R\alpha \times Var_t(c_{t+1})}{\left(1 - \frac{\gamma}{R}\right)(R-1)^2} \right] \right) - \psi \\ &= \frac{\gamma}{R} c_{t-1} + \left(1 - \frac{\gamma}{R}\right) \left(\frac{R-1}{R} \right) w_t + \Lambda \end{aligned}$$

where

$$\Lambda = \frac{\alpha}{2} \frac{Var_t(c_{t+1})}{R-1} - \frac{1}{R-1} \frac{1}{\alpha} \ln(R\delta).$$

We can rewrite the value function in the following simplified form (redefining the earlier value function notation):

$$V(c_{t-1}, w_t) = \frac{-1}{\alpha(1-\delta)} \exp\left(-\alpha \left[c_t - \gamma c_{t-1}\right]\right)$$

Relative risk aversion is given by:

$$\begin{aligned} CRRA &= -\frac{w\, V_{ww}}{V_w} \\ &= w\, \alpha\, \frac{\partial c_t}{\partial w_t} \\ &= w\alpha \left(1 - \frac{\gamma}{R}\right) \left(\frac{R-1}{R} \right) \\ &\simeq c\alpha \left(1 - \frac{\gamma}{R}\right). \end{aligned}$$

3.3 Calibration

To match historical data, we assume that the quarterly risk free (world) interest rate is $R = 1.0025$, implying that the annualized (net) rate is 0.01. We set the product of the discount factor and the gross interest rate to unity: $\delta R = 1$. The (quarterly) habit parameter is set to $\gamma = 0.9$, implying a half-life of adjustment of about 6 quarters.[28] Finally, we set the curvature of the utility function so that

$$\alpha = \frac{4}{d_0\left(1 - \frac{\gamma}{R}\right)}.$$

This implies that relative risk aversion (at initialization) is approximately four.[29]

The stochastic driving process is exogenous capital income, which are calibrated using real capital income from the NIPA accounts.[30] We use capital income from the NIPA accounts because it is seasonally adjusted (unlike corporate earnings) and is not artificially smoothed (like corporate dividends[31]). Figure 1 plots the natural log of real capital income at a quarterly frequency from 1947:1 to 2010:3.

Using this data, Figure 2 plots the impulse response functions resulting from estimating a range of ARIMA(p,1,0) models, with $p = 1, 10, 20, 30, 40$. We will henceforth refer to these as AR(p) models of the *differenced* data. Figure 2 reveals that the order of the model is critical in determining inferences about persistence. Low order models ($p = 1$ and 10) generate persistence estimates that are greater than or equal to 1. As the order of the model increases, estimated persistence falls dramatically. For an AR(40), estimated persistence

[28] Recall that the consumption function is $c_t = \frac{\gamma}{R}c_{t-1} + \left(1 - \frac{\gamma}{R}\right)\left(\frac{R-1}{R}\right)w_t + \text{constant}$.

[29] At initialization, with $b_0 = 0$, relative risk aversion in our CARA economy is

$$\left(\frac{R-1}{R}\right)w\alpha\left(1 - \frac{\gamma}{R}\right) \simeq c_0\alpha\left(1 - \frac{\gamma}{R}\right) \simeq d_0\alpha\left(1 - \frac{\gamma}{R}\right).$$

[30] Specifically, we study the natural log of real net operating surplus of private enterprises as reported in the U.S. National Income and Product Accounts (Bureau of Economic Analysis, 1947:1 to 2010:3). The net operating surplus of private enterprises is reported in NIPA Table 1.10, line 12. This definition is net of capital depreciation. To adjust for inflation, we use the GDP deflator.

[31] Dividends also have the problem that they miss cash-flow that is returned to shareholders through buy-backs.

after 40 quarters is about 1/3.

For AR(p) models with $p \leq 20$, there is a substantial gap between the persistence implied by the estimated model and the persistence implied by what we are assuming to be the true model: the AR(40). Hence, parsimonious AR(p) models yield a form of extrapolation bias in *levels* (not growth rates). In other words, "low" order AR(p) models imply that shocks are far more persistent than they are actually likely to be; low order models imply excess optimism in good times and excess pessimism in bad times.

The associated coefficient estimates from the five AR(p) models in Figure 2 are reported in Table 1. There are few significant coefficients after the 10th lag.[32]

In what follows, we will always assume that the *true data generating process* is the AR(40) estimated in the final column of Table 1 and plotted as an impulse response function in Figure 2. We are not loyal to this particular specification. Instead it is merely a convenient way of capturing hump-shaped dynamics: short-run momentum and long-run partial mean reversion. We will assume that this is the true data generating process because this case is interesting, and *not* because we believe that there is strong evidence for this specification. We only wish to argue that this hump-shaped process—the point estimate from the AR(40) case—is plausible.

The only remaining thing to specify is the beliefs of the agents in the economy. We don't have strong views about this either. We will therefore study the beliefs generated by every AR(p) model with $p = 1, 2, ..., 40$. In other words, we will assume that the true model is the AR(40) in the last column of Table 1, but that agents believe some AR(p) model which may not be the AR(40). We will study the properties of all of these belief assumptions: $p = 1, ..., 40$. When we assume $p = 40$, we are implicitly assuming that the agents know the true model (exactly); $p = 40$ is the rational expectations benchmark. When $p < 40$, we are assuming that agents use a simpler model than the true model.

[32]It is well-known that estimating AR coefficients is small samples yields biased coefficients (e.g., Shaman and Stine 1988). Monte Carlo simulations indicate that with our sample size, this does not significantly affect estimated persistence.

4 Simulation and Empirical Evaluation

To characterize the qualitative predictions of our model, we begin by reporting two impulse response functions—one for asset returns and one for consumption—that summarize the key mechanisms that are at work in the model. We then describe eight empirical moments that enable us to quantitatively evaluate the model's predictions. These empirical moments characterize the joint evolution of asset prices and consumption. The moments are chosen because they represent a test of the anomalous predictions that the model makes—e.g., cyclical fluctuations in asset returns and consumption growth. We find that these anomalous predictions match the actual empirical evidence well. Specifically, when agents hold beliefs generated by an AR(p) with $p \leq 20$, and unlike in the rational expectations benchmark, the model provides a good quantitative fit to the empirical moments.

4.1 Impulse Response Functions

We begin by reporting the impulse response function for cumulative excess returns. In our economy with exponential (CARA) utility, the excess return is more naturally expressed as an "excess gain," which is defined as

$$g_t = p_t + d_t - Rp_{t-1.} \tag{5}$$

Note that

$$\begin{aligned} \frac{g_t}{p_{t-1}} &= \frac{p_t + d_t - Rp_{t-1}}{p_{t-1}} \\ &= \frac{p_t + d_t}{p_{t-1}} - R, \end{aligned} \tag{6}$$

which is the standard definition of excess return, since R is the gross risk free rate. We study g_t, without dividing through by scaling factor p_{t-1} since prices can fall below zero in our CARA economy (assuming no free disposal of negative earnings), so the usual definition

of excess return (6) is not appropriate. Henceforth, we will refer to excess gains—equation (5)—when studying asset returns in our model, but the reader should intuitively think of excess gains as excess returns, since excess gains are an appropriately rescaled version of excess returns.

Figure 3 reports the impulse response function for the 'cumulative excess gain' following a positive (one-unit) shock to the earnings process at date zero. The cumulative excess gain is defined as,

$$\text{cumulative excess gain } t \text{ periods after the shock} = \sum_{s=0}^{t} g_t.$$

The cumulative excess gain is the CARA analog of the cumulative excess return from date 0 to date t.

In Figure 3, the impulse response function is reported for five simulated economies. In every case, the *true* process for earnings growth is the same—the AR(40) described in section 3.3. The economies differ only with respect to the *beliefs* that agents hold. Figure 3 analyzes economies in which agents generate forecasts using the following AR(p) models for earnings growth: $p = 1, 10, 20, 30, 40$.

Since the true data generating process is an AR(40), let's start with the case in which agents use the same model, an AR(40), to form beliefs—the rational expectations case. The impulse response function in this economy is the (flat) purple line in Figure 3; i.e., the cumulative excess gain jumps up at date zero and immediately plateaus. In the rational expectations benchmark, an impulse at date 0 does not forecast additional excess returns in periods after date zero.

When agents generate beliefs from an AR(p) model with $p < 40$, we are studying an economy with natural expectations, in which excess gains become more volatile and predictable. For example, when agents use an AR(1), they mistakenly infer that earnings impulses are highly persistent (cf. Figure 2). This belief in persistence causes the asset price to jump

up more than it should at date zero. Thereafter, the asset price tends to decline (and this decline is not foreseen by the agents). This decline is easy to see on the impulse response function, since there is only one impulse at date 0 and no impulse at any other point in time. In a full-blown simulation of the economy, the stream of negative excess gains would be largely masked by other contemporaneous shocks in the economy. Hence, the predictable negative excess gains would be very hard to infer.

As the order of the estimated AR(p) rises from $p = 1$ to $p = 40$, the impulse response functions look more and more like the rational expectations impulse response function. As p rises, the initial jump in the asset price becomes smaller, since agents with high p models correctly anticipate more mean reversion in the true data generating process. As p rises, the magnitude of the negative excess returns gets smaller.

Hence, models with low levels of p are characterized by a high degree of overreaction in asset prices. Models with p equal to or near 40 are characterized by little or no overreaction and little or no return predictability.

Figure 4 reports the impulse response function for consumption following a positive (one-unit) shock to the earnings process at date zero, and mimics the reporting conventions of Figure 3. Specifically, the impulse response function is reported for five simulated economies. In every case, the true process for earnings growth is the AR(40) described in section 3.3, while agents generate forecasts using an AR(p) model. We again plot the cases $p = 1, 10, 20, 30, 40$.

Let's again start with the case in which agents have AR(40) beliefs, which is the rational expectations case. In this world, a shock at date 0 would forecast a slow rise in consumption because the agents in our setting have a preference for slow adjustment, due to the habit term in their utility function. On the other hand, if the habit parameter, γ, were zero, the consumption series would jump up at date zero and immediately plateau.

When agents have natural expectations, i.e. they hold beliefs generated from an AR(p) model with $p < 40$, the impulse response function for consumption tends to become hump-shaped. For example, when agents use an AR(1), they mistakenly infer that the earnings

impulse is highly persistent, leading them to overreact to the earnings impulse (despite their preference for slow adjustment). About ten quarters after the initial impulse, they realize that they have gone too far and begin to reduce consumption. Over the long run, consumption falls from its peak level as households realize that their wealth is not as great as they initially thought (mirroring the negative excess gains in Figure 3). Since the NPV of these consumption lines must be the same across the p values— a consequence of the budget constraint—the long-run consumption asymptotes are different, reflecting pay-back for the short-run over-consumption.

As the estimated AR(p) rises from $p = 1$ to $p = 40$ in Figure 4, the impulse response functions again move towards the rational expectations benchmark. The initial growth in consumption becomes smaller, since agents that use high-p models believe in more mean reversion in the earnings process and therefore don't believe that the impulse to earnings has made them much more wealthy. Therefore, as p is increased, the magnitude of the predictable fall in consumption also gets smaller.

In summary, models with low levels of p generate overreaction of consumption. Models with p equal to or near 40 are characterized by little or no overreaction and little or no eventual declines in consumption.

4.2 Empirical Moments for Model Evaluation

We use the following eight moments to evaluate the empirical performance of our model under different assumptions. All data is annual, and goes from 1929 to 2010 (we use annual data so we can go further back in time, as the quarterly series in NIPA only start in 1947). To distinguish this *annual* data from the *quarterly* observation frequency of the model/calibration, we use the new time subscript τ.

The moments that follow reveal evidence for reversals that start around two years after an initial impulse. Our model predicts that these reversals will be delayed for two reasons. First, equity prices won't mean revert until agents start to notice a break between their forecasts

and the realizations of earnings (see Figure 2). For the simple models that we estimate, this divergence tends to occur about two years after an initial impulse. Second, consumption growth in the short-run is positively auto-correlated because of habits (see Figure 4). The presence of habits delays the onset of mean reversion in consumption and slows down the process of mean reversion. Because it takes longer for the consumption dynamics to play out, we always extend the consumption analysis an extra year beyond the window of the asset price analysis. After an initial impulse at date τ, the asset price dynamics are measured from τ+2 to τ+5. The consumption dynamics are measured from τ+2 to τ+6.

1. The correlation between excess returns of equity over the risk-free rate in year τ and cumulative excess returns from year $\tau + 2$ to year $\tau + 5$. This equals -0.22.[33]

2. The correlation between the ratio of S&P price at the end of year τ and average earnings over years $\tau - 9$ to τ (the *P/E_{10} ratio*) and excess returns from year $\tau + 2$ to year $\tau + 5$. This equals -0.38.[34]

3. The correlation between the one-year change in log consumption (i.e., $ln(c_\tau) - ln(c_{\tau-1})$) and cumulative excess returns from year $\tau + 2$ to year $\tau + 5$. This equals -0.30.[35]

4. The correlation between the P/E_{10} ratio and the change in log consumption from year $\tau + 2$ to year $\tau + 6$ (i.e., $ln(c_{\tau+6}) - ln(c_{\tau+2})$). This equals -0.17.

[33]Excess returns are defined as "the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate (from Ibbotson Associates)." (source: Ken French's online data library). Early evidence on long-term mean reversion in stock prices was presented in Fama and French (1988a) and Poterba and Summers (1988). The significance of this evidence was subsequently challenged on statistical grounds, e.g. by Richardson (1993). Cutler, Poterba and Summers (1991) look at a variety of stocks, bonds and foreign exchange markets and find a relatively slight negative autocorrelation in returns in many markets over a horizons of 3 to 5 years. Balvers, Wu and Gilliland (2000) consider a panel of 18 countries and document strong mean reversion in national equity indexes relative to the world index.

[34]Data on stock prices and earnings come from Robert Shiller's website. Earnings and dividend yields as predictors of future returns are studied e.g. by Campbell and Shiller (1988ab, 2005) and Fama and French (1988b).

[35]As our consumption measure, we use real per capita expenditures on non-durable consumption and services, from NIPA. The nominal data on expenditures are in NIPA table 2.3.5, while price indices are in table 2.3.4. The population data come from table 2.1. Consumption growth is a relatively little studied predictor for excess returns. However, in recent work, Moller (2008) and Moller and Rangvid (2011) show that high consumption growth between quarters 3 and 4 significantly predicts low excess returns over the following year.

5. The correlation between $ln(c_\tau) - ln(c_{\tau-1})$ and $ln(c_{\tau+6}) - ln(c_{\tau+2})$. This equals -0.25.

6. The mean annual excess return of equity over the risk-free rate (the "equity premium"). This equals 7.44%.

7. The standard deviation of annual excess return of equity over the risk-free rate. This equals 20.83%.

8. The standard deviation of $ln(c_\tau) - ln(c_{\tau-1})$. This equals 0.022.

While the directions and magnitudes of the correlations reported under 1) to 5) are suggestive, one has to be careful with the interpretation: as investigated in the finance literature by Stambaugh (1999), Volkanov (2003) and others, predictive regressions in short samples and with overlapping data can yield biased coefficients, and standard statistical inference may not be valid.[36]

To get an idea of the statistical significance of our results as well as the magnitude of a potential bias, we conduct a series of bootstrap exercises (similar to e.g. Nelson and Kim 1993). For moment 1), we draw repeatedly (with replacement) from the observed realization of excess returns under the assumption that excess returns are not serially correlated, and construct the measure of subsequent cumulative excess returns that we use in the empirical data. We then calculate the correlation between the two series and repeat this 10,000 times for samples of length 75 (corresponding to our empirical sample size), so we can construct a confidence interval under the null of independent returns. For moment 5), we instead assume one-period log consumption growth follows an AR(1) process.[37] We estimate this process and then simulate time series of consumption (of length equal to our empirical sample) by repeatedly drawing from the residuals. From this series, we construct the change in log consumption between years $\tau + 2$ and $\tau + 6$, and correlate it with the one-period change in

[36] See Lettau and Ludvigson (2010) for an overview of the issues as well as potential remedies.

[37] This is the approximate process that consumption growth should follow when a consumer has habit preferences; see e.g. Dynan (2000).

log consumption. Again, this is repeated 10,000 times, and we construct confidence intervals based on the obtained correlations.

For moments 2) to 4), we proceed in a similar manner: e.g. for 2), i) we assume and estimate an AR(1) process for P/E_{10} (accounting for the small-sample bias in the autoregression coefficient); ii) we assume that excess returns are independently distributed (the null); and iii) draw residual pairs under the null (with replacement), construct the predicted variable we use in the data (for a sample of equal length to our data), and estimate the correlation between the two series. We again repeat this 10,000 times and construct 90 and 95% confidence intervals for the correlation coefficient.

The results in Table 2 show that the correlations 3) and 5) are outside of the 90% confidence interval constructed under the null, but not outside the 95% interval. The "long-horizon autocorrelation" of excess returns (1) as well as the correlation of P/E_{10} with subsequent excess returns (2) are not quite significant at 10%, while the correlation between P/E and subsequent consumption growth (4) is far from significant.[38]

These relatively "weak" results do not come as a surprise, as a large literature in finance has established that predictability patterns that had appeared highly statistically significant in the early (1980s) literature are much less significant once statistical complications are accounted for. That said, our use of annual rather than quarterly data (as is more common in the literature) presumably weakens our statistical power. Furthermore, there exist more sophisticated ways of testing for predictability, and overall the finance literature seems to have concluded that "despite complexities with statistical inference in return predictability regressions, it is difficult to reconcile the historical behavior of the U.S. stock market without admitting some degree of predictability in excess returns" (Lettau and Ludvigson 2010, p. 620). Finally, even where we cannot reject the null hypothesis of no predictability, we certainly also cannot reject an alternative null with economically high predictability—the

[38]The wide confidence intervals for 2) and 4) may seem surprising; they are due to the high persistence of the P/E_{10} and the use of overlapping data for consumption growth and excess returns (cf. Boudoukh, Richardson, and Whitelaw, 2008).

currently available data just do not allow us to know.

4.3 Asset Return Predictability

For the next eight exhibits, we will illustrate the implications of the model for the full range of natural expectations cases. Specifically, beliefs about earnings are generated for all AR(p) models, for $p = 1, ..., 40$. As usual, we always assume that the true model generating earnings is an AR(40). The rational expectations case therefore coincides with the case $p = 40$; natural expectations correspond to the cases $1 \leq p < 40$. It turns out that when p is between 10 and 20 the natural expectations model does a good job of matching the point estimates for the eight empirical moments.

In the exhibits that follow, we always study simulations with 328 periods of quarterly data, matching the duration of our available empirical data (annual U.S. data since 1929). We then calculate the eight moments on these 328 simulated periods. We repeat this exercise 200 times and average the eight moments over these 200 (independent) simulations. We match the sample duration of our empirical data and the sample duration of our simulated data to make finite sample biases comparable across the empirical and simulated moments.

Figure 5 studies the correlation between excess returns in year τ and cumulative excess returns from year $\tau + 2$ to $\tau + 5$. The empirical correlation is -0.22. For low values of p, the simulated moment is about -0.3, rising to about -0.15 as p rises to 20. For the rational expectations case, $p = 40$, the correlation is approximately[39] zero.

Similarly, Figure 6 studies the correlation between the P/E_{10} ratio at year-end τ and cumulative excess returns from year $\tau + 2$ to $\tau + 5$. The empirical correlation is -0.38. For low values of p, the simulated moment is about -0.5, rising to about -0.35 as p rises to 20. For the rational expectations case, $p = 40$, the correlation is again approximately zero.

The natural expectations model matches these moments, because good earnings news at date zero generates over-reaction in asset prices (and increase the P/E_{10} ratio, since the

[39]The rational expectations benchmark is slightly below zero due to a variant of Hurwicz bias.

denominator is relatively unresponsive given that it is averaged over 40 quarters). Over the next five years, some of that initial return will be reversed as agents discover that the good earnings news was not as persistent as they had anticipated.

In our model, Lettau and Ludvigson's (2001) *cay* variable is positively correlated with excess returns (for the same reason that P/E predicts future excess returns). The *cay* variable is constructed by identifying the co-integration residual between consumption and wealth: $cay_t = c_t - \Psi w_t$. In our model, a positive earnings shock raises asset values but does relatively little to consumption in the short run (due to slow adjustment). Thus a positive earnings shock causes *cay* to fall. Over the next five years, asset returns will tend to be low, due to unanticipated mean reversion in earnings. Hence, our model provides an alternative explanation for the observed positive correlation between *cay* and future excess returns.

Figure 7 studies the correlation between consumption growth at year τ and equity returns from year $\tau+2$ to $\tau+5$. The empirical correlation is -0.30. For low values of p, the simulated moment is about -0.5, rising to about -0.35 as p rises to 20. For the rational expectations case, $p = 40$, the correlation is approximately zero.[40] A high level of lagged consumption growth is a proxy for a positive earnings shock (at that date or some preceding date). Natural expectations agents subsequently discover that the earnings news was not as persistent as they had anticipated, leading to below average excess gains/returns.

4.4 Consumption Growth Predictability

Figure 8 studies the correlation between the P/E_{10} ratio at year-end τ and consumption growth from year $\tau + 2$ to $\tau + 6$. The empirical correlation is -0.17. For low values of p, the simulated moment is about -0.35, rising to about -0.15 as p rises to 20. For the rational expectations case, $p = 40$, the simulated correlation is approximately $+0.25$; this large positive value in the rational expectations case is due to habit formation and slow adjustment in consumption. For the natural expectations agents, $p < 40$, the negative

[40] Once again, Hurwicz bias lowers the rational expectations correlation below zero.

correlations arise because unanticipated mean reversion in earnings swamps the effects of habit formation. Good earnings news increases the P/E ratio, since the denominator is relatively unresponsive given that it is averaged over 40 quarters. Over the next five years, some of that initial return will be reversed as natural expectations agents discover that the earnings news was not as persistent as they had anticipated. This causes consumption to decline from year $\tau + 2$ to $\tau + 6$ for the natural expectations agents. This decline is strong enough to overcome the positive correlation that would otherwise arise because of habit formation. The consumption impulse response function—Figure 4—illustrates the effects of the countervailing forces of unanticipated mean reversion and habit formation. For the natural expectations agents, unanticipated mean reversion offsets the pure effects of habit formation.

Figure 9 studies the correlation between consumption growth at year τ and consumption growth from year $\tau + 2$ to $\tau + 6$. The empirical correlation is -0.25. For low values of p, the simulated moment is about -0.3, rising to about -0.2 as p rises to 20. For the rational expectations case, $p = 40$, the simulated correlation is approximately $+0.15$; once again, the positive correlation in the rational expectations case is due to habit formation. The consumption of natural expectations agents ($p < 40$) initially overreacts to a positive earnings shock. Subsequently, they realize that the initial positive earnings shock was not as persistent as they had anticipated, which in turn causes consumption to decline from year $\tau + 2$ to $\tau + 6$. This decline is strong enough to overcome the positive correlation that would otherwise arise because of habit-based slow adjustment.

The fact that consumption growth is only relatively weakly predicted by lagged equity prices is a strength of our model. Beeler and Campbell (2009) point out that Bansal and Yaron's (2004) long-run risks model is problematic because it implies much more predictability of consumption growth from lagged price/dividend ratios than is present in the data, while at the same time generating too little excess return predictability.[41] Beeler and Campbell

[41]For an alternative view, see Bansal, Kiku, and Yaron (2009), who argue that consumption growth becomes much more predictable once additional predictor variables are considered.

also note the empirical evidence for medium-run mean reversion in the level of consumption, while the long-run risks model implies that all autocorrelations of consumption growth should be positive.

4.5 The Equity Premium and Standard Deviations

In the natural expectations economy, agents perceive that equities are very risky. Agents believe that earnings innovations are highly persistent, so that equity returns are therefore highly volatile and anticipated to strongly covary with long-run consumption growth. This leads agents to require a high premium to hold equities for low values of p. By contrast, when $p = 40$ – the rational expectations case – the equity premium is only 1/10 of 1%. Figure 10 plots the (annual) equity premium for different values of p as well as the historical premium of 7.4%.[42]

Finally, we are able to roughly match the equity premium without generating counterfactually high levels of asset return volatility – see Figure 11 – or consumption volatility – see Figure 12. No matter what the order of the forecasting model, the standard deviation of consumption volatility is less than the historical standard deviation of consumption volatility. We don't see this as a problem for two reasons. First, we can easily raise the standard deviation of consumption changes by lowering the (high) level of the habit: e.g., calibrate $\gamma = 0.75$ rather than the chosen value of 0.9. Second, the post-War variation in consumption is substantially lower than the full-sample variation, which is used to generate the empirical (green) line.

5 Agents with Rational Expectations

Starting with the natural expectations economy described above (with $p < 40$), we now add a vanishingly small measure of agents with rational expectations. These 'RE agents' know

[42]See Appendix B for the formulae that we use to calculate the equity premium. The X marks on the upper edge of Figure 10 represent equity premia that exceed the vertical scale of the figure.

all of the coefficients in the true data generating process for dividends—the AR(40). This perfect knowledge is meant to proxy for the more plausible assumption that these new agents have a high degree of sophistication relative to the agents with natural expectations.

We solve for the behavior of the RE agents assuming that they have no effect on equilibrium prices. This *small mass* assumption is made for tractability and future work should consider equilibria in which the rational agents are wealthy enough to influence equilibrium prices. The current analysis is indicative of the results that will apply in that more realistic case.

Appendix C derives the optimal policy function – both asset allocation and consumption – of the RE agents. Later in this section, we present the optimal policy rules for the RE agent. Those policy rules take a large set of other objects as inputs. We now discuss and define those inputs before turning to the key result of this section.

Let Φ be the vector of true autoregressive coefficients in the earnings process, which as always is an AR(40). Let $\hat{\Phi}$ be the vector of autoregressive coefficients in the earnings process as perceived by the NE agents. Let σ_ε be the true standard deviation of shock to dividend process. Let $\hat{\sigma}_c^2$ be the conditional variance of consumption changes as perceived by the NE agents. Let $e_{1,p} = [1\ 0\ 0\ ...\ 0]'$ be a standard basis vector of length p. Define the following variables:

$$\sigma_g = \frac{R}{R-1}\sigma_\varepsilon e_{1,p}' \left(I - \frac{1}{R}\hat{\Phi}\right)^{-1} e_{1,p}.$$

$$\mu = \frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)}\hat{\sigma}_c^2.$$

$$M = \frac{R}{R-1}e_{1,p}' \left[\Phi - \hat{\Phi}\left(I - \frac{1}{R}\hat{\Phi}\right)^{-1}\left(I - \frac{1}{R}\Phi\right)\right].$$

$$B = \frac{1}{2R\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g^2}\sum_{k=0}^{\infty} R^{-k}\hat{\Phi}'^k M' M \hat{\Phi}^k.$$

$$A' = \frac{2\mu}{R\sigma_g}\left(\frac{1}{2(1-\frac{\gamma}{R})\alpha\sigma_g}M + \sigma_e e_{1,p}' B\hat{\Phi}\right)\left(I - \frac{1}{R}\hat{\Phi}\right)^{-1}.$$

$$\lambda = (1 - \frac{\gamma}{R})\sigma_e^2 e'_{1,p} B e_{1,p}$$

$$q = \frac{1 + 2\alpha\lambda}{2(R-1)(1-\frac{\gamma}{R})\alpha\sigma_g^2}\mu^2 - \frac{\mu\sigma_e}{\sigma_g(R-1)}A'e - \frac{1}{2\alpha(R-1)(1-\frac{\gamma}{R})}\ln(1+2\alpha\lambda).$$

We can now describe the policy function of RE agents. We introduce a new variable, θ_t, which is the quantity of units of the Lucas tree held by an RE agent. By comparison, the quantity of units of the Lucas tree held by an NE agent is normalized to one. Hence, when $\theta_t > 1$, each RE agent holds more equities than each NE agent.

Theorem 1 *For an agent with rational expectations, the optimal asset allocation, θ_t, is given by*

$$\theta_t = \frac{R}{R-1}\frac{1}{\sigma_g}\left(\frac{\mu + M\Delta\vec{d}_t}{\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g} + \frac{2\mu\sigma_e^2}{\sigma_g}e'_{1,p}Be_{1,p} - \sigma_e A'e - 2\sigma_e e'_{1,p}B\hat{\Phi}\Delta\vec{d}_t\right),$$

and the optimal consumption rule is given by:

$$c_t = \left(1-\frac{\gamma}{R}\right)\left(\frac{R-1}{R}w_t + A'\Delta\vec{d}_t + \Delta\vec{d}_t' B\Delta\vec{d}_t + q\right) + \frac{\gamma}{R}c_{t-1}.$$

The proof is provided in Appendix C. The consumption policy rule implies that consumption is still a weighted average of the annuity value of wealth and lag consumption, with weights $\left(1-\frac{\gamma}{R}\right)$ and $\frac{\gamma}{R}$. There are two additional terms in the consumption policy, reflecting the impact of dividend shocks excluding their effect on wealth and prices. These terms represent the ability of the fully informed investor to predict future "returns" (gains), and adjust consumption accordingly. In our setting, where the rational agent knows about mean reversion that is not "priced in" to asset prices, and generally has positive equity holdings, the first order effect (A) is mostly negative—that is, the agent consumes more in response to a positive dividend shock, but less than their wealth gain would suggest. The second order effect (B) is mostly positive—the agent has some ability to predict future gains, and the benefits of this ability are larger when large shocks have occurred in the recent past.

The asset allocation result is harder to interpret than the consumption rule, other than that it is a linear function of the dividend history. However, we can write a simple lower

bound on the average equity holding:

Corollary 2 *On average, rational agents will hold a quantity of equities greater than*

$$\overline{\theta}_{min} = \frac{\widehat{\sigma}_{\varepsilon}^2}{\sigma_{\varepsilon}^2}.$$

Note that $\frac{\widehat{\sigma}_{\varepsilon}^2}{\sigma_{\varepsilon}^2}$ is the ratio of variances of impulses to earnings. The numerator, $\widehat{\sigma}_{\varepsilon}^2$, is the variance perceived by the NE agents. The denominator, σ_{ε}^2, is the variance perceived by the RE agents. Since the NE agents estimate a simpler model than the (true) model used by the RE agents, it follows that $\widehat{\sigma}_{\varepsilon}^2 > \sigma_{\varepsilon}^2$.

The average value of θ_t (given in Appendix C) includes two terms related to the consumption policy matrices A and B. These terms reflect the contrarian nature of the consumption policy: because dividend shocks have a muted effect on consumption via A, and because shocks of any sort increase consumption via B, the Lucas tree is less risky for the rational investors than would otherwise be the case, and the equilibrium outcome is for them to hold even more equities. Because both of these terms have a positive impact on average equity holdings, discarding them and simplifying results in the lower bound described above.

We emphasize this lower bound because it reflects what a rational agent would do if they knew there was an equity premium, but were not able to time returns. In our framework, this would occur if $\hat{\sigma}_{\varepsilon} > \sigma_{\varepsilon}$ but $M = 0$.

In Figure 13, we report the average asset allocation, and this lower bound, over different model orders chosen by the NE agents. The extent to which NE agents choose low order models determines both the excessive variance they perceive, and the ability of rational agents to predict returns, and it is therefore unsurprising that the equity holdings for rational agents is essentially monotonically decreasing in the order of model chosen by the NE agents.

Finally, RE agents will make counter-cyclical equity investments. Because of overreaction among the NE agents, the (rational) expected return on equities falls after positive shocks, leading the RE agents to lower their equity exposure during good times. Conversely the RE

agents increase their equity exposure during bad times.

6 Conclusion

This paper examines an open endowment economy in which (1) fundamentals have hump-shaped dynamics (momentum in the short run and partial mean reversion in the long run), and (2) agents don't know these dynamics and generate beliefs by fitting parsimonious models. We show that these two assumptions have a large set of empirically valid implications: endogenous extrapolation bias; pro-cyclical excess optimism; over-reaction in asset pricing; mean reversion in returns; amplified dynamics in consumption; cyclical dynamics in consumption; a high equity premium; relatively highly levered equity exposure among sophisticated investors; and counter-cyclical asset allocation among sophisticated investors. The framework that we have described provides a parsimonious and psychologically plausible explanation for these stylized facts.

We believe that our two assumptions also characterize other 'macro' markets. For example, the same approach could be used to study dynamics in housing prices, residential investment, non-residential investment, inventory accumulation, international capital flows, bond markets, or commodity prices.[43] Wherever our two assumptions apply, markets will be characterized by amplified cycles, overreaction, excess volatility, and asset returns that are negatively autocorrelated over the medium-run.

Finally, we wish to highlight a gap in our analysis. Our model is about belief formation, but all of our evidence about beliefs is indirect – for instance, consumption and asset price dynamics. We believe that researchers should also measure beliefs directly. Some research on bubbles has moved in this direction (e.g., Vissing-Jorgenson 2003; Shiller 2005; Piazessi and Schneider 2009, 2011) and we hope that more will do so in the future.

[43]For some alternative approaches, see for example Lansing (2009), Hassan and Mertens (2010), Adam and Marcet (2010), LeBaron (2010), Burnside, Eichenbaum, and Rebelo (2011), and Piazzesi and Schneider (2011).

References

[1] Abraham, Jesse M., and Patric H. Hendershott. 1996. "Bubbles in Metropolitan Housing Markets," Journal of Housing Research, 7(2), 191-207.

[2] Adam, Klaus, and Adam Marcet. 2010. "Booms and Busts in Asset Prices," Bank of Japan Institute for Monetary and Economic Studies IMES Discussion Paper Series 2010-E-2.

[3] Alessie, Rob, and Annamaria Lusardi. 1997. "Consumption, saving and habit formation," Economic Letters, 55: 103–108.

[4] Bacchetta, Philippe, Elmar Mertens, and Eric van Wincoop. 2009. "Predictability in financial markets: What do survey expectations tell us?" Journal of International Money and Finance, 28: 406–426.

[5] Baker, Malcolm, and Jeffrey Wurgler. 2007. "Investor Sentiment in the Stock Market," Journal of Economic Perspectives, 21(2): 129–151.

[6] Balvers, R., Y. Wu, and E. Gilliland. 2000. "Mean Reversion across National Stock Markets and Parametric Contrarian Investment Strategies," Journal of Finance, 55(2), 745–772

[7] Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," Journal of Finance, 59(4): 1481–1509.

[8] Bansal, Ravi, Dana Kiku, and Amir Yaron. 2009. "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," Working Paper.

[9] Barberis, Nicholas C. 2010. "Psychology and the Financial Crisis of 2007-2008," Working Paper, Yale School of Management.

[10] Barberis, Nicholas C., A. Shleifer, and R. W. Vishny. 1998. "A Model of Investor Sentiment," Journal of Financial Economics, 49(3), 307-343.

[11] Barsky, Robert B., and J. Bradford De Long. 1993. "Why Does the Stock Market Fluctuate?" Quarterly Journal of Economics, 108(2), 291-311.

[12] Beeler, Jason, and John Y. Campbell. 2009. "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," NBER Working Paper No. w14788.

[13] Benartzi, Shlomo. 2001. "Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock." The Journal of Finance, 56(5): 1747-1764.

[14] Benartzi, Shlomo, and Richard H. Thaler. 2007. "Heuristics and Biases in Retirement Savings Behavior," Journal of Economic Perspectives, 21(3), 81-104.

[15] Bergman, N., and S. Roychowdhury. 2008. "Investor Sentiment and Corporate Disclosure," Journal of Accounting Research, 46(5), 1057–1083.

[16] Boudoukh, Jacob, Matthew Richardson, and Robert F. Whiltelaw. 2008. "The Myth of Long-Horizon Predictability," Review of Financial Studies, 21(4), 1577–1605.

[17] Branch William A., and George W. Evans. 2007. "Model uncertainty and endogenous volatility," Review of Economic Dynamics, 10, 207- 37

[18] Branch, William A., and George W. Evans. 2010. "Asset Return Dynamics and Learning." Review of Financial Studies, 23(4):1651-1680

[19] Brav, Alon, and J.B. Heaton. 2002. "Competing Theories of Financial Anomalies," Review of Financial Studies, 15(2), 575–606.

[20] Burnside, Craig,, Martin Eichenbaum, and Sergio Rebelo. 2011. "Understanding Booms and Busts in Housing Markets." Mimeo.

[21] Caballero, Ricardo J. 1990. "Consumption puzzles and precautionary savings," Journal of Monetary Economics, 25(1), 113–136.

[22] Campbell, John Y. 2003. "Consumption-based Asset Pricing" Chapter 13 in: Handbook of the Economics of Finance Vol. IB, George Constantinides, Milton Harris, and Rene Stulz eds., Amsterdam: North-Holland, 803-887.

[23] Campbell, John Y., and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," Journal of Political Economy, 107, 205-251.

[24] Campbell, John Y., and N. Gregory Mankiw. 1987. "Are Output Fluctuations Transitory?" Quarterly Journal of Economics, 102(4): 857-880.

[25] Campbell, John Y., and Robert J. Shiller. 1988a. "Stock Prices, Earnings, and Expected Dividends." The Journal of Finance, 43(3): 661-676.

[26] Campbell, John Y., and Robert J. Shiller. 1988b. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." The Review of Financial Studies, 1(3): 195-228.

[27] Campbell, John Y., and Robert J. Shiller. 2005. "Valuation Ratios and the Long-Run Stock Market Outlook: An Update," Chapter 5 in Richard H. Thaler ed. Advances in Behavioral Finance, Volume II, Princeton University Press, 173–201

[28] Carroll, Christopher D., Martin Sommer, and Jiri Slacalek. Forthcoming,"International Evidence on Sticky Consumption Growth," Review of Economics and Statistics.

[29] Case, Karl E., and Robert J. Shiller. 2003. "Is There a Bubble in the Housing Market?" Brookings Papers on Economic Activity, 2003(2): 299-342.

[30] Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark. 2000. "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?" American Economic Review, 90(4): 787–804.

[31] Chalmers, John, and Jonathan Reuter. 2009. "How Do Retirees Value Life Annuities? Evidence from Public Employees." NBER Working Paper 15608.

[32] Chan, Louis K.C., Jason Karceski and Josef Lakonishok. 2003. "The Level and Persistence of Growth Rates." Journal of Finance, 58: 643–684.

[33] Chen, L., C.E. Moise, and X.S. Zhao. 2009. "Myopic Extrapolation, Price Momentum, and Price Reversal," Working Paper, http://ssrn.com/abstract=1429612

[34] Chevalier, Judith A., and Glenn Ellison. 1997. "Risk Taking by Mutual Funds as a Response to Incentives." The Journal of Political Economy, 105(6): 1167-1200.

[35] Choi, James J. 2006. "Extrapolative Expectations and the Equity Premium." Yale University mimeo.

[36] Choi, James J., David I. Laibson, Brigitte C. Madrian, and Andrew Metrick. 2004. "Employees' Investment Decisions About Company Stock." In Pension Design and Structure: New Lessons from Behavioral Finance, ed. Olivia S. Mitchell and Stephen P. Utkus, 121-136. New York: Oxford University Press.

[37] Choi, James J., David I. Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. "Reinforcement Learning and Savings Behavior." The Journal of Finance, 64(6): 2515-2534.

[38] Cochrane, John H. 1988. "How Big is the Random Walk in GNP?" Journal of Political Economy, 96(5): 893-920.

[39] Cochrane, John H. 2007. "Financial Markets and the Real Economy." In: R. Mehra, ed., Handbook of the Equity Premium, Amsterdam: North Holland, 273–325.

[40] Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1991. "Speculative Dynamics." The Review of Economic Studies, 58(3): 529-546.

[41] Daniel, K. D., D. Hirshleifer, and A. Subrahmanya. 1998. "Investor psychology and security market under- and overreactions," Journal of Finance, 53, 1839- 1886.

[42] De Bondt, Werner F. M. 1993. "Betting on Trends: Intuitive Forecasts of Financial Risk and Return." International Journal of Forecasting, 9(3): 355-371.

[43] De Bondt, Werner F. M., and Richard H. Thaler. 1985. "Does the Stock Market Overreact?" The Journal of Finance, 40(3): 793-805.

[44] De Bondt, Werner F. M., and Richard H. Thaler. 1989. "Anomalies: A Mean-Reverting Walk Down Wall Street." Journal of Economic Perspectives, 3(1): 189-202.

[45] De Bondt, Werner F.M, and Richard H. Thaler. 1990. "Do Security Analysts Overreact," American Economic Review, 80(2), 52–57.

[46] De Grauwe, Paul. 2010. "Top-Down versus Bottom-Up Macroeconomics." CESifo Working Paper 3020.

[47] De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert Waldmann. 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." The Journal of Finance, 45(2): 379-395.

[48] Dwyer, Gerald P., Arlington W. Williams, Raymond C. Battalio, Timothy I. Mason. 1993. "Tests of Rational Expectations in a Stark Setting." The Economic Journal, 103(418): 586-601.

[49] Dynan, Karen E. 2000. "Habit Formation in Consumer Preferences: Evidence from Panel Data," American Economic Review, 90(3), 391-406.

[50] Dynan, Karen E., and Dean M. Maki. 2001. "Does Stock Market Wealth Matter for Consumption?" FEDS Discussion Paper No. 2001-23.

[51] Easterwood, J. C., and S. R. Nutt. 1999. "Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism?" Journal of Finance, 54(5), 1777–1797

[52] Eusepi, S., and B. Preston. Forthcoming. "Expectations, Learning and Business Cycle Fluctuations," American Economic Review.

[53] Evans, George W., and Seppo Honkapohja. 2001. Learning and Expectations in Macroeconomics. Princeton: Princeton University Press.

[54] Evans, George W., and Seppo Honkapohja. 2009. "Learning and Macroeconomics." Annual Review of Economics, 1: 421-451.

[55] Fama, E., and K. French. 1988a."Permanent and temporary components of stock prices," Journal of Political Economy 96, 246–273.

[56] Fama, E., and K. French. 1988b."Dividend Yields and Expected Stock Returns," Journal of Financial Economics 25, 23–49.

[57] Friedman, Benjamin. 1979. "Optimal Expectations and the Extreme Information Assumptions of 'Rational Expectations' Macromodels," Journal of Monetary Economics 5, 23-41.

[58] Fuster, Andreas, David Laibson, and Brock Mendel. 2010. "Natural Expectations and Macroeconomic Fluctuations," Journal of Economic Perspectives, 24(4): 67–84.

[59] Gabaix, Xavier. 2011. "A Sparsity-Based Model of Bounded Rationality", mimeo.

[60] Gabaix, Xavier and David I. Laibson. 2002. "The 6D Bias and the Equity Premium Puzzle" NBER Macroeconomics Annual, 16:257-311.

[61] Gabaix, Xavier, David Laibson, Guillermo Moloche and Stephen Weinberg. 2006. "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model." American Economic Review 96 (4): 1043-1068.

[62] Gerardi, Kristopher S., Andreas Lehnert, Shane M. Sherlund, and Paul S. Willen. 2008. "Making Sense of the Subprime Crisis." Brookings Papers on Economic Activity, 2008: 69-145.

[63] Gigerenzer, Gerd, and Daniel G. Goldstein. 1996. "Reasoning the fast and frugal way: Models of bounded rationality,"Psychological Review, 103(4): 650–669.

[64] Glaeser, Edward L., Joshua D. Gottlieb, Joseph Gyourko. 2010. "Can Cheap Credit Explain the Housing Boom?" NBER Working Paper 16230.

[65] Goetzmann, William N., Liang Peng, and Jacqueline Yen. 2009. "The Subprime Crisis and House Price Appreciation." NBER Working Paper 15334.

[66] Gourinchas, Pierre-Olivier, and Aaron Tornell. 2004. "Exchange rate puzzles and distorted beliefs," Journal of International Economics, 64: 303–333.

[67] Greenwood, Robin, and Samuel G. Hanson. 2010. "Issuer Quality and Corporate Bond Returns," Working Paper, Harvard University.

[68] Hannan, E.J. 1980. "The Estimation of the Order of an ARMA Process," Annals of Statistics, 8, 1071–1081.

[69] Haruvy, Ernan, Yaron Lahav, and Charles N. Noussair. 2007. "Traders' Expectations in Asset Markets: Experimental Evidence." The American Economic Review, 97(5): 1901-1920.

[70] Hassan, Tarek A. and Mertens, Thomas M. 2010. "The Social Cost of Near-Rational Investment," Working Paper, http://ssrn.com/abstract=1106347

[71] Hey, John D. 1994. "Expectations Formation: Rational or Adaptive or É?" Journal of Economic Behavior and Organization, 25(3): 329-349.

[72] Hommes, Cars, Joep Sonnemans, Jan Tulnstra, and Henk van de Velden. 2008. "Expectations and Bubbles in Asset Pricing Experiments." Journal of Economic Behavior and Organization, 67(1): 116-133.

[73] 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

[74] Hong, Harrison, Jeremy C. Stein, and Jialin Yu. 2007. "Simple Forecasts and Paradigm Shifts." The Journal of Finance, 62(3): 1207-1242.

[75] Huang, K. X., Liu, Z. and Zha, T. 2009. "Learning, Adaptive Expectations and Technology Shocks," The Economic Journal, 119, 377- 405

[76] Kahneman, Daniel, and Amos N. Tversky. 1973. "On the Psychology of Prediction." Psychological Review, 80(4): 237-251.

[77] La Porta, Rafael 1996. "Expectations and the Cross-Section of Stock Returns," Journal of Finance, 51, 1715–42.

[78] Lakonishok, J., A. Shleifer, and R. W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk," Journal of Finance, 49(5), 1541–1578

[79] Lansing, Kevin J. 2006."Lock-in of extrapolative expectations in an asset pricing model," Macroeconomic Dynamics, 10, 317-348

[80] Lansing, Kevin J. 2009. "Speculative Growth, Overreaction, and the Welfare Cost of Technology-Driven Bubbles." Federal Reserve Bank of San Francisco Working Paper 2008-08.

[81] LeBaron, Blake. 2010. "Heterogeneous Gain Learning and the Dynamics of Asset Prices," Working Paper, Brandeis University.

[82] LeBaron, Blake, W. Brian Arthur, and Richard G. Palmer. 1999. "Time Series Properties of an Artificial Stock Market." Journal of Economic Dynamics and Control, 23(9): 1487-1516.

[83] LeRoy, Stephen F., and Richard D. Porter. 1981. "The present value relation: Tests based on implied variance bounds," Econometrica, 49: 555–574.

[84] Lettau, Martin, and Sydney C. Ludvigson. 2001. "Consumption, Aggregate Wealth and Expected Stock Returns," Journal of Finance, 56(3): 815–849.

[85] Lettau, Martin, and Sydney C. Ludvigson. 2010. "Measuring and Modeling Variation in the Risk-Return Tradeoff" in: Handbook of Financial Econometrics, ed. by Y. Ait-Sahalia and L. P. Hansen vol. 1, pp. 617-690. Elsevier Science B.V., North Holland, Amsterdam.

[86] Lim, T. 2001. "Rationality and Analysts' Forecast Bias," Journal of Finance, 56: 369-385.

[87] Malmendier, Ulrike M., and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" Quarterly Journal of Economics.

[88] Marcet, A., and J.P. Nicolini. 2003. "Recurrent Hyperinflations and Learning," American Economic Review, 93(5), 1476–1498

[89] McQuarrie, Allan D.R, and Chih-Ling Tsai. 1998. "Regression and Time Series Model Selection," Singapore: World Scientific Publishing.

[90] Moller, Stig V. 2008. "Consumption growth and time-varying expected stock returns." Finance Research Letters, 5: 129-136.

[91] Moller, Stig V., and Jesper Rangvid. 2011. "The fourth-quarter consumption growth rate and expected returns," Working Paper, Aarhus School of Business and Copenhagen Business School.

[92] Morley, J.C., C.R. Nelson, and E. Zivot. 2003. "Why are Beveridge-Nelson and unobserved-component decompositions of GDP so different?" Review of Economics and Statistics, 85, 235–243.

[93] Muellbauer, John N., and Anthony Murphy. 1997. "Booms and Busts in the UK Housing Market." The Economic Journal, 107(445): 1701-1727.

[94] Neath, A.A., and J.E. Cavanaugh. 1997. "Regression and time series model selection using variants of the Schwarz information criterion," Communications in Statistics - Theory and Methods, 26, 559-580.

[95] Nelson, C.R., and M.J. Kim. 1993. "Predictable Stock Returns: The Role of Small Sample Bias," Journal of Finance, 48, 641–661.

[96] Parker, Jonathan A. 2001. "The Consumption Risk of the Stock Market," Brookings Papers on Economic Activity, 2, 279–348.

[97] Pastor, Lubos,and Pietro Veronesi. 2009. "Learning in Financial Markets." Annual Review of Financial Economics, 1: 361-381.

[98] Perron, Pierre, and Tatsuma Wada. 2009. "Let's take a break: Trends and cycles in US real GDP," Journal of Monetary Economics, 56, 749–765.

[99] Piazzesi, Monika, and Martin Schneider. 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." American Economic Review, 99(2): 406-411.

[100] Piazzesi, Monika and Martin Schneider. 2011. "Trend and Cycle in Bond Premia," mimeo.

[101] Poterba, J., and L. Summers. 1988. "Mean reversion in stock prices: Evidence and implications," Journal of Financial Economics 22, 27–59.

[102] Previtero, Alessandro. 2010. "Stock Market Returns and Annuitization." Working Paper, UCLA Anderson.

[103] Rabin, Matthew. 2002. "Inference by Believers in the Law of Small Numbers," Quarterly Journal of Economics, 117, 775-816.

[104] Rabin, Matthew, and Dimitri Vayanos. 2010. "The Gambler's and Hot-Hand Fallacies: Theory and Applications," Review of Economic Studies, 77(2), 730–778.

[105] Reinhart, C., and K. Rogoff. 2009. "This Time is Different: Eight Centuries of Financial Folly." Princeton NJ: Princeton University Press.

[106] Richardson, M. 1993. "Temporary Components of Stock Prices: A Skeptic's View," Journal of Business & Economic Statistics, 11(2),199–207.

[107] Sargent, Thomas J. 1993. Bounded Rationality in Macroeconomics. Oxford: Oxford University Press.

[108] Shaman, Paul, and Robert A. Stine. 1988. "The Bias of Autoregressive Coefficient Estimators," Journal of the American Statistical Association, 83(403), 842–848.

[109] Shiller, Robert J. 1981. "Do stock prices move too much to be justified by subsequent changes in dividends?" American Economic Review, 71: 421–436.

[110] Shiller, Robert J. 2005. Irrational Exuberance: Second Edition. Princeton: Princeton University Press.

[111] Sims, C.A. 1998. "Stickiness," Carnegie-Rochester Conference Series on Public Policy 49, 317-356.

[112] Sims, C.A. 2003. "Implications of Rational Inattention," Journal of Monetary Economics 50(3), 665-690

[113] Sirri, Erik R., and Peter Tufano. 1998. "Costly Search and Mutual Fund Flows." The Journal of Finance, 53(5): 1589-1622.

[114] Stambaugh, R. F. 1999. "Predictive Regressions," Journal of Financial Economics, 54, 375–421.

[115] Stock, James H., 1994. "Unit Roots, Structural Breaks and Trends." in: Handbook of Econometrics, Volume IV, eds. R.F. Engle and D.L. McFadden.

[116] Tesfatsion, Leigh S., and Kenneth L. Judd, ed. 2006. Handbook of Computational Economics: Agent-Based Computational Economics. Vol. 2. Amsterdam: North-Holland Publishing Company.

[117] Timmermann, A. G. 1993. "How learning in financial markets generates excess volatility and predictability in stock prices," Quarterly Journal of Economics, 108, 1135–1145.

[118] Tversky, Amos N., and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probabilities." Cognitive Psychology, 5: 207-232.

[119] Tversky, Amos N., and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." Science, 185(4157): 1124-1131.

[120] Valkanov, R. 2003. "Long-Horizon Regressions: Theoretical Results and Applications," Journal of Financial Economics, 68, 201–232.

[121] Vissing-Jorgensen, Annette. 2003. "Perspectives on Behavioral Finance: Does 'Irrationality' Disappear With Wealth? Evidence from Expectations and Actions." NBER Macroeconomics Annual, 18(1): 139–194.

[122] Woodford, Michael. 2009. "Information-Constrained State-Dependent Pricing," Journal of Monetary Economics 56(S): 100–124.

7 Appendix A: Matrix Representation of Consumption Model

7.1 Solving for CARA policy for ARIMA(p,1,0) case

$$\Delta d_t = \Phi(L)\Delta d_t + \sigma_e \epsilon_t$$

We can represent this system as a AR(1) system with evolution operator Φ:

$$\begin{bmatrix} \Delta d_t \\ \Delta d_{t-1} \\ ... \\ \Delta d_{t-p+1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & ... & \phi_p \\ 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & & & \\ 0 & 0 & ... & 0 \end{bmatrix} \begin{bmatrix} \Delta d_{t-1} \\ \Delta d_{t-2} \\ ... \\ \Delta d_{t-p} \end{bmatrix} + \begin{bmatrix} \sigma_e \\ 0 \\ ... \\ 0 \end{bmatrix} \epsilon_t$$

There is also the foreign debt variable, b_t, which evolves as

$$b_{t+1} = c_t + Rb_t - d_t - y_t$$

Assume R and y_t are constant. Define the AR(1) representation state vector:

$$z_t = \begin{bmatrix} B_t \\ c_{t-1} \\ 1 \\ y_t \\ d_t \\ \Delta d_t \\ \Delta d_{t-1} \\ ... \\ \Delta d_{t-p+1} \end{bmatrix}$$

Use Alessie-Lusardi (1997) habits with CARA utility:

$$u(c_t, c_{t-1}) = -\frac{1}{\alpha}\exp(-\alpha(c_t - \gamma c_{t-1}))$$

Guess a linear policy function, $c_t = P' z_t$. Because the policy function is linear, we can define the AR1 evolution operator:

$$M = \bar{M} + NP' = \begin{bmatrix} R & 0 & 0 & -1 & -1 & \vec{0} \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 1 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 1 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 1 & e'_{1,p}\Phi \\ \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \Phi \end{bmatrix} + (e_{1,p+5} + e_{2,p+5})P'$$

Where $e_{i,n}$ is the $i'th$ basis vector of length n. This satisfies

$$z_t = Mz_{t-1} + C\epsilon_t$$

where C is a column vector that applies the random shock to Δd_t and is otherwise empty. An alternative evolution equation is

$$z_t = \bar{M}z_{t-1} + Nc_{t-1} + C\epsilon_t$$

We guess that the value function is of the form

$$V(z_t) = -\frac{\Psi}{\alpha}\exp(-\alpha(c_t - \gamma c_{t-1}))$$

Define $\hat{P} = P - \gamma e_{2,p+5}$, and plug into the Bellman:

$$-\frac{\Psi}{\alpha}e^{-\alpha\hat{P}'z_t} = -\frac{1}{\alpha}e^{-\alpha\hat{P}'z_t} + \delta E\left[-\frac{\Psi}{\alpha}e^{-\alpha\hat{P}'z_{t+1}}\right]$$

The expectation for $\hat{P}'z_{t+1}$ is $\hat{P}'Mz_t$, and the variance is $C'\hat{P}\hat{P}'C$. So the bellman simplifies to

$$-\frac{\Psi}{\alpha}e^{-\alpha\hat{P}'z_t} = -\frac{1}{\alpha}e^{-\alpha\hat{P}'z_t} + -\delta\frac{\Psi}{\alpha}e^{-\alpha(\hat{P}'Mz_t - \frac{\alpha}{2}C'\hat{P}\hat{P}'C)}$$

Right now, without worrying about optimality, i'd like to solve the bellman. Dividing through by common terms,

$$\Psi = 1 + \delta\Psi e^{-\alpha(\hat{P}'(M-I)z_t - \frac{\alpha}{2}C'\hat{P}\hat{P}'C)}$$

For this equation to be solved for all z_t, for some constant κ,

$$\hat{P}'(M-I)z_t = \kappa$$

Next, we need to derive an optimality condition for c_t. The first-order condition is that

$$\exp(-\alpha\hat{P}'z_t) + \delta E[\Psi\exp(-\alpha\hat{P}'z_{t+1})\hat{P}'N] = 0$$

Expanding the expectation, and noting that $\hat{P}'N$ is a scalar constant,

$$\exp(-\alpha\hat{P}'z_t) + \hat{P}'^{-\alpha(\hat{P}'Mz_t - \frac{\alpha}{2}C'\hat{P}\hat{P}'C)} = 0$$

Divding through,

$$0 = 1 + \hat{P}'^{-\alpha(\hat{P}'(M-I)z_t - \frac{\alpha}{2}C'\hat{P}\hat{P}'C)}$$

Again, note that if $\hat{P}(M-I)z_t = \kappa$ for all t, this equation can be satisfied for some constant Ψ. Combining the two equations, we can see that

$$0 = 1 + \hat{P}'N(\Psi - 1)$$

Solving,

$$\Psi = 1 - \frac{1}{\hat{P}'N}$$

At this point, we will try to guess P and show that our guess satisfies the equations above. For some constants K and Q,

$$P = \begin{bmatrix} -(R-1)(1-\frac{\gamma}{R}) \\ \frac{\gamma}{R} \\ Q \\ \frac{R-\gamma}{R} \\ \frac{R-\gamma}{R} \\ \left[Ke'_{1,p}(I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi\right]' \end{bmatrix}$$

We can first solve for $\hat{P}'N$.

$$\hat{P}'N = (P - \gamma e_{2,p+5})N = -(R-1)(1-\frac{\gamma}{R}) + \frac{\gamma}{R} - \gamma = 1 - R$$

Therefore,

$$\Psi = 1 - \frac{1}{1-R} = \frac{R}{R-1}$$

Returning to the first order condition,

$$0 = 1 - R\delta e^{-\alpha(\hat{P}'(M-I)z_t - \frac{\alpha}{2}C'\hat{P}\hat{P}'C)}$$

Next, we need to confirm that $\hat{P}(M-I)z_t = \kappa$ for all t.

$$M - I = \bar{M} - I + NP' = \begin{bmatrix} R-1 & 0 & 0 & -1 & -1 & \vec{0} \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & e'_{1,p}\Phi \\ \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \Phi - I \end{bmatrix} + NP'$$

$$NP' = \begin{bmatrix} -(R-1)(1-\frac{\gamma}{R}) & \frac{\gamma}{R} & Q & \frac{R-\gamma}{R} & \frac{R-\gamma}{R} & Ke'_{1,p}(I-\frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi \\ -(R-1)(1-\frac{\gamma}{R}) & \frac{\gamma}{R} & Q & \frac{R-\gamma}{R} & \frac{R-\gamma}{R} & Ke'_{1,p}(I-\frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & 0 \end{bmatrix}$$

$$M - I = \begin{bmatrix} (R-1)\frac{\gamma}{R} & \frac{\gamma}{R} & Q & \frac{-\gamma}{R} & \frac{-\gamma}{R} & Ke'_{1,p}(I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi \\ -(R-1)(1-\frac{\gamma}{R}) & \frac{-R+\gamma}{R} & Q & \frac{R-\gamma}{R} & \frac{R-\gamma}{R} & Ke'_{1,p}(I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & \vec{0}' \\ 0 & 0 & 0 & 0 & 0 & e'_{1,p}\Phi \\ \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \Phi - I \end{bmatrix}$$

$$\hat{P}'(M-I) = \begin{bmatrix} 0 \\ 0 \\ -Q(R-1) \\ 0 \\ 0 \\ -(R-1)Ke_{1,p}(I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi + \frac{R-\gamma}{R}e_{1,p}\Phi + Ke_{1,p}(I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi(\Phi - I) \end{bmatrix}$$

We need to solve for K so that the last element becomes zero. Simplifying,

$$0 = -RKe_{1,p}\left(I - \frac{1}{R}\Phi\right)^{-1}\frac{1}{R}\Phi + \frac{R-\gamma}{R}e_{1,p}\Phi + RKe_{1,p}\left(I - \frac{1}{R}\Phi\right)^{-1}\frac{1}{R}\Phi\frac{1}{R}\Phi$$

Noting that

$$(I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi\frac{1}{R}\Phi = (I - \frac{1}{R}\Phi)^{-1}\frac{1}{R}\Phi - \frac{1}{R}\Phi$$

We can simplify to

$$0 = \frac{R-\gamma}{R}e_{1,p}\Phi - RKe_{1,p}\frac{1}{R}\Phi$$

Solving,

$$K = \frac{R-\gamma}{R} = \left(1 - \frac{\gamma}{R}\right)$$

For this value of K, and any value of Q, $\hat{P}(M-I)z_t = \kappa = -Q(R-1)$ for all t. To solve for Q, we can rewrite the FOC, replacing for our value of Ψ,

$$0 = 1 - \exp\left(-\alpha\left[\frac{-1}{\alpha}\ln(R\delta) - Q(R-1) - \frac{\alpha}{2}C'\hat{P}\hat{P}'C\right]\right)$$

$$\hat{P}'C = \sigma_e\left(1 - \frac{\gamma}{R}\right)e'_{1,p}\left(I - \frac{1}{R}\Phi\right)^{-1}e_{1,p} = \sigma_c$$

$$Q = \frac{1}{R-1}\left[-\frac{1}{\alpha}\ln(R\delta) - \frac{\alpha}{2}\sigma_c^2\right]$$

We have now fully solved for the linear policy function, and shown that it was optimal.

8 Appendix B: Asset pricing

Next, we derive a price for the dividend stream. Our timing convention is that the price at time t does not include the dividend at time t. To calculate the equilibrium price of the

Lucas tree, we consider the asset allocation problem (as opposed to the planner's problem). We then solve for the asset price that leads the representative agent to hold one unit of the equity tree.

Return to the asset allocation problem:

$$V\left(c_{t-1}, w_t, p_t, \overrightarrow{d}_t\right) = \sup_{\theta_t, c_t} u(c_t, c_{t-1}) + E_t \delta V\left(c_t,\, \left(w_t - c_t - \theta_t p_t\right) R + \theta_t \left(d_{t+1} + p_{t+1}\right),\, \overrightarrow{d}_{t+1}\right)$$

since

$$w_{t+1} = \left(w_t - c_t - \theta_t p_t\right) R + \theta_t \left(d_{t+1} + p_{t+1}\right)$$

We can rewrite this as

$$V\left(c_{t-1}, x_t, p_t, \overrightarrow{d}_t\right) = \frac{-1}{\alpha(1-\delta)} \exp\left(-\alpha\left[c_t - \gamma c_{t-1}\right]\right)$$

where

$$c_t = \frac{\gamma}{R} c_{t-1} + \left(1 - \frac{\gamma}{R}\right) x_t - \psi. \tag{7}$$

Consumption is a weighted average of lagged consumption and the (risk-neutral) annuity value of future dividends, x_t, shifted down by an additive constant ψ. Let's write x_t so that we allow the agent to buy more or less of the risky tree. Buying more means raise θ above unity.

$$x_t = \frac{R-1}{R}\left[-Rb_t - (\theta - 1)\, p_t + d_t + \theta \sum_{s=1}^{\infty} \frac{E_t d_{t+s}}{R^s}\right] \tag{8}$$

$$\psi = \frac{1}{R-1}\left[\frac{1}{\alpha}\ln(R\delta) + \frac{\alpha}{2} Var_t(c_{t+1})\right]. \tag{9}$$

In equilibrium, $\theta = 1$. In other words, supply equals demand. The associated first order condition is:

$$\frac{dV}{d\theta} = 0 \text{ at } \theta = 1.$$

We can expand this derivative.

$$\frac{dV}{d\theta} = \frac{\partial V}{\partial c_t}\left[\frac{\partial c_t}{\partial x_t}\frac{\partial x_t}{\partial \theta} + \frac{\partial c_t}{\partial \psi}\frac{\partial \psi}{\partial \theta}\right] = 0$$

which implies that the asset price, p_t, is chosen such that

$$\frac{\partial c_t}{\partial x_t}\frac{\partial x_t}{\partial \theta} + \frac{\partial c_t}{\partial \psi}\frac{\partial \psi}{\partial \theta} = 0 \text{ at } \theta = 1 \tag{10}$$

Let's evaluate each of these partial derivatives in turn:

$$\begin{aligned}
\frac{\partial c_t}{\partial x_t} &= \left(1-\frac{\gamma}{R}\right) \\
\frac{\partial x_t}{\partial \theta} &= \frac{R-1}{R}\left[-p_t+\sum_{s=1}^{\infty}\frac{E_t d_{t+s}}{R^s}\right] \\
\frac{\partial c_t}{\partial \psi} &= -1 \\
\frac{\partial \psi}{\partial \theta_t} &= \frac{1}{R-1}\frac{\alpha}{2}\frac{\partial Var_t(c_{t+1})}{\partial \theta} \\
&= \frac{1}{R-1}\frac{\alpha}{2}2\theta Var_t(c_{t+1})
\end{aligned}$$

Now we are ready to use our equilibrium condition (equation (10)):

$$\begin{aligned}
\frac{\partial c_t}{\partial x}\frac{\partial x}{\partial \theta} &= -\frac{\partial c_t}{\partial \psi}\frac{\partial \psi}{\partial \theta} \quad \text{(evaluated at } \theta=1\text{)} \\
\left(1-\frac{\gamma}{R}\right)\frac{R-1}{R}\left[-p_t+\sum_{s=1}^{\infty}\frac{E_t d_{t+s}}{R^s}\right] &= \frac{1}{R-1}\frac{\alpha}{2}2Var_t(c_{t+1})
\end{aligned}$$

Rearranging we have

$$\begin{aligned}
-p_t+\sum_{s=1}^{\infty}\frac{E_t d_{t+s}}{R^s} &= \frac{R\alpha \times Var_t(c_{t+1})}{\left(1-\frac{\gamma}{R}\right)(R-1)^2} \\
p_t &= \sum_{s=1}^{\infty}\frac{E_t d_{t+s}}{R^s}-\frac{R\alpha \times Var_t(c_{t+1})}{\left(1-\frac{\gamma}{R}\right)(R-1)^2}
\end{aligned}$$

We can also express the expectation of the discounted stream of dividends using our earlier matrix notation, implying that

$$p_t = \frac{1}{R-1}d_t+\frac{R}{R-1}e'_{1,p}\left(I-\frac{1}{R}\Phi\right)^{-1}\frac{1}{R}\Phi\Delta\vec{d_t}-\frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)^2}\sigma_c^2$$

8.1 Equity premium

Assume that $d_t = d$ and $\Delta\vec{d_t} = 0$, so that mistakes don't matter (and note that our theory also has a prediction about the high frequency relationship between $\Delta\vec{d_t}$ and the equity premium). The average quarterly return is $\frac{d}{p_t}$ where

$$p_t = \frac{d}{R-1}-\frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)^2}\sigma_c^2.$$

Hence,

$$\frac{d}{p} = \frac{d}{\frac{d}{R-1} - \frac{R\alpha}{\left(1-\frac{\gamma}{R}\right)(R-1)^2}\sigma_c^2}$$

Therefore the annualized equity premium when dividends have a flat history (i.e., $\Delta\vec{d}_t = 0$) is given by

$$4 \times \left[\frac{d}{\frac{d}{R-1} - \frac{R\alpha}{\left(1-\frac{\gamma}{R}\right)(R-1)^2}\sigma_c^2} - (R-1) \right]$$

9 Appendix C: Rational Expectations Investors in the Consumption Model

9.1 Excess Gains in Natural Expectations

We start with the definition of the price in the natural expectations framework:

$$p_t = \frac{1}{R-1}d_t + \frac{R}{R-1}e'_{1,p}\left(I - \frac{1}{R}\Phi\right)^{-1}\frac{1}{R}\Phi\Delta\vec{d}_t - \frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)^2}\sigma_c^2$$

We would like to understand the gains process, defined as

$$g_{t+1} = p_{t+1} + d_{t+1} - Rp_t$$

We replace the Φ matrix and σ_c with their 'hat' equivalents, to denote misspecification. Expanding the definitions of p_t and p_{t+1}, we find that

$$\begin{aligned} g_{t+1} = {} & d_{t+1} + \frac{1}{R-1}(d_{t+1} - Rd_t) + \\ & \frac{R}{R-1}e'_{1,p}\left(I - \frac{1}{R}\hat{\Phi}\right)^{-1}\frac{1}{R}\hat{\Phi}\left(\Delta\vec{d}_{t+1} - R\Delta\vec{d}_t\right) + \frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)}\hat{\sigma}_c^2 \end{aligned}$$

Define

$$\mu = \frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)}\hat{\sigma}_c^2$$

Using the definition of the AR process driving d_t, we can rewrite this as

$$\begin{aligned} g_{t+1} = {} & \frac{R}{R-1}e'_{1,p}\left(\Phi\Delta\vec{d}_t + \sigma_e\epsilon_{t+1}e_{1,p}\right) + \\ & \frac{R}{R-1}e'_{1,p}\left(I - \frac{1}{R}\hat{\Phi}\right)^{-1}\hat{\Phi}\left(\frac{1}{R}\Phi\Delta\vec{d}_t - \Delta\vec{d}_t + \frac{1}{R}\sigma_e\epsilon_{t+1}e_{1,p}\right) + \mu \end{aligned}$$

Regrouping terms,

$$g_{t+1} = \frac{R}{R-1}e'_{1,p}\left(\Phi - \hat{\Phi}\left(I - \frac{1}{R}\hat{\Phi}\right)^{-1}\left(I - \frac{1}{R}\Phi\right)\right)\Delta\vec{d_t} + \frac{R}{R-1}\sigma_e e'_{1,p}\left(I - \frac{1}{R}\hat{\Phi}\right)^{-1}e_{1,p}\epsilon_{t+1} + \mu$$

We can define the vector M and constant σ_g so that the equation above is

$$g_{t+1} = M\Delta\vec{d_t} + \sigma_g\epsilon_{t+1} + \mu$$

9.2 Budget Constraint

Let w_t be the agent's wealth in period t before consumption is chosen. Assume the agent can hold either risk-free assets with return R, or a risky asset. There are no shorting/leverage constraints. The agent's budget constraint is

$$\begin{aligned} w_{t+1} &= (w_t - c_t - \theta_t p_t)R + \theta_t(d_{t+1} + p_{t+1}) \\ &= (w_t - c_t)R + \theta_t g_{t+1} \end{aligned}$$

where the choice variables c_t and θ_t are the consumption and dollar amount in the risky asset. We can rewrite the budget constraint in terms of the evolution of g_t,

$$w_{t+1} = R(w_t - c_t) + \theta_t M\Delta\vec{d_t} + \theta_t\sigma_g\epsilon_{t+1} + \theta_t\mu.$$

9.3 Utility and Value Functions

The agent has flow utility of the form

$$u(c_t, c_{t-1}) = -\frac{1}{\alpha}e^{-\alpha(c_t - \gamma c_{t-1})}$$

It is convenient to define

$$\hat{c}_t = c_t - \gamma c_{t-1}$$

The state is captured entirely by w_t, c_{t-1}, and $\Delta\vec{d_t}$. The Bellman is

$$V(w_t, c_{t-1}, \Delta\vec{d_t}) = \max_{\hat{c}_t, \theta_t} u(\hat{c}_t) + \delta E[V(w_{t+1}, \hat{c}_t + \gamma c_{t-1}, \Delta\vec{d}_{t+1})]$$

The first order conditions are

$$u'(\hat{c}_t) = \delta E\left[R\frac{\partial V_{t+1}}{\partial w_{t+1}} - \frac{\partial V_{t+1}}{\partial c_t}\right] \tag{11}$$

and

$$\delta E\left[\frac{\partial V_{t+1}}{\partial w_{t+1}}\left(M\Delta\vec{d_t} + \mu + \sigma_g\epsilon_{t+1}\right)\right] = 0 \tag{12}$$

The envelope condition for w_t is

$$\frac{\partial V_t}{\partial w_t} = \delta E_t \left[\frac{\partial V_{t+1}}{\partial w_{t+1}} R \right] \tag{13}$$

From the inter-temporal budget constraint (see 20)

$$\frac{\partial V_t}{\partial c_{t-1}} = \frac{-\gamma}{1 - \frac{\gamma}{R}} \frac{\partial V_t}{\partial w_t} \tag{14}$$

We use this to derive the Euler equations, which will be verified after deriving a solution.

Rewriting (11), and then using (13),

$$\begin{aligned} u'(\hat{c}_t) &= \delta E \left[R \frac{\partial V_{t+1}}{\partial w_{t+1}} + \frac{\gamma}{1 - \frac{\gamma}{R}} \frac{\partial V_{t+1}}{\partial w_{t+1}} \right] \\ \left(1 - \frac{\gamma}{R}\right) u'(\hat{c}_t) &= \delta E \left[R \frac{\partial V_{t+1}}{\partial w_{t+1}} \right] = \frac{\partial V_t}{\partial w_t} \end{aligned}$$

Advancing time by one unit, and taking expectations,

$$\left(1 - \frac{\gamma}{R}\right) E\left[u'(\hat{c}_{t+1})\right] = E\left[\frac{\partial V_{t+1}}{\partial w_{t+1}}\right]$$

The consumption Euler equation, assuming $\delta R = 1$, is therefore

$$u'(\hat{c}_t) = E\left[u'(\hat{c}_{t+1})\right]$$

The asset Euler equation is

$$0 = E\left[\left(1 - \frac{\gamma}{R}\right) u'(\hat{c}_{t+1}) \left(M \Delta \vec{d}_t + \mu + \sigma_g \epsilon_{t+1}\right)\right] \tag{15}$$

9.4 Guess and check

We guess that

$$\hat{c}_t = D c_{t-1} + \left(1 - \frac{\gamma}{R}\right) \left(\frac{R-1}{R} w_t + A' \Delta \vec{d}_t + \Delta \vec{d}_t' B \Delta \vec{d}_t + q \right)$$

where A is a vector and B is a matrix. From this guess, we can see that

$$\hat{c}_{t+1} = D c_t + \left(1 - \frac{\gamma}{R}\right) \left(\frac{R-1}{R} w_{t+1} + A' \Delta \vec{d}_{t+1} + \Delta \vec{d}_{t+1}' B \Delta \vec{d}_{t+1} + q \right)$$

Expanding the budget and evolution of $\Delta\vec{d_t}$,

$$\begin{array}{l}\hat{c}_{t+1} = Dc_t + \\ \left(1-\frac{\gamma}{R}\right)\left[\begin{array}{c}\frac{R-1}{R}\left(R(w_t-c_t)+\theta_t M\Delta\vec{d_t}+\theta_t\mu+\theta_t\sigma_g\epsilon_{t+1}\right)+ \\ A'\Phi\Delta\vec{d_t}+A'e_{1,p}\sigma_e\epsilon_{t+1}+\left(\Phi\Delta\vec{d_t}+e_{1,p}\sigma_e\epsilon_{t+1}\right)'B\left(\Phi\Delta\vec{d_t}+e_{1,p}\sigma_e\epsilon_{t+1}\right)+q\end{array}\right]\end{array}$$

Taking expectations,

$$\begin{array}{l}E_t[\hat{c}_{t+1}] = Dc_t + \\ \left(1-\frac{\gamma}{R}\right)\left[\frac{R-1}{R}\left(R(w_t-c_t)+\theta_t M\Delta\vec{d_t}+\theta_t\mu\right)+A'\Phi\Delta\vec{d_t}+\Phi\Delta\vec{d_t}B\Phi\Delta\vec{d_t}+\sigma_e^2 e'_{1,p}Be_{1,p}+q\right]\end{array}$$

We can then write

$$\begin{array}{rcl}\hat{c}_{t+1} & = & E[\hat{c}_{t+1}]+\left(1-\frac{\gamma}{R}\right) \\ & & \left[-\sigma_e^2 e'_{1,p}Be_{1,p}+\left(\frac{(R-1)}{R}\theta_t\sigma_g+\sigma_e A'e_{1,p}+\sigma_e e'_{1,p}B\Phi\Delta\vec{d_t}+\sigma_e\Delta\vec{d_t}'\Phi'Be_{1,p}\right)\epsilon_{t+1}+\sigma_e^2 e'_{1,p}Be_{1,p}\epsilon_{t+1}^2\right].\end{array}$$

Define the following constants:

$$\begin{array}{rcl}k_t & = & \left(1-\frac{\gamma}{R}\right)\left(\frac{(R-1)}{R}\theta_t\sigma_g+\sigma_e A'e_{1,p}+\sigma_e e'_{1,p}B\Phi\Delta\vec{d_t}+\sigma_e\Delta\vec{d_t}'\Phi'Be_{1,p}\right) \\ \lambda & = & \left(1-\frac{\gamma}{R}\right)\sigma_e^2 e'_{1,p}Be_{1,p}.\end{array}$$

We can rewrite $u'(\hat{c}_{t+1})$ as

$$u'(\hat{c}_{t+1}) = e^{-\alpha\left(E[\hat{c}_{t+1}]-\lambda+k_t\epsilon_{t+1}+\lambda\epsilon_{t+1}^2\right)}.$$

We can now compute expectations.

$$E[u'(\hat{c}_{t+1})] = e^{-\alpha(E[\hat{c}_{t+1}]-\lambda)}\times\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\alpha k_t x-\alpha\lambda x^2}e^{-\frac{x^2}{2}}dx$$

Define a second variable, which is a constant,

$$s=\sqrt{\frac{1}{1+2\alpha\lambda}}$$

We can complete the square on the integral as follows:

$$-\alpha k_t x-\alpha\lambda x^2-0.5x^2 = -\frac{x^2+2\alpha k_t s^2}{2s^2} = -\frac{(x+\alpha k_t s^2)^2}{2s^2}+\frac{\alpha^2}{2}k_t^2 s^2$$

$$E[u'(\hat{c}_{t+1})] = e^{-\alpha\left(E[\hat{c}_{t+1}]-\lambda-\frac{1}{2}\alpha k_t^2 s^2\right)} s \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{\left(x+\alpha k_t s^2\right)^2}{2s^2}} dx$$

We therefore conclude that

$$E[u'(\hat{c}_{t+1})] = e^{-\alpha\left(E[\hat{c}_{t+1}]-\lambda-\frac{1}{2}\alpha k_t^2 s^2-\frac{1}{2\alpha}\ln s^2\right)} \tag{16}$$

Next, consider the integral from the second Euler equation.

$$E[u'(\hat{c}_{t+1})\sigma_g \epsilon_{t+1}] = e^{-\alpha\left(E[\hat{c}_{t+1}]-\lambda-\frac{1}{2}\alpha k_t^2 s^2\right)} \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \sigma_g x e^{-\frac{\left(x+\alpha k_t s^2\right)^2}{2s^2}} dx$$

For the integral, use the following transformation:

$$u = \frac{x + \alpha k_t s^2}{s}$$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \sigma_g x e^{-\frac{\left(x+\alpha k_t s^2\right)^2}{2s^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \sigma_g (s\,u - \alpha k_t s^2) e^{-\frac{u^2}{2}} s du = -\alpha \sigma_g k_t s^3$$

We combine 15 and the equation above to see that

$$e^{-\alpha\left(E_t[\hat{c}_{t+1}]-\lambda-\frac{1}{2}\alpha k_t^2 s^2-\frac{1}{2\alpha}\ln s^2\right)}(\mu + M\Delta\vec{d}_t) = e^{-\alpha\left(E_t[\hat{c}_{t+1}]-\lambda-\frac{1}{2}\alpha k_t^2 s^2\right)}\alpha\sigma_g k_t s^3$$

Simplifying,

$$\mu + M\Delta\vec{d}_t = \alpha\sigma_g k_t s^2$$

9.5 Solving for the policy rules

Next, we use consumption Euler equation to see that

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - \lambda - \frac{1}{2}\alpha k_t^2 s^2 - \frac{1}{2\alpha}\ln s^2$$

From our conjecture about $\hat{c}_t$, we can expand $\hat{c}_{t+1}$:

$$\hat{c}_t + \lambda + \frac{1}{2}\alpha k_t^2 s^2 \frac{1}{2\alpha}\ln s^2 = \left(1-\frac{\gamma}{R}\right)\left(\frac{(R-1)}{R}E_t[w_{t+1}] + A' E_t[\Delta\vec{d}_{t+1}] + E_t[\Delta\vec{d}'_{t+1} B \Delta\vec{d}_{t+1}] + q\right) + Dc_t$$

$$\hat{c}_t + \lambda + \frac{1}{2}\alpha k_t^2 s^2 + \frac{1}{2\alpha}\ln s^2 = \left(1-\frac{\gamma}{R}\right)\left(\frac{(R-1)}{R}E_t[w_{t+1}] + A'\Phi\Delta\vec{d}_t + \lambda + \Delta\vec{d}'_t\Phi' B \Phi\Delta\vec{d}_t + q\right) + Dc_t$$

Expanding using the budget constraint,

$$\hat{c}_t - Dc_t + \frac{\gamma}{R}\lambda + \frac{1}{2}\alpha k_t^2 s^2 + \frac{1}{2\alpha}\ln s^2 = \\ \left(1-\frac{\gamma}{R}\right)\left(\frac{R-1}{R}E_t[g_{t+1}]\theta_t + (R-1)(w_t - c_t) + A'\Phi\Delta\vec{d}_t + \Delta\vec{d}_t'\Phi' B\Phi\Delta\vec{d}_t + q\right)$$

Simplifying, and replacing c_t with $\hat{c}_t + \gamma c_{t-1}$,

$$\hat{c}_t\left(1 - D + \left(1-\frac{\gamma}{R}\right)(R-1)\right) + \left(\left(1-\frac{\gamma}{R}\right)(R-1) - D\right)\gamma c_{t-1} + \frac{\gamma}{R}\lambda + \frac{1}{2}\alpha k_t^2 s^2 + \frac{1}{2\alpha}\ln s^2 = \\ \left(1-\frac{\gamma}{R}\right)\left(\frac{R-1}{R}(M\Delta\vec{d}_t + \mu)\theta_t + (R-1)w_t + A'\Phi\Delta\vec{d}_t + \Delta\vec{d}_t'\Phi' B\Phi\Delta\vec{d}_t + q\right)$$

Let $Q = 1 - D + (1-\frac{\gamma}{R})(R-1)$. Simplifying, and expanding the terms of $\hat{c}_t$,

$$((Q-1)\gamma + QD)c_{t-1} + Q\left(1-\frac{\gamma}{R}\right)\left(\frac{R-1}{R}w_t + A'\Delta\vec{d}_t + Q\Delta\vec{d}_t' B\Delta\vec{d}_t + Qq\right) + \frac{\gamma}{R}\lambda + \frac{1}{2}\alpha k_t^2 s^2 + \frac{1}{2\alpha}\ln s^2 = \\ \left(1-\frac{\gamma}{R}\right)\left((R-1)w_t + \frac{R-1}{R}(M\Delta\vec{d}_t + \mu)\theta_t + A'\Phi\Delta\vec{d}_t + \Delta\vec{d}_t'\Phi' B\Phi\Delta\vec{d}_t + q\right)$$

We guess that $D = \frac{\gamma}{R} - \gamma$, and see that this simplifies nicely:

$$\begin{aligned} Q &= R \\ (Q-1)\gamma + QD &= 0 \end{aligned}$$

Therefore, noting that w_t cancels,

$$R\left(1-\frac{\gamma}{R}\right)\left(A'\Delta\vec{d}_t + \Delta\vec{d}_t' B\Delta\vec{d}_t + q\right) + \frac{\gamma}{R}\lambda + \frac{1}{2}\alpha k_t^2 s^2 + \frac{1}{2\alpha}\ln s^2 = \\ \left(1-\frac{\gamma}{R}\right)\left(\frac{R-1}{R}(M\Delta\vec{d}_t + \mu)\theta_t + A'\Phi\Delta\vec{d}_t + \Delta\vec{d}_t'\Phi' B\Phi\Delta\vec{d}_t + q\right)$$

Using the asset allocation Euler result, we see that

$$\frac{1}{2}\alpha k_t^2 s^2 = \frac{1}{2\alpha\sigma_g^2 s^2}\left(\mu^2 + 2\mu M\Delta\vec{d}_t + \Delta\vec{d}_t' M'M\Delta\vec{d}_t\right)$$

Expanding using the definition of s,

$$\frac{1}{2}\alpha k_t^2 s^2 = \frac{1+2\alpha\lambda}{2\alpha\sigma_g^2}\left(\mu^2 + 2\mu M\Delta\vec{d}_t + \Delta\vec{d}_t' M'M\Delta\vec{d}_t\right) \tag{17}$$

We also need to replace θ_t using the asset allocation Euler.

$$\mu + M\Delta\vec{d}_t = \alpha\sigma_g s^2 k_t = \left(1-\frac{\gamma}{R}\right)\alpha\sigma_g s^2\left(\frac{(R-1)}{R}\theta_t\sigma_g + \sigma_e A'e_{1,p} + \sigma_e e_{1,p}' B\Phi\Delta\vec{d}_t + \sigma_e\Delta\vec{d}_t'\Phi' Be_{1,p}\right)$$

Solving,

$$\theta_t = \frac{R}{R-1}\frac{1}{\sigma_g}\left(\frac{\mu + M\Delta\vec{d_t}}{(1-\frac{\gamma}{R})\alpha\sigma_g s^2} - \sigma_e A' e_{1,p} - \sigma_e e'_{1,p} B\Phi\Delta\vec{d_t} - \sigma_e \Delta\vec{d_t}'\Phi' B e_{1,p}\right)$$

To replace the term in the consumption equation,

$$\frac{R-1}{R}(M\Delta\vec{d_t} + \mu)\theta_t =$$
$$\frac{1}{\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g^2 s^2}\left(\mu^2 + 2\mu M\Delta\vec{d_t} + \Delta\vec{d_t}' M' M\Delta\vec{d_t}\right) -$$
$$\frac{\mu\sigma_e}{\sigma_g}A' e_{1,p} - \frac{\sigma_e}{\sigma_g}\left(A' e_{1,p} M + \mu e'_{1,p}(B+B')\Phi\right)\Delta\vec{d_t} - \frac{\sigma_e}{\sigma_g}\Delta\vec{d_t}' M' e'_{1,p} B\Phi\Delta\vec{d_t} - \frac{\sigma_e}{\sigma_g}\Delta\vec{d_t}'\Phi' B e_{1,p} M\Delta\vec{d_t}$$

We can now substitute all of these results:

$$R\left(1-\frac{\gamma}{R}\right)\left(A'\Delta\vec{d_t} + \Delta\vec{d_t}' B\Delta\vec{d_t} + q\right) + \frac{\gamma}{R}\lambda + \frac{1}{2}\alpha k_t^2 s^2 + \frac{1}{2\alpha}\ln s^2 = \alpha k_t^2 s^2 +$$
$$\left(1-\frac{\gamma}{R}\right)\left(-\frac{\mu\sigma_e}{\sigma_g}A' e_{1,p} - \frac{\sigma_e}{\sigma_g}(A' e_{1,p} M + \mu e'_{1,p}(B+B')\Phi)\Delta\vec{d_t} - \frac{\sigma_e}{\sigma_g}\Delta\vec{d_t}' M' e'_{1,p} B\Phi\Delta\vec{d_t} - \frac{\sigma_e}{\sigma_g}\Delta\vec{d_t}'\Phi' B e_{1,p} M\Delta\vec{d_t}\right) +$$
$$\left(1-\frac{\gamma}{R}\right)\left(A'\Phi\Delta\vec{d_t} + \Delta\vec{d_t}'\Phi' B\Phi\Delta\vec{d_t} + q\right)$$

The final version of the system is:

$$R\left(1-\frac{\gamma}{R}\right)(A'\Delta\vec{d_t} + \Delta\vec{d_t}' B\Delta\vec{d_t} + q) + \frac{\gamma}{R}\lambda + \frac{1}{2\alpha}\ln s^2 =$$
$$\frac{1+2\alpha\lambda}{2\alpha\sigma_g^2}(\mu^2 + 2\mu M\Delta\vec{d_t} + \Delta\vec{d_t}' M' M\Delta\vec{d_t}) +$$
$$\left(1-\frac{\gamma}{R}\right)\left(-\frac{\mu\sigma_e}{\sigma_g}A' e_{1,p} - \frac{\sigma_e}{\sigma_g}\left(A' e_{1,p} M + \mu e'_{1,p}(B+B')\Phi\right)\Delta\vec{d_t} + \Delta\vec{d_t}'(\Phi - \frac{\sigma_e}{\sigma_g}e_{1,p} M)' B\Phi\Delta\vec{d_t}\right)$$
$$+\left(1-\frac{\gamma}{R}\right)\left(\Delta\vec{d_t}'\Phi' B(\Phi - \frac{\sigma_e}{\sigma_g}e_{1,p} M)\Delta\vec{d_t} - \Delta\vec{d_t}'\Phi' B\Phi\Delta\vec{d_t} + A'\Phi\Delta\vec{d_t} + q\right) \qquad (18)$$

This equation must hold for all value of $\Delta\vec{d_t}$, so we use term matching. Note that λ is actually the upper left element of B, scaled by $(1-\frac{\gamma}{R})\sigma_e^2$. Beginning with the 2nd order terms,

$$\Delta\vec{d_t}'\left(RB - \frac{1+2\alpha\lambda}{2\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g^2}M' M - (\Phi - \frac{\sigma_e}{\sigma_g}M e_{1,p})' B\Phi - \Phi' B(\Phi - \frac{\sigma_e}{\sigma_g}M e_{1,p}) + \Phi' B\Phi\right)\Delta\vec{d_t} = 0$$

Define

$$\Lambda = \Phi - \hat{\Phi}$$

and note that

$$M = \frac{R}{R-1} e'_{1,p} \left[\Phi - \hat{\Phi} \left(I - \frac{1}{R} \hat{\Phi} \right)^{-1} \left(I - \frac{1}{R} \Phi \right) \right]$$
$$= \frac{R}{R-1} e'_{1,p} \left[\Lambda + \hat{\Phi} - \hat{\Phi} \left(I - \frac{1}{R} \hat{\Phi} \right)^{-1} \left(I - \frac{1}{R} \hat{\Phi} - \frac{1}{R} \Lambda \right) \right]$$
$$= \frac{R}{R-1} e'_{1,p} \left[\Lambda + \hat{\Phi} \left(I - \frac{1}{R} \hat{\Phi} \right)^{-1} \frac{1}{R} \Lambda \right]$$
$$= \frac{R}{R-1} e'_{1,p} \left(I - \frac{1}{R} \hat{\Phi} \right)^{-1} \Lambda$$

Because Λ has non-zeros only in the top row, and $e_{1,p}e'_{1,p}$ has non-zeros only in the upper left element, which is one,

$$\Lambda = e_{1,p} e'_{1,p} \Lambda$$

Therefore,

$$e_{1,p} M = \frac{R}{R-1} e_{1,p} e'_{1,p} \left(I - \frac{1}{R} \hat{\Phi} \right)^{-1} e_{1,p} e'_{1,p} \Lambda$$
$$= \frac{\sigma_g}{\sigma_e} \Lambda$$

and

$$\Phi - \frac{\sigma_e}{\sigma_g} M e_{1,p} = \Phi - \Lambda = \hat{\Phi}$$

The equation for B can be rewritten:

$$\Delta \vec{d}'_t \left(RB - \frac{1 + 2\alpha\lambda}{2\left(1 - \frac{\gamma}{R}\right)\alpha\sigma_g^2} M'M - \hat{\Phi}' B \Phi - \Phi' B \hat{\Phi} + \Phi' B \Phi \right) \Delta \vec{d}_t = 0$$

Substituting for λ and regrouping terms,

$$\Delta \vec{d}'_t \left(RB - \frac{1}{2\left(1 - \frac{\gamma}{R}\right)\alpha\sigma_g^2} M'M - \frac{\sigma_e^2}{\sigma_g^2} M' e'_{1,p} B e_{1,p} M - \hat{\Phi}' B \hat{\Phi} + \Lambda' B \Lambda \right) \Delta \vec{d}_t = 0$$

By our earlier result relating $e_{1,p}M$ and Λ, and assuming that this holds for all $\Delta \vec{d}_t$,

$$B = \frac{1}{2R\left(1 - \frac{\gamma}{R}\right)\alpha\sigma_g^2} M'M + \frac{1}{R} \hat{\Phi}' B \hat{\Phi}$$

This is a discrete time Lyapunov equation. We can apply the standard convergence results to see that, because the eigenvalues of $\frac{1}{\sqrt{R}}\hat{\Phi}$ are entirely less than 1 (no unit root), convergence is certain. Therefore,

$$B = \frac{1}{2R(1-\frac{\gamma}{R})\alpha\sigma_g^2}\sum_{k=0}^{\infty} R^{-k}\hat{\Phi}'^k M' M \hat{\Phi}^k$$

From this, we can easily solve for

$$\lambda = \left(1-\frac{\gamma}{R}\right)\sigma_e^2 e'_{1,p} B e_{1,p}$$

Moving on to the 1st order terms in equation 18, and using the symmetry of B,

$$\left(R\left(1-\frac{\gamma}{R}\right)A' - \frac{1+2\alpha\lambda}{2\alpha\sigma_g^2}2\mu M - \left(1-\frac{\gamma}{R}\right)A'\Phi + \left(1-\frac{\gamma}{R}\right)\frac{\sigma_e}{\sigma_g}(e'_{1,p}AM + 2\mu e'_{1,p}B\Phi)\right)\Delta\vec{d}_t = 0$$

Subsituting for λ, and regrouping,

$$\left(RA' - \frac{1}{2\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g^2}2\mu M - \frac{2\mu\sigma_e}{\sigma_g}\frac{\sigma_e}{\sigma_g}e'_{1,p}Be_{1,p}M - A'\Phi + \frac{\sigma_e}{\sigma_g}A'e_{1,p}M + \frac{2\mu\sigma_e}{\sigma_g}e'_{1,p}B\Phi\right)\Delta\vec{d}_t = 0$$

Again using the relation between $e_{1,p}M$ and Λ,

$$\left(RA'(I-\frac{1}{R}\hat{\Phi}) - \frac{1}{2\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g^2}2\mu M + \frac{2\mu\sigma_e}{\sigma_g}e'_{1,p}B\hat{\Phi}\right)\Delta\vec{d}_t = 0$$

This is solved by inversion:

$$A' = \frac{2\mu}{R\sigma_g}\left(\frac{1}{2(1-\frac{\gamma}{R})\alpha\sigma_g}M + \sigma_e e'_{1,p}B\hat{\Phi}\right)\left(I-\frac{1}{R}\hat{\Phi}\right)^{-1}$$

Finally, we solve the constants.

$$(R-1)\left(1-\frac{\gamma}{R}\right)q + \frac{\gamma}{R}\lambda - \frac{1+2\alpha\lambda}{2\alpha\sigma_g^2}\mu^2 + \left(1-\frac{\gamma}{R}\right)\frac{\mu\sigma_e}{\sigma_g}A'e_{1,p} + \frac{1}{2\alpha}\ln s^2 = 0$$

Consequently,

$$q = \frac{1+2\alpha\lambda}{2(R-1)(1-\frac{\gamma}{R})\alpha\sigma_g^2}\mu^2 - \frac{\mu\sigma_e}{\sigma_g(R-1)}A'e_{1,p} - \frac{1}{2\alpha(R-1)(1-\frac{\gamma}{R})}\ln(1+2\alpha\lambda).$$

At this point, it is worth simplifying the expression for asset allocation.

$$\theta_t = \frac{R}{R-1}\frac{1}{\sigma_g}\left(\frac{\mu + M\Delta\vec{d}_t}{\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g s^2} - \sigma_e A'e - 2\sigma_e e'B\Phi\Delta\vec{d}_t\right)$$

Applying the standard substitution for s and then λ,

$$\theta_t = \frac{R}{R-1}\frac{1}{\sigma_g}\left(\frac{\mu + M\Delta\vec{d_t}}{\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g} + \frac{2\mu\sigma_e^2}{\sigma_g}e'_{1,p}Be_{1,p} - \sigma_e A'e - 2\sigma_e e'_{1,p}B\hat{\Phi}\Delta\vec{d_t}\right)$$

The average value of θ (assuming $\Delta\vec{d_t} = 0$) is

$$\bar{\theta} = \frac{R}{R-1}\frac{\mu}{\left(1-\frac{\gamma}{R}\right)\alpha\sigma_g^2} + \frac{R}{R-1}\frac{2\mu\sigma_e^2}{\sigma_g^2}e'_{1,p}Be_{1,p} - \frac{R}{R-1}\frac{\sigma_e}{\sigma_g}A'e$$

Recall that

$$\mu = \frac{R\alpha}{(1-\frac{\gamma}{R})(R-1)}\hat{\sigma}_c^2$$
$$\hat{\sigma}_c = \frac{R-1}{R}(1-\gamma/R)\frac{\hat{\sigma}_e}{\sigma_e}\sigma_g$$

We can rewrite the average asset allocation as

$$\bar{\theta} = \frac{\hat{\sigma}_e^2}{\sigma_e^2} + \frac{R}{R-1}\frac{2\mu\sigma_e^2}{\sigma_g^2}e'_{1,p}Be_{1,p} - \frac{R}{R-1}\frac{\sigma_e}{\sigma_g}A'e \tag{19}$$

9.6 Inter-temporal Budget

It will be helpful to solve for an inter-temporal budget constraint. The transversality condition is

$$\lim_{k\to\infty} E_t[R^{-k}w_{t+k}] = 0$$

Seeing that

$$\begin{aligned} E[w_{t+2}] &= E[R(w_{t+1} - c_{t+1} + y) + \theta_{t+1}M\Delta\vec{d}_{t+1} + \theta_{t+1}\mu] \\ &= E[R^2(w_t - c_t + y) + R(-c_{t+1} + y)\theta_{t+1}M\Delta\vec{d}_{t+1} + \theta_{t+1}\mu + R\theta_t M\Delta\vec{d_t} + R\mu] \end{aligned}$$

We conjecture that

$$E_t[R^{-k}w_{t+k}] = w_t + \frac{1}{R}\sum_{j=0}^{k} R^{-j}E_t[\theta_{t+j}M\Delta\vec{d}_{t+j} + \theta_{t+j}\mu - Rc_{t+k} + Ry]$$

This holds trivially for $k = 0$ and $k = 1$. Assume it holds for k.

$$\begin{aligned} E[R^{-k-1}w_{t+k+1}] &= E[R^{-k}w_{t+k} - R^{-k}c_{t+k} + R^{-k}y + R^{-k-1}\theta_{t+k}M\Delta\vec{d}_{t+k} + R^{-k-1}\theta_{t+k}\mu] \\ &= w_t + \frac{1}{R}\sum_{j=0}^{k+1} R^{-j}E_t[\theta_{t+j}M\Delta\vec{d}_{t+j} + \theta_{t+j}\mu - Rc_{t+k} + Ry] \end{aligned}$$

By induction, it holds in the limit, and therefore the inter-temporal budget constraint is

$$\sum_{j=0}^{\infty} E_t[R^{-j} c_{t+j}] = w_t + \frac{1}{R-1} y + \sum_{j=0}^{\infty} E_t[R^{-j} \theta_{t+j} M \Delta \vec{d}_{t+j} + R^{-j} \theta_{t+j} \mu]$$

We can rewrite the sum of c_{t+j} in terms of $\hat{c}_{t+j} = c_{t+j} - \gamma c_{t+j-1}$.

$$\begin{aligned}
\sum_{j=0}^{\infty} E_t[R^{-j} c_{t+j}] &= \sum_{j=0}^{\infty} E_t[R^{-j}(\hat{c}_{t+j} + \gamma c_{t+j-1})] \\
&= \gamma c_{t-1} + \sum_{j=0}^{\infty} E_t[R^{-j} \hat{c}_{t+j}] + \frac{\gamma}{R} \sum_{j=1}^{\infty} E_t[R^{-j+1} c_{t+j-1}] \\
&= \gamma c_{t-1} + \sum_{j=0}^{\infty} E_t[R^{-j} \hat{c}_{t+j}] + \frac{\gamma}{R} \sum_{k=0}^{\infty} E_t[R^{-k} c_{t+k}]
\end{aligned}$$

We can then solve to see that

$$\sum_{j=0}^{\infty} E_t[R^{-j} c_{t+j}] = \frac{1}{1 - \frac{\gamma}{R}} (\gamma c_{t-1} + \sum_{j=0}^{\infty} E_t[R^{-j} \hat{c}_{t+j}])$$

Rewriting the inter-temporal budget constraint,

$$\sum_{j=0}^{\infty} E_t[R^{-j} \hat{c}_{t+j}] = -\gamma c_{t-1} + (1 - \frac{\gamma}{R}) \left(w_t + \frac{1}{R-1} y + \sum_{j=0}^{\infty} E_t[R^{-j} \theta_{t+j} M \Delta \vec{d}_{t+j} + R^{-j} \theta_{t+j} \mu] \right) \tag{20}$$

Table 1
Estimated coefficients from AR(p)

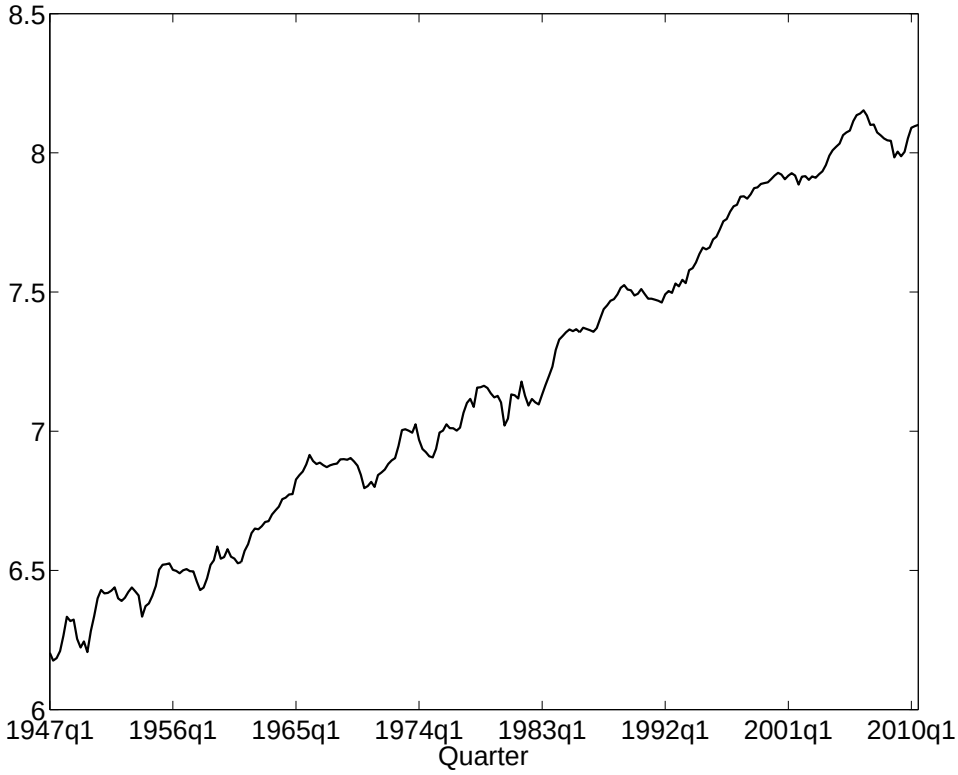
	AR(1)		AR(10)		AR(20)		AR(30)		AR(40)	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
1st lag	0.2011	3.27	0.1841	2.83	0.1894	2.80	0.18	2.57	0.2113	2.79
2nd lag			0.0319	0.48	0.0159	0.23	0.01	0.15	-0.0493	-0.64
3rd lag			0.1242	1.93	0.1121	1.64	0.16	2.19	0.1969	2.57
4th lag			-0.1009	-1.57	-0.0801	-1.16	-0.08	-1.11	-0.1282	-1.64
5th lag			-0.1487	-2.31	-0.1622	-2.36	-0.19	-2.68	-0.1057	-1.35
6th lag			-0.0904	-1.41	-0.0276	-0.40	-0.04	-0.62	-0.0880	-1.13
7th lag			0.1181	1.86	0.0626	0.90	0.05	0.68	0.0682	0.87
8th lag			-0.1318	-2.04	-0.1718	-2.42	-0.15	-2.03	-0.2069	-2.59
9th lag			0.0853	1.31	0.0432	0.62	0.04	0.55	0.0755	0.94
10th lag			-0.0130	-0.20	0.0144	0.21	-0.05	-0.69	-0.0826	-1.02
11th lag					-0.0158	-0.23	-0.02	-0.27	-0.0047	-0.06
12th lag					-0.1696	-2.48	-0.15	-2.07	-0.1831	-2.32
13th lag					-0.0227	-0.33	0.00	0.05	0.0327	0.41
14th lag					-0.0099	-0.15	-0.03	-0.43	-0.0642	-0.82
15th lag					-0.0573	-0.85	-0.03	-0.34	-0.0231	-0.30
16th lag					0.0208	0.31	-0.03	-0.35	-0.0592	-0.75
17th lag					-0.0580	-0.89	-0.07	-0.99	-0.0873	-1.12
18th lag					0.0877	1.34	0.01	0.19	0.0039	0.05
19th lag					-0.0231	-0.35	-0.05	-0.74	-0.0624	-0.81
20th lag					-0.0908	-1.41	-0.06	-0.88	-0.0705	-0.92
21st lag							-0.01	-0.08	-0.0062	-0.08
22nd lag							-0.10	-1.48	-0.1343	-1.76
23rd lag							0.06	0.81	0.0944	1.23
24th lag							-0.04	-0.58	-0.0541	-0.70
25th lag							-0.05	-0.68	-0.0352	-0.46
26th lag							-0.06	-0.90	-0.0876	-1.15
27th lag							0.02	0.29	0.0317	0.41
28th lag							0.05	0.71	-0.0012	-0.02
29th lag							-0.02	-0.31	-0.0109	-0.15
30th lag							-0.17	-2.62	-0.2144	-2.95
31st lag									0.0547	0.74
32nd lag									-0.1021	-1.39
33rd lag									0.0556	0.78
34th lag									-0.0994	-1.39
35th lag									0.1435	2.01
36th lag									-0.1413	-1.95
37th lag									-0.0149	-0.21
38th lag									-0.1210	-1.71
39th lag									0.0858	1.20
40th lag									-0.0630	-0.90
N	253		244		234		224		214	
R squared	0.0407		0.1244		0.1658		0.2071		0.2483	
Root MSE	0.0249		0.0236		0.0228		0.0225		0.0227	

Table 2: Bootstrap confidence intervals under the null of no predictability

Moment	US data	*Bootstrap* Mean	90% C.I.	95% C.I.
1) corr(XR_τ, $XR_{\tau+2\to\tau+5}$)	-0.220	-0.049	$[-0.249, 0.165]$	$[-0.284, 0.206]$
2) corr($(P/E_{10})_\tau$, $XR_{\tau+2\to\tau+5}$)	-0.377	-0.141	$[-0.432, 0.192]$	$[-0.478, 0.249]$
3) corr($ln(c_\tau) - ln(c_{\tau-1})$, $XR_{\tau+2\to\tau+5}$)	-0.297	-0.009	$[-0.267, 0.254]$	$[-0.313, 0.297]$
4) corr($(P/E_{10})_\tau$, $ln(c_{\tau+6}) - ln(c_{\tau+2})$)	-0.167	-0.022	$[-0.424, 0.388]$	$[-0.493, 0.450]$
5) corr($ln(c_\tau) - ln(c_{\tau-1})$, $ln(c_{\tau+6}) - ln(c_{\tau+2})$)	-0.245	0.008	$[-0.237, 0.257]$	$[-0.283, 0.302]$

Note: Sources for empirical moments and bootstrap procedures are described in the text.

Figure 1: Log of Real Net Operating Surplus ("Earnings") of U.S. Private Enterprises from 1947Q1 to 2010Q3



Source: U.S. National Income and Product Accounts (Bureau of Economic Analysis), Table 1.10, line 12. Adjusted for inflation using the GDP deflator.

Figure 2: Estimated Impulse Response Function for Earnings (in Levels), for Different AR(p) Models of Earnings Growth

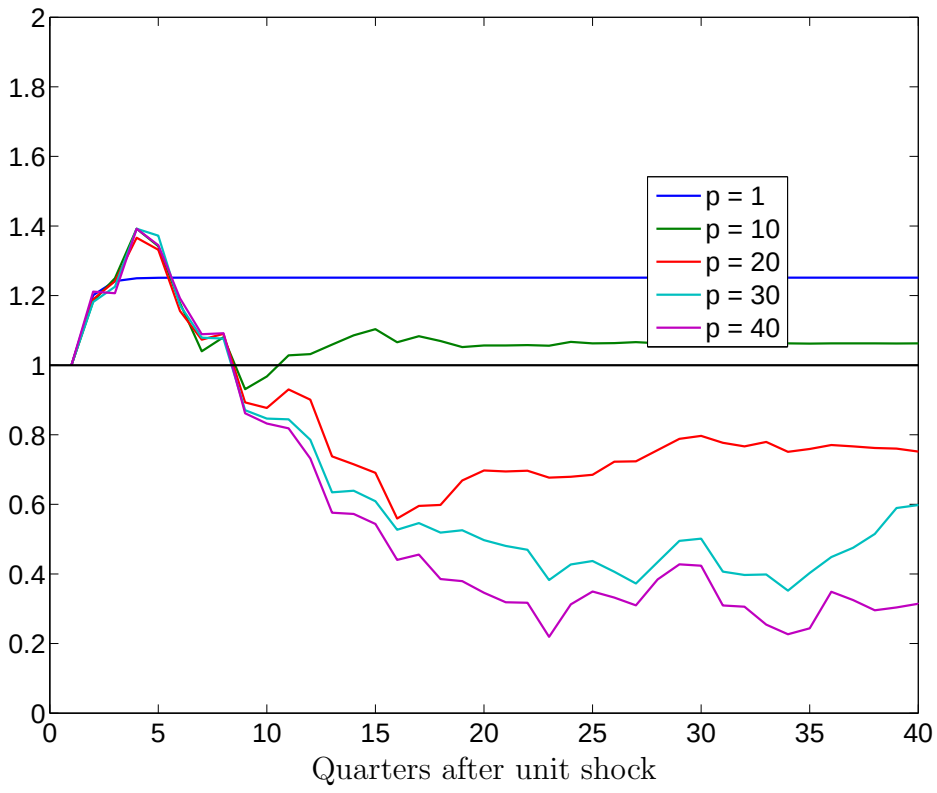


Figure 3: Cumulative Impulse Response Function for Gains, for Different AR(p) Models of Earnings Growth

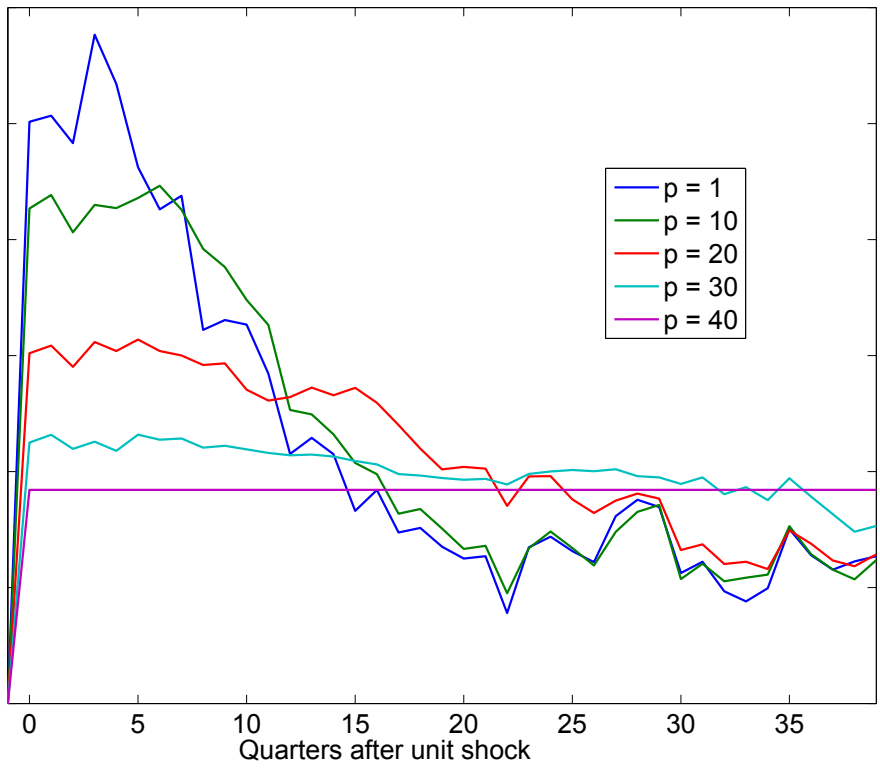


Figure 4: Impulse Response Function for Consumption, for Different AR(p) Models of Earnings Growth

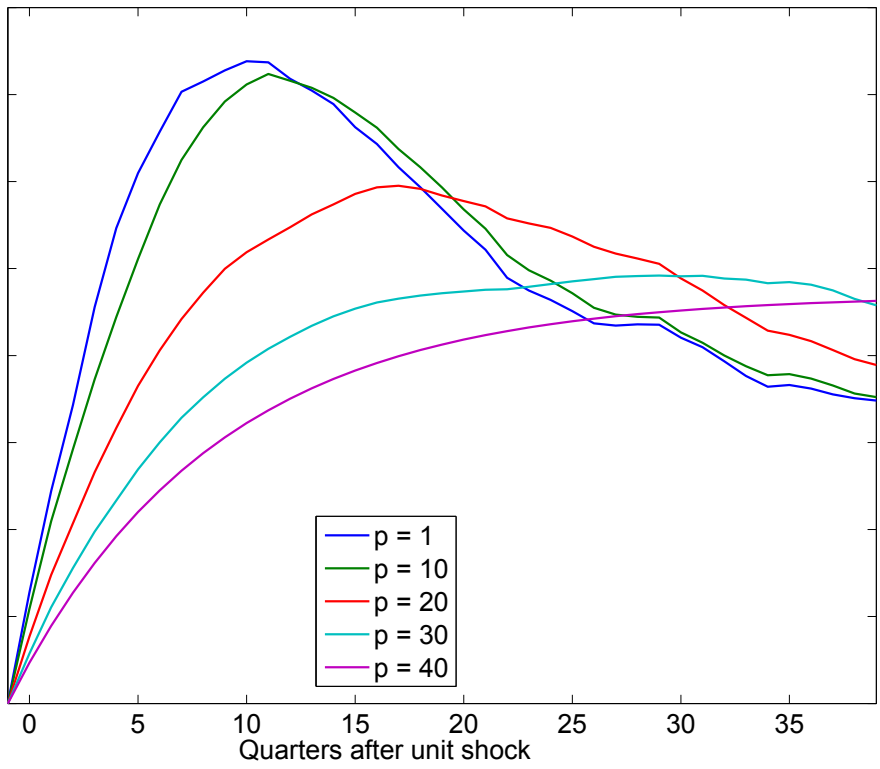
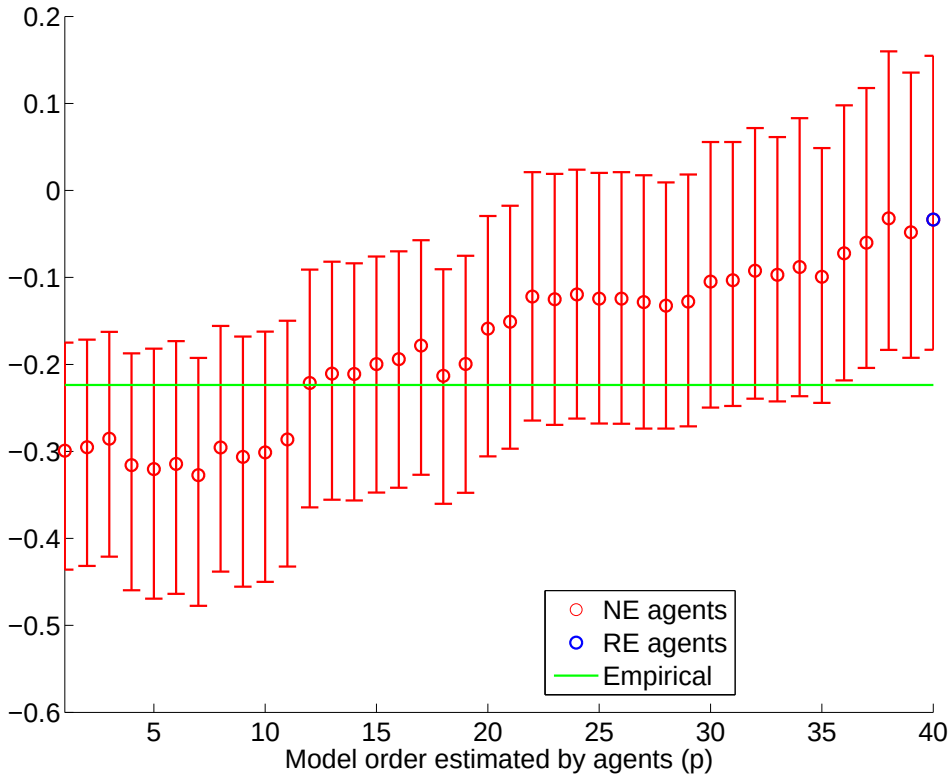
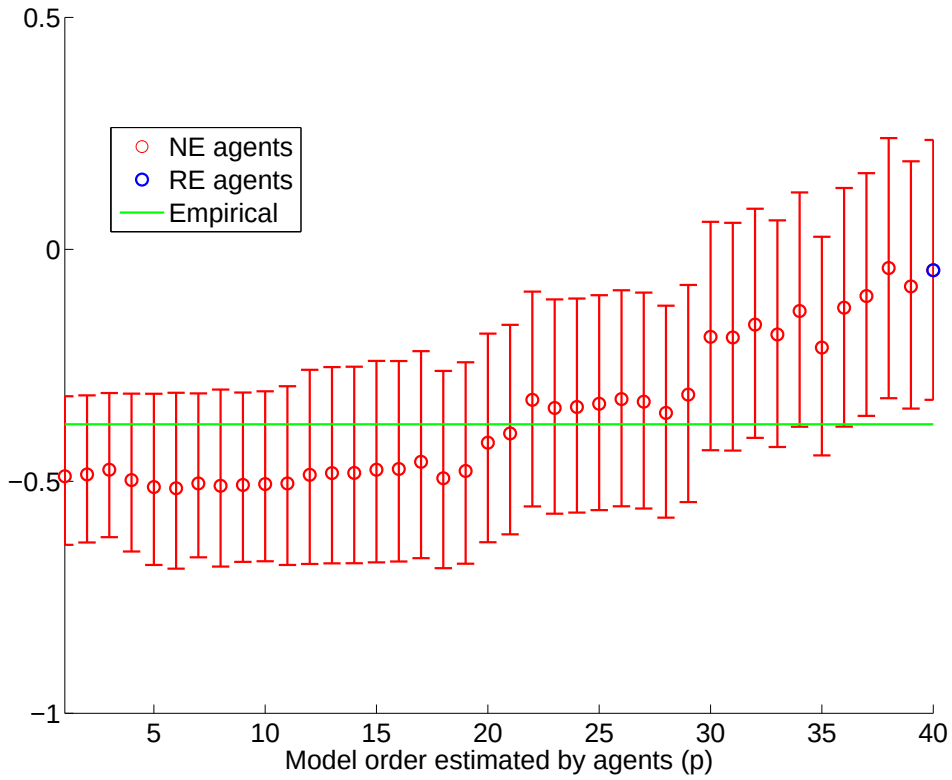


Figure 5: Correlation of Excess Returns in Year τ with Cumulative Excess Returns for Years $\tau + 2$ to $\tau + 5$, for Different AR(p) Models of Earnings Growth



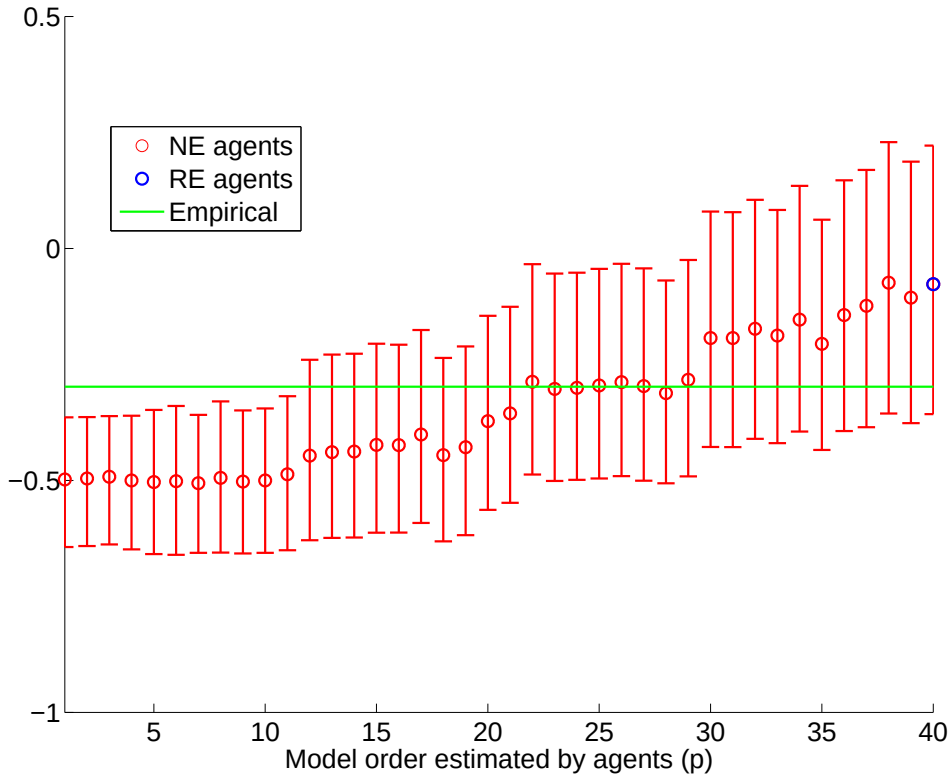
Note: Circles depict mean correlations over 200 simulations of length equal to the empirical sample, while bars depict the 5th and 95th percentile of simulated values. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 6: Correlation of the P/E_{10} Ratio in Year τ with Cumulative Excess Returns for Years $\tau + 2$ to $\tau + 5$, for Different AR(p) Models of Earnings Growth



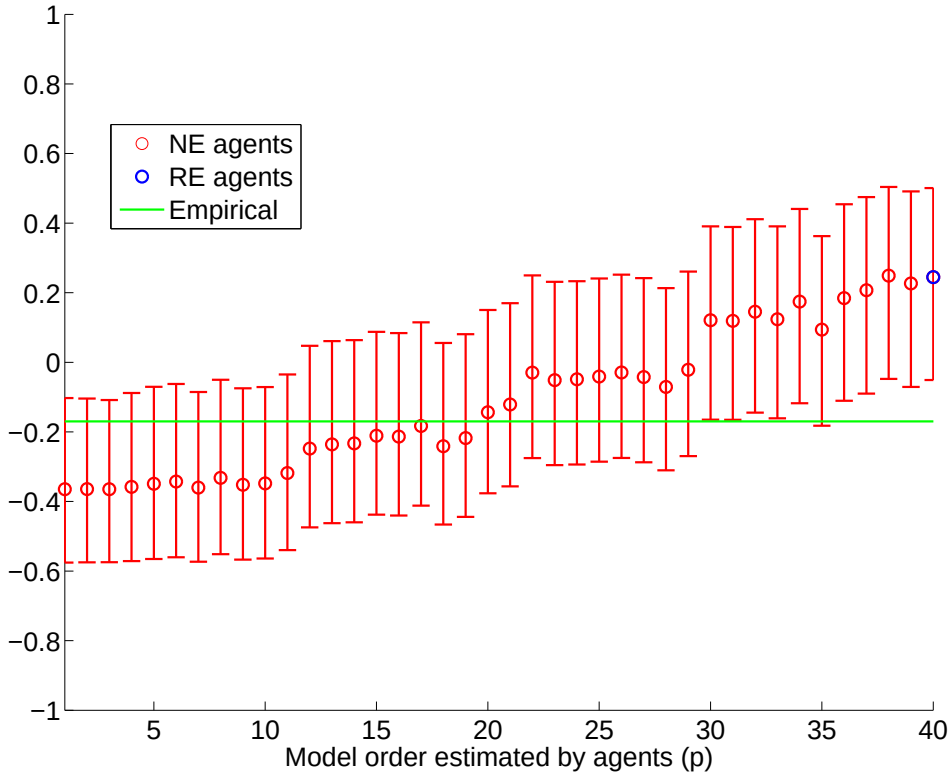
Note: Circles depict mean correlations over 200 simulations of length equal to the empirical sample, while bars depict the 5th and 95th percentile of simulated values. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 7: Correlation of Consumption Growth in Year τ with Cumulative Excess Returns for Years $\tau + 2$ to $\tau + 5$, for Different AR(p) Models of Earnings Growth



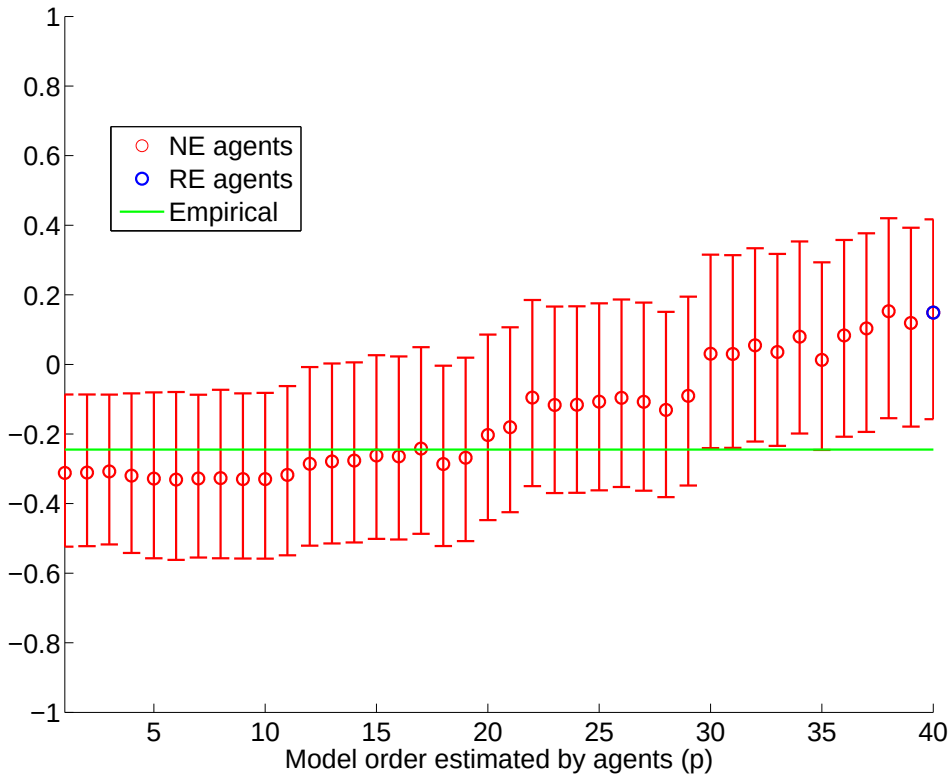
Note: Circles depict mean correlations over 200 simulations of length equal to the empirical sample, while bars depict the 5th and 95th percentile of simulated values. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 8: Correlation of the P/E_{10} Ratio in Year τ with Consumption Growth from Year $\tau + 2$ to $\tau + 6$, for Different AR(p) Models of Earnings Growth



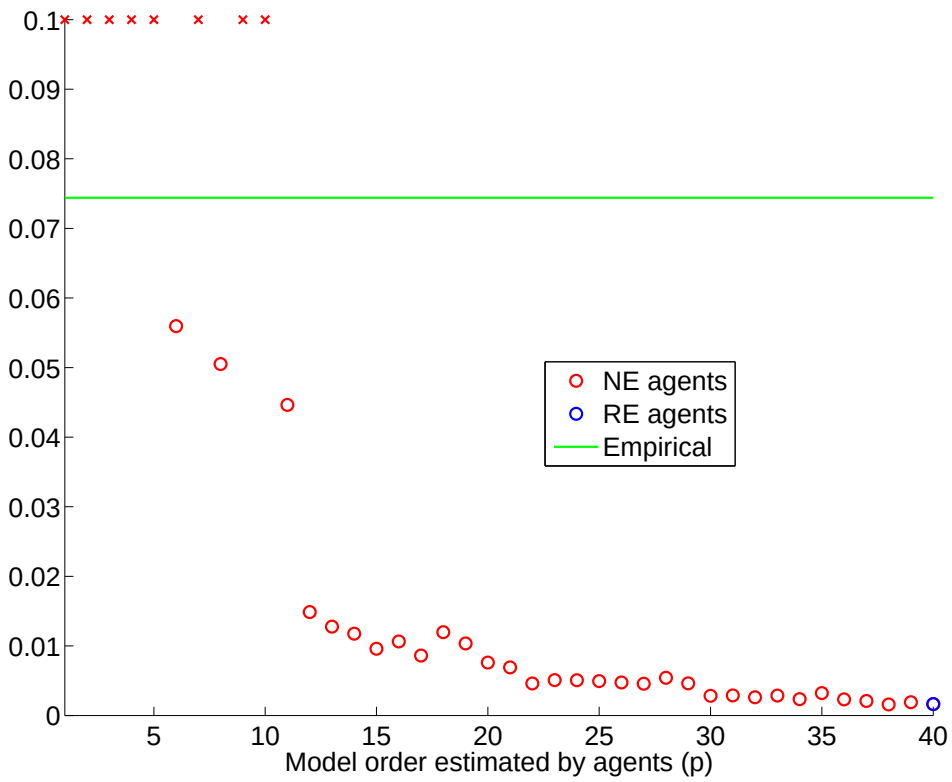
Note: Circles depict mean correlations over 200 simulations of length equal to the empirical sample, while bars depict the 5th and 95th percentile of simulated values. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 9: Correlation of Consumption Growth in Year τ with Consumption Growth from Year $\tau + 2$ to $\tau + 6$, for Different AR(p) Models of Earnings Growth



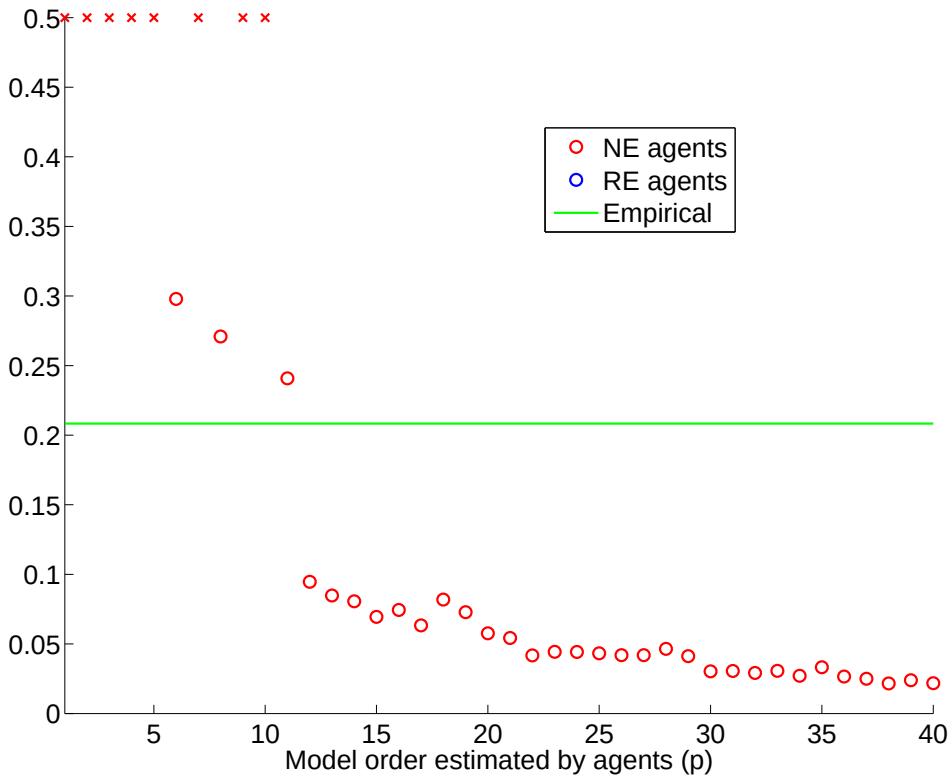
Note: Circles depict mean correlations over 200 simulations of length equal to the empirical sample, while bars depict the 5th and 95th percentile of simulated values. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 10: Equity Premium (= Mean Excess Return of Equity over the Riskfree Rate) for Different AR(p) Models of Earnings Growth



Note: "x" denote equity premia above 10%. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 11: Standard Deviation of Excess Returns for Different AR(p) Models of Earnings Growth



Note: "x" denote standard deviations above 50%. "NE agents" use models with too few lags (AR(1) to AR(39)) while "RE agents" use the correct (AR(40)) model for earnings growth.

Figure 12: Standard Deviation of Consumption Growth for Different AR(p) Models of Earnings Growth

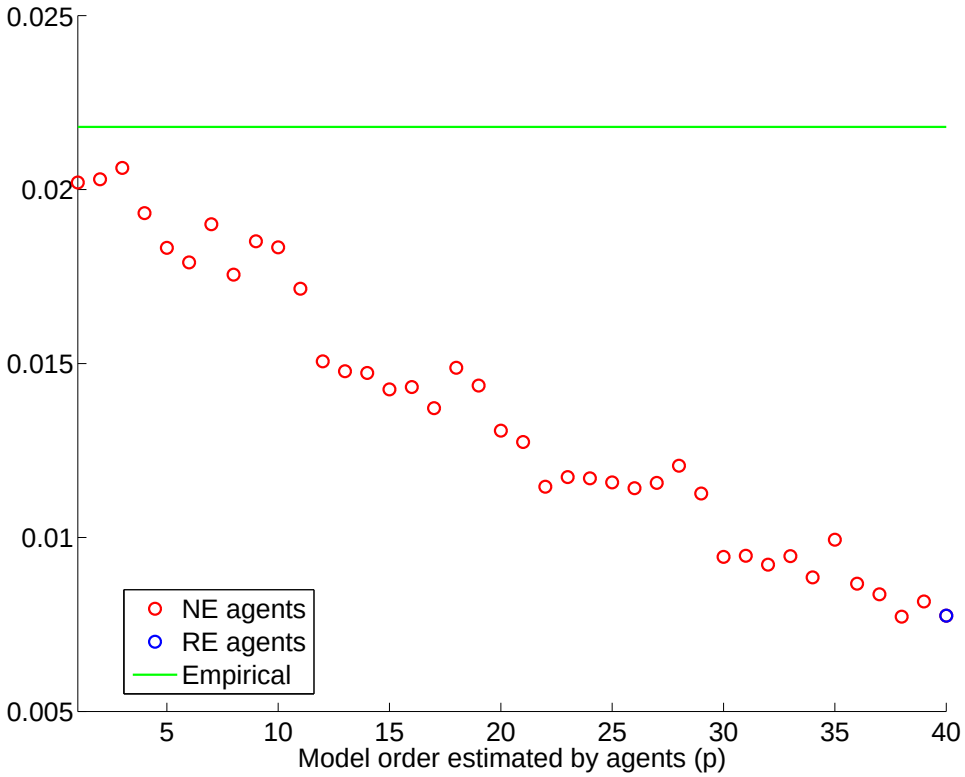


Figure 13: Equity Allocation of Rational Agents for Different AR(p) Models of Earnings Growth held by NE Agents: Average (Top Panel) and Lower Bound (Bottom Panel)

