



Heterogeneous Agents and Long Horizon Features of Asset Prices

Blake LeBaron, International Business School, Brandeis University

Working Paper Series

Heterogeneous Agents and Long Horizon Features of Asset Prices

Blake LeBaron *

International Business School

Brandeis University

Preliminary: July 1, 2013

Revised: Sept 29, 2013

Abstract

Heterogeneous agent models for financial markets have provided explanations for many empirical regularities of relatively high frequency (hourly/daily) financial time series. They have been much quieter when it comes to longer range features. This paper examines a simplified computational heterogeneous agent model in the context of various longer range time series properties for equity returns. The model is compared to a specially created long range data set, and is found to perform well in terms of replicating features, and even revealing some aspects of the data that have not been well quantified to date. By matching empirical properties at both short and long horizons this sets a higher standard in terms of validation which this model is able to match.

Keywords: Learning, Asset Pricing, Financial Time Series, Evolution, Memory

*415 South Street, Mailstop 32, Waltham, MA 02453 - 2728, blebaron@brandeis.edu, www.brandeis.edu/~blebaron. The author is also a research associate at the National Bureau of Economic Research. This research is partially supported by funding from the Institute for New Economic Thinking, INET.

1 Introduction

Over the past two decades much has been learned about the dynamics of models for financial markets designed around heterogeneous agents. Whether analytic or computational, they have emphasized some basic properties. Among these are the ability to amplify the volatility of relatively calm or constant fundamentals, and a demonstration that convergence to a common rational expectations equilibrium may be a more difficult task than is often assumed. Beyond these basic theoretical insights, agent-based markets have demonstrated an ability to replicate certain stylized facts. These include fat tailed return distributions, persistent volatility, and large, persistent deviations from fundamental values.¹ Some of these common features are not even considered by DSGE asset pricing models, yet they appear almost generic in multi-agent models.

Agent-based approaches have been quieter about empirical features at longer horizons. There may be several reasons for this. One is that for univariate time series, longer range features are relatively weak, and not well agreed on. More importantly, most of facts at longer horizons are related to cross sectional asset pricing models.² It may also be true that agent-based models have more difficulty fitting longer horizon features.³ Finally, almost no papers have tried to fit features at both higher and lower frequencies simultaneously. Doing this sets a very high standard for model validation and should help sort through the very crowded space of agent-based approaches to asset pricing.

The goal of this paper is to show that a relatively simple agent-based computational model introduced in LeBaron (2012) is capable of replicating empirical features at multiple horizons. Some of the longer horizon properties that are matched are not well known, and they have required building a special long range data series, splicing various series together and adjusting for share repurchases. The model gives some new insights into the previous empirical work, by suggesting new places to look for important nonlinearities. An example of this is a dramatic asymmetry in return momentum that was discovered first in the model, and then confirmed with financial data.

The model is also simple enough to give some important insights into what is driving the results. It emphasizes several critical components. The most important of these is the fact that agents must learn to

¹ Various surveys of agent-based financial markets discuss many of these papers. These include Hommes (2006), LeBaron (2006), Chiarella, Dieci & He (2009), Hommes & Wagener (2009), Lux (2009), and more recently, Chen (2012). The number of models capable of generating fat tailed distributions alone must be enormous.

² See Cochrane (2011) for some connections between cross sectional and time series relations in asset returns.

³ Recently, more papers have appeared which have turned to longer horizons. This includes Boswijk, Hommes & Manzan (2007), Kouwenberg & Zwinkels (2010), Chiarella, He & Zwinkels (2009), and Chiarella, He, Huang & Zheng (2012). Also, LeBaron (2001a) began trying to take a realistic dividend process and measured price reversals by looking at price/dividend regressions.

forecast returns and risk.⁴ This mechanism allows for very large movements in the price/dividend ratio due to common changes in risk perception, while not requiring large shifts in wealth across different strategy types. A second major feature is the fact that learners are using heterogeneous gain perspectives. They assign differing weights to past information as it comes in. The critical gain level reflects beliefs about both the stationarity of the observed times series, and also the signal to noise ratio of incoming information.⁵ It is conjectured that multiple gain levels can lead to unusual time series behavior such as long memory and volatility time asymmetry.⁶ This paper demonstrates that they are capable of generating a rich set of empirical features.⁷

These features include well known facts such as leptokurtic return distributions, persistent volatility, and the leverage effect. Beyond these relatively short horizon features, the model also displays characteristics which are well known from longer horizon returns including persistent swings around fundamentals and return momentum and reversals. Finally, several features are replicated which have not been well documented. These include nonlinearities in the risk/return relationship, connections between price/dividend ratios and volatility, and asymmetries between short and long horizon volatility. This long list sets a high threshold in terms of model validation, and presents a strong case for many of the internal features of this model, especially the aggregation across different learning gain levels.

Section 2 will briefly describe the model. Section 3 describes the empirical results. There is a description of the data set used, some basic empirical features common to many agent-based models, and then the more unusual empirical results that are specific to this paper. Section 4 concludes.

2 Model description

This section will briefly describe the structure of the model. It combines components from many parts of the learning literature. The goal is to build a heterogeneous agent asset market which is as parsimonious as possible, but can still do a good job of replicating financial market features. Also, its inner workings should be simple enough for detailed analysis, meet a general plausibility test, and yet be rich enough to

⁴See LeBaron, Arthur & Palmer (1999) for an early example of this, and Branch & Evans (2011) for a simple model of how this mechanism could operate.

⁵ Other papers have considered heterogeneous gain learning. These include Anufriev & Dindo (2010), Honkapohja & Mitra (2006), Levy, Levy & Solomon (1994), Mitra (2005), Thurner, Dockner & Gaunersdorfer (2002), and Zschischang & Lux (2001). LeBaron (2002) considers the ecological battle between learners with differing perspectives, and how short term learning perspectives can survive by creating an environment in which they thrive.

⁶ See Corsi (2009), Dacorogna, Gencay, Muller, Olsen & Pictet (2001), Ding, Granger & Engle (1993) and LeBaron (2001b) for empirical examples. All are probably connected in some way to Granger's aggregation theories in Granger (1980).

⁷LeBaron (2013) deals directly with the necessity of multiple gain levels, and shows that without them simulated markets generate relatively unrealistic pricing time series.

understand several aspects of how wealth moves around in a learning investment environment.

2.1 Assets

The market consists of only two assets. First, there is a risky asset paying a stochastic dividend,

$$d_{t+1} = d_g + d_t + \epsilon_t, \quad (1)$$

where d_t is the log of the dividend paid at time t . Time will be incremented in units of weeks. Lower case variables will represent logs of the corresponding variables, so the actual dividend is given by,

$$D_t = e^{d_t}. \quad (2)$$

The shocks to dividends are given by ϵ_t which is independent over time, and follows a Gaussian distribution with zero mean, and variance, σ_d^2 , that will be calibrated to actual long run dividends from the U.S. The dividend growth rate would then be given by $e^{g+(1/2)\sigma_d^2}$ which is approximately $D_g = d_g + (1/2)\sigma_d^2$.

The return on the stock with dividend at date t is given by

$$R_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}}, \quad (3)$$

where P_t is the price of the stock at time t . Timing in the market is critical. Dividends are paid at the beginning of time period t . Both P_t and D_t are part of the information set used in forecasting future returns, R_{t+1} . There are I individual agents in the model indexed by i . The total supply of shares is fixed, and set to unity,

$$\sum_{i=1}^I S_{t,i} = 1. \quad (4)$$

There is also a risk free asset that is available in infinite supply, with agent i holding $B_{t,i}$ units at time t . The risk free asset pays a rate of R_f which will be assumed to be zero. This is done for two important reasons. It limits the injection of outside resources to the dividend process only. Also, it allows for an interpretation of this as a model with a perfectly storable consumption good along with the risky asset. The standard intertemporal budget constraint holds for each agent i ,

$$W_{t,i} = P_t S_{t,i} + B_{t,i} + C_{t,i} = (P_t + D_t) S_{t-1,i} + (1 + R_f) B_{t-1,i}, \quad (5)$$

where $W_{t,i}$ represents the wealth at time t for agent i .

2.2 Preferences

Portfolio choices in the model are determined by a simple myopic power utility function in future wealth. The agent's portfolio problem corresponds to,

$$\max_{\alpha_{t,i}} \frac{E_t^i W_{t+1,i}^{1-\gamma}}{1-\gamma}, \quad (6)$$

$$st. \quad W_{t+1,i} = (1 + R_{t+1,i}^p)(W_{t,i} - C_{t,i}), \quad (7)$$

$$R_{t+1,i}^p = \alpha_{t,i} R_{t+1} + (1 - \alpha_{t,i}) R_f. \quad (8)$$

$\alpha_{t,i}$ represents agent i 's fraction of savings ($W - C$) in the risky asset. It is well known that the solution to this problem yields an optimal portfolio weight given by,

$$\alpha_{t,i} = \frac{E_t^i(r_{t+1}) - r_f + \frac{1}{2}\sigma_{t,i}^2}{\gamma\sigma_{t,i}^2} + \epsilon_{t,i}, \quad (9)$$

with $r_t = \log(1 + R_t)$, $r_f = \log(1 + R_f)$, $\sigma_{t,i}^2$ is agent i 's estimate of the conditional variance at time t , and $\epsilon_{t,i}$ is an individual shock designed to make sure that there is some small amount of heterogeneity to keep trade operating.⁸ It is distributed normally with variance, σ_ϵ^2 .

In the current version of the model neither leverage nor short sales are allowed. The fractional demand is restricted to $\alpha_{t,i}$ to $\alpha_L \leq \alpha_{t,i} \leq \alpha_H$. The addition of both these features is important, but adds significant model complexity. One key problem is that with either one of these, one must address problems of agent bankruptcy, and borrowing constraints. Both of these are not trivial, and involve many possible implementation details. One further modification is that portfolio choices are tapered in their extremes.

$$\tilde{\alpha}_{t,i} = \begin{cases} \alpha_{t,i} & 0.1 \leq \alpha_{t,i} \leq 0.9 \\ 0.9 + 0.02(\alpha_{t,i} - 0.9) & \alpha_{t,i} > 0.9 \\ 0.1 + 0.02(\alpha_{t,i} - 0.1) & \alpha_{t,i} < 0.1 \end{cases} \quad (10)$$

This shrinks the extreme levels of α to the central range. This can be viewed as a slight behavioral mod-

⁸The derivation of this follows Campbell & Viceira (2002). It involves taking a Taylor series approximation for the log portfolio return.

ification that agents require extra empirical evidence to push their portfolios weights to the most extreme positions. This has little impact on most of the simulation runs, and it could be easily removed without affecting most runs. However, it is important in near rational expectations benchmarks which cause the agents to hold almost all equity since volatility is driven to very low levels.⁹ The above restriction keeps the agents in the interior of their portfolio decision and helps the market to converge to a well defined equilibrium in this case.¹⁰

Consumption will be assumed to be a constant fraction of wealth, λ . This is identical over agents, and constant over time. The intertemporal budget constraint is therefore given by

$$W_{t+1,i} = (1 + R_{t+1}^p)(1 - \lambda)W_{t,i}. \quad (11)$$

This also gives the current period budget constraint,

$$P_t S_{t,i} + B_{t,i} = (1 - \lambda)((P_t + D_t)S_{t-1,i} + (1 + R_f)B_{t-1,i}). \quad (12)$$

This simplified portfolio strategy will be used throughout the chapter. It is important to note that the fixed consumption/wealth, myopic strategy approach given here would be optimal in a standard intertemporal model for consumption portfolio choice subject to two key assumptions. First, the intertemporal elasticity of substitution would have to be unity to fix the consumption wealth ratio, and second, the correlation between unexpected returns and certain state variables would have to be zero to eliminate the demand for intertemporal hedging.¹¹

2.3 Expected Return Forecasts

The basic problem faced by agents is to forecast both expected returns and the conditional variance one period into the future. This section will describe the forecasting tools used for expected returns. A forecast strategy, indexed by j , is a method for generating an expected return forecast $E^j(r_{t+1})$. Agents, indexed by i , will choose forecast rules, indexed by j , according to expected utility objectives.

All the forecasts will use long range forecasts of expected values using constant gain learning algorithms

⁹This is an appearance of the equity premium puzzle when this model is near the rational expectations equilibrium.

¹⁰It also corresponds to the neural network based strategies used in LeBaron (2001a), but in a crude piecewise linear fashion.

¹¹See Campbell & Viceira (1999) for the basic framework. Also, see Giovannini & Weil (1989) for early work on determining conditions for myopic portfolio decisions. Hedging demands would only impose a constant shift on the optimal portfolio, so it is an interesting question how much of an impact this might have on the results.

equipped with the minimum gain level, denoted g_L .

$$\bar{r}_t = (1 - g_L)\bar{r}_{t-1} + g_L r_t \quad (13)$$

$$\overline{pd}_t = (1 - g_L)\overline{pd}_{t-1} + g_L pd_{t-1} \quad (14)$$

$$\bar{\sigma}_t^2 = (1 - g_L)\bar{\sigma}_{t-1}^2 + g_L (r_t - \bar{r}_t)^2 \quad (15)$$

$$\bar{\sigma}_{pd,t}^2 = (1 - g_L)\bar{\sigma}_{pd,t-1}^2 + g_L (pd_t - \overline{pd}_t)^2 \quad (16)$$

$$\overline{\Delta p}_t = (1 - g_L)\overline{\Delta p}_{t-1} + g_L \Delta p_t \quad (17)$$

$$\Delta p_t = p_t - p_{t-1} \quad (18)$$

The log price/dividend ratio is given by $pd_t = \log(P_t/D_t)$. The forecasts used will combine four linear forecasts drawn from well known forecast families. The first of these is a form of adaptive expectations which corresponds to,

$$f_t^j = f_{t-1}^j + g_j(r_t - f_{t-1}^j). \quad (19)$$

Forecasts of expected returns are dynamically adjusted based on the latest forecast error. This forecast format is simple and generic. It has roots connected to adaptive expectations, trend following technical trading, and also Kalman filtering.¹² The critical parameter is the gain level represented by g_j . This determines the weight that agents put on recent returns and how this impacts their expectations of the future. Forecasts with a large range of gain parameters will compete against each other in the market. Finally, this forecast will be trimmed in that it is restricted to stay between the values of $[-h_j, h_j]$. These will be set to relatively large values, and are randomly distributed across rules.

¹²See Cagan (1956), Friedman (1956), Muth (1960), and Phelps (1967) for original applications. A nice summary of the connections between Kalman filtering, adaptive expectations, and recursive least squares is given in Sargent (1999). A recent paper demonstrating a more general connection to state space models and expected returns is Pastor & Stambaugh (2009). Empirically, Frankel & Froot (1987) and Taylor & Allen (1992) provide evidence that at least some forecasters do use these rules. Finally, Hommes (2011) surveys some of the laboratory evidence that experimental subjects also use extrapolative methods for forecasting time series.

The second forecasting rule is based on a classic fundamental strategy. It is based on the log linear approximation for returns,¹³

$$r_{t+1} \approx k + \rho p_{t+1} + (1 - \rho)d_{t+1} - p_t, \quad (20)$$

$$\rho = \frac{1}{1 + e^{-(\bar{p}-d)}} \quad k = -\log(\rho) - (1 - \rho)\log\left(\frac{1}{\rho} - 1\right). \quad (21)$$

This approximation allows for a separation of forecasts into a dividend and price component. Agents are assumed to have observed the random walk dividend process, and d_{t+1} is easy to forecast. The mapping of dividends into price movements is much more complicated, and here they resort to a linear regression of $\Delta p_{t+1} = p_{t+1} - p_t$ on current price dividend ratios, pd_t .

$$f_t^j = E_{t,j}(\Delta p_{t+1}) = \overline{\Delta p_t} + \beta_t^j(pd_t - \overline{pd_t}). \quad (22)$$

where pd_t is $\log(P_t/D_t)$, and

$$E_{t,j}(p_{t+1}) = f_t^j + p_t \quad (23)$$

$$E_{t,j}(d_{t+1}) = d_t \quad (24)$$

$$E_{t,j}r_{t+1} \approx k + \rho E_{t,j}p_{t+1} + (1 - \rho)E_{t,j}d_{t+1} - p_t. \quad (25)$$

The third forecast rule will be based on simple linear regressions. It is a predictor of returns at time t given by

$$f_t^j = \bar{r}_t + \sum_{i=0}^{M_{AR}-1} \beta_{t,i}^j(r_{t-i} - \bar{r}_t) \quad (26)$$

This strategy works to eliminate short range autocorrelations in returns series through its behavior, and $M_{AR} = 3$ for all runs in this chapter. It will be referred to as the Short AR forecast.¹⁴

The previous two rules will be estimated each period using recursive least squares. There are many examples of this for financial market learning.¹⁵ The key difference is that this model will stress hetero-

¹³ See Campbell, Lo & MacKinlay (1996) for a textbook exposition.

¹⁴ These simple forecasting agents who use only recent returns in their models, fighting against return correlations, share some features with the momentum traders of Hong & Stein (1999).

¹⁵ See Evans & Honkapohja (2001) for many examples, and also very extensive descriptions of recursive least squares learning

geneity in the learning algorithms with wealth shifting across many different rules, each using a different gain parameter in its online updating.¹⁶

The final rule is a buy and hold strategy using the long run mean, \bar{r}_t , for the expected return, and the long run variance, $\bar{\sigma}_t^2$, as the variance estimate. This portfolio fraction is then determined by the demand equation used by the other forecasting rules. This gives a useful passive benchmark strategy which can be monitored for relative wealth accumulation in comparison with the other active strategies.

2.4 Regression Updates

Forecasting rules are continually updated. The adaptive forecast only involves fixed forecast parameters, so its updates are trivial, requiring only the recent return. The two regression forecasts are updated each period using recursive least squares.

All the rules assume a constant gain parameter, but each rule in the family corresponds to a different gain level. This again corresponds to varying weights for the forecasts looking at past data.

The fundamental regression is updated according to,

$$\beta_{t+1}^j = \beta_t^j + \frac{g_j}{\bar{\sigma}_{pd,t}^2} (pd_{t-1} - \bar{p}d_{t-1}) u_{t,j} \quad (27)$$

$$u_{t,j} = ((p_t - p_{t-1}) - f_{j,t-1})$$

Also, β_t^j is restricted to be between 0 and -0.05 . The zero upper bound on β makes sure this strategy is mean reverting, with an overall stabilizing impact on the market.

For the lagged return regression the update follows,

$$\beta_{t+1,i}^j = \beta_{t,i}^j + \frac{g_j}{\bar{\sigma}_{r,t}^2} (r_{t-i} - \bar{r}_{t-i}) u_{t,j}, \quad (28)$$

$$u_{t,j} = (r_t - f_t^j)$$

where g_j is again the critical gain parameter, and it varies across forecast rules.¹⁷

methods.

¹⁶Another recent model stressing heterogeneity in an OLS learning environment is Georges (2008) in which OLS learning rules are updated asynchronously.

¹⁷This format for multivariate updating is only an approximation to the true recursive estimation procedure. It is assuming that the variance/covariance matrix of returns is diagonal. Generated returns in the model are close to uncorrelated, so this approximation is reasonable. This is done to avoid performing many costly matrix inversions. Also, the standard recursive least squares is simplified by using the long run estimates for the mean in both regressions. Only the linear coefficient is estimated with a heterogeneous learning model. This is done to simplify the learning model, and concentrate heterogeneity on the linear parameters, β_t^j .

2.5 Variance Forecasts

The optimal portfolio choice demands a forecast of the conditional variance as well as the conditional mean.¹⁸ The variance forecasts will be generated from adaptive expectations as in,

$$\hat{\sigma}_{t+1,j}^2 = \hat{\sigma}_{t,j}^2 + g_{j,\sigma}(e_{t,j}^2 - \hat{\sigma}_{t,j}^2) \quad (29)$$

$$e_{t,j}^2 = (r_t - f_{t-1}^j)^2, \quad (30)$$

where $e_{t,j}^2$ is the squared forecast error at time t , for rule j . The above conditional variance estimate is used for all the rules. There is no attempt to develop a wide range of variance forecasting rules, reflecting the fact that while there may be many ways to estimate a conditional variance, they often produce similar results.¹⁹ This forecast method has many useful characteristics as a benchmark forecast. First, it is essentially an adaptive expectations forecast on second moments, and therefore shares a functional form similar to that for the adaptive expectations family of return forecasts. Second, it is closely related to other familiar conditional variance estimates.²⁰ Finally, the gain level for the variance in a forecast rule, $g_{j,\sigma}$, is allowed to be different from that used in the mean expectations, g_j . This allows for rules to have a different time series perspective on returns and volatility.

Finally, agents do not update their estimates of the variance each period. They do this stochastically by updating their variance estimate each period with probability 0.25. This is done for several reasons. First it introduces more heterogeneity into the variance estimation part of the model since its construction yields a lot of similarity in variance forecasts. Also, if variance updating occurred simultaneous to return forecasts, the market would be unstable. Spirals of ever falling prices, and increasing variance estimates would be impossible to avoid in this case.

¹⁸ Several other papers have explored the dynamics of risk and return forecasting. This includes Branch & Evans (2011) and Gaunersdorfer (2000). In LeBaron (2001a) risk is implicitly considered through the utility function and portfolio returns. Obviously, methods that parameterize risk in the variance may miss other components of the return distribution that agents care about, but the gain in tractability is important.

¹⁹ See Nelson (1992) for early work on this topic.

²⁰ See Bollerslev, Engle & Nelson (1995) or Andersen, Bollerslev, Christoffersen & Diebold (2006) for surveys of the large literature on volatility modeling.

2.6 Market Clearing

The market is cleared by setting the individual share demands equal to the aggregate share supply of unity,

$$1 = \sum_{i=1}^I Z_{t,i}(P_t). \quad (31)$$

Writing the demand for shares as its fraction of current wealth, remembering that $\alpha_{t,i}$ is a function of the current price gives

$$P_t Z_{t,i} = (1 - \lambda) \alpha_{t,i}(P_t) W_{t,i}, \quad (32)$$

$$Z_{t,i}(P_t) = (1 - \lambda) \alpha_{t,i}(P_t) \frac{(P_t + D_t) S_{t-1,i} + B_{t-1,i}}{P_t}. \quad (33)$$

This market is cleared for the current price level P_t . This needs to be done numerically given the complexities of the various demand functions and forecasts, and also the boundary conditions on $\alpha_{t,i}$.²¹ It is important to note again, that forecasts are conditional on the price at time t , so the market clearing involves finding a price which clears the market for all agent demands, allowing these demands to be conditioned on their forecasts of R_{t+1} given the current price and dividend.²²

2.7 Gain Levels

An important design question for the simulation is how to set the range of gain levels for the various forecast rules. These will determine the dynamics of forecasts. Given that this is an entire distribution of values it will be impossible to accomplish much in terms of sensitivity analysis on this. Therefore, a reasonable mechanism will be used to generate these, and this will be used in all the simulations.

Gain levels will be thought of using their half-life equivalents, since the gain numbers themselves do not offer much in the way of economic or forecasting intuition. For this think of the simple exponential forecast mechanism with

$$f_{t+1}^j = (1 - g_j) f_t^j + g_j e_{t+1}. \quad (34)$$

²¹ A binary search is used to find the market clearing price using starting information from P_{t-1} .

²² The current price determines R_t which is an input into both the adaptive, and short AR forecasts. Also, the price level P_t enters into the P_t/D_t ratio which is required for the fundamental forecasts. All forecasts are updated with this time t information in the market clearing process.

This easily maps to the simple exponential forecast rule,

$$f_t = \sum_{k=1}^{\infty} (1 - g_j)^k e_{t-k}. \quad (35)$$

The half-life of this forecast corresponds to the number of periods, m_h , which drops the weight to $1/2$,

$$\frac{1}{2} = (1 - g_j)^{m_h}, \quad (36)$$

or

$$g_j = 1 - 2^{-1/m_h}. \quad (37)$$

The distribution of m_h then is the key object of choice here. It is chosen so that $\log_2(m_h)$ is distributed uniformly between a given minimum and maximum value. The gain levels are further simplified to use only 5 discrete values. These are given in table 2. These distributions are used for all forecasting rules. All forecast rules need a gain both for the expected return forecast, and the variance forecast. These will be chosen independently from each other. This allows for agents to have differing perspectives on the importance of past data for the expected return and variance processes.

2.8 Adaptive rule selection

The design of the models used here allows for both passive and active learning. Passive learning corresponds to the long term evolution of wealth across strategies. Beyond passive learning, the model allows for active learning, or adaptive rule selection. This mechanism addresses the fact that agents will seek out strategies which best optimize their estimated objective functions. In this sense it is a form of adaptive utility maximization.

Implementing such a learning process opens a large number of design questions. This chapter will stay with a relatively simple implementation. The first question is how to deal with estimating expected utility. Expected utility will be estimated using an exponentially weighted average over the recent past,

$$\hat{U}_{t,j} = \hat{U}_{t-1,j} + g_u^i (U_{t,j} - \hat{U}_{t-1,j}), \quad (38)$$

where $U_{t,j}$ is the realized utility for rule j received at time t . This corresponds to,

$$U_{t,j} = \frac{1}{1-\gamma} (1 + R_{t,j}^p)^{(1-\gamma)} \quad (39)$$

with $R_{t,j}^p$ the portfolio holdings of rule j at time t . Each rule reports this value for the 5 discrete agent gain parameters, g_u^i . Agents choose rules optimally using the objective that corresponds to their specific perspective on the past, g_u^i , which is a fixed characteristic. The gain parameter g_u^i follows the same discrete distribution as that for the expected return and variance forecasts.

The final component to the learning dynamic is how the agents make the decision to change rules. The mechanism is simple, but designed to capture a kind of heterogeneous updating that seems plausible. Each period a certain fraction, L , of agents is chosen at random. Each one randomly chooses a new rule out of the set of all rules. If this rule exceeds the current one in terms of estimated expected utility, then the agent switches forecasting rules.

3 Empirical results

3.1 Data description

This paper uses a data set for U.S. stock returns and dividends as a comparison. It is designed both to cover as long a horizon as is reasonably possible for dividends, prices, and return volatility. It merges three different data sets. For the period from 1926 through 2012, the benchmark CRSP data is used with the value weighted index as the proxy for market returns, and interest rates given by the 90 day U.S. treasury bill. The difference between the returns on the index with and without dividends are used to construct price dividend ratios as is standard in the finance literature. Also, the daily returns data are used to build a monthly volatility series with a simple monthly realized volatility.²³ For the period from 1871 to 1925 the data set constructed by Shiller is used, which available at his website.²⁴ This includes stock prices, dividends, interest rates, and inflation. The Shiller data is only available monthly frequency, so it is also augmented with an early daily series from Schwert which is used to generate early monthly realized volatilities back to 1886.²⁵ For most results the range of prices, returns and volatility will cover 1886-2012

²³This method is very commonly used with intraday data, but was originally used in French, Schwert & Stambaugh (1987) at the daily frequency.

²⁴ This is the main input for Shiller (2000), and many papers.

²⁵Schwert (1990) gives extensive details on the construction of this series which is available at his website.

with a splice into the standard CRSP series in 1926.

The real dividend series in the Shiller data set is used for calibration. Figure 1 plots the price dividend ratio for the entire range of the data set. The solid line gives ratios which are adjusted for share repurchase payouts, the dashed line is the raw dividend series. For most of the early data there is no difference, but in later years it is clear that ignoring share repurchase payouts is a problem. The payout adjusted series could be argued as reasonably stationary, while the unadjusted series moves into a very different territory in the 1990's and beyond.²⁶

When thinking about a single candidate for the aggregate fundamental series for aggregate stock returns there are many candidates. Table 1 compares some of the available options. This most obvious would be the aggregate dividend series which is taken from the Shiller data set. It is presented in both regular form and adjusted for payouts. The latter case shifts its growth rate close to a real two percent growth rate, putting it more in line with the standard macro measures of the post-war U.S. economy. It should be noted that volatility of aggregate dividends falls well above that of other measures of aggregate income flows. In particular, consumption is roughly one fifth the annual standard deviation of dividends. This leaves one with a bit of a dilemma in terms of calibration, since a key feature should be the model's ability to magnify the variability of these rather quiet fundamentals. In this paper, the philosophy will be to take the adjusted dividend as the best measure of the of the aggregate cash flow. However, aggregate information suggests that lowering its volatility might be a sensible experiment.

3.2 Basic features

The basic parameters for the simulation are given in table 2. This table gives the number of agents and rules, and the basic simulation parameters. Results will be presented for two different simulation specifications. First, the multiple gain runs correspond to a full run using all the different gain levels across forecast rules and agents. These runs serve as a benchmark for replicating actual returns in the U.S. data set. The simulation is allowed to run for 200,000 weeks, and the last 100,000 weeks are used in most of the time series analysis, giving approximately 1900 years of data.

²⁶The adjustment factor used is from Boudoukh, Michaely, Richardson & Roberts (2007). The adjustment used here augments dividends with total share repurchases. There are several other measures one might use, none of which perfectly accounts for the attempts of firms to get cash into shareholders hands. Boudoukh et al. (2007) try several, and give a lot of discussion on the debate about which measure one should use. Their data ends in 2003, and it is augmented with data from S&P which can be found on the web at, <http://finance.yahoo.com/news/p-500-stock-buybacks-increase-133000759.html> and <http://www.indexuniverse.com/publications/journalofindexes/joi-articles/9431-the-importance-of-dividends-and-buybacks-ratios-for-gauging-equity-values.html>. The annual data is interpolated to a monthly frequency.

A figure showing the final 100 years of a run is presented in figure 2. The upper panel shows the price dividend ratio which is moving through time, displaying large and persistent swings. The pattern is somewhat irregular with no clear periodicity or maximum or minimum level which would allow investors to call the top or the bottom. The middle panel presents the weekly returns corresponding to the price/dividend ratios. These give temporal pockets of high and low volatility common in most financial time series. The lower panel displays trading volume which also moves counter cyclically to the P/D ratio, and is strongly correlated with the level of volatility in the return series.

The overall structure of the agent strategies is presented in figure 3. This graph presents the wealth fractions for the various forecast families over the entire 200,000 week run. In terms of long range patterns some key features are very obvious. First, the short range arbitrage strategy is nearly driven out of the market. It essentially does this to itself by eliminating all short run predictability. Second, the fraction of adaptive or trend following strategies is significantly larger than the fundamental strategies. They control close to 40 percent of the wealth, and the fundamental types only hold about 10 percent. This relative wealth comparison for destabilizing and stabilizing strategies is important for the market dynamics. It is also interesting to note that the passive buy and hold strategy controls almost 50 percent of the wealth. This is consistent with some results from the model which suggest that predictability in the series is not too strong. Passive strategies therefore hold their own against the more active ones.²⁷

The second set of runs is a counterfactual run with the gain levels set to very long half lives of about 100 years. This market converges to something very different from the multiple gain runs. It is again run for 200,000 weeks, and the final 100 years are summarized in figure 4. The three panels display P/D ratios, returns, and trading volume. It is obvious this run is very different from the previous one. The price dividend ratio stays relatively constant around a level of 35. Returns are uniform with no dramatic changes in volatility. Trading volume is also flat, and more than an order of magnitude lower than that in the multiple gain benchmark. This casual picture appears to show a market which has converged to something close to a simple rational expectations equilibrium.

Table 3 presents basic summary statistics for the data and the two different simulation runs. The multiple gain simulations give reasonable representations for the long range properties of the mean and standard deviation of excess returns. They are slightly high in both instances.²⁸ The small gain comparison run gen-

²⁷Detailed analysis of the underlying market fractions and wealth distributions are presented in LeBaron (2012).

²⁸ The volatility in many simulation runs is often too high relative to aggregate data. One key reason for this is that the simulated market is driven by a single stock which is traded by all the agents. There is no opportunity to smooth out any of the trading variability across stocks as there is in real markets where the benchmark is an average over many stocks.

erates returns which are too low, and a standard deviation which is less than half that of the data. In this paper the dynamics of the fundamental series are crucial to the market. The lower part of the table presents the mean and standard deviation for the price/dividend ratio. The multiple gain simulation is closely aligned with the payout adjusted levels from the long range data set. The small gain run confirms the figure by showing a variability in the price/dividend ratio which is about 1/3 the actual data. This lower risk in the aggregate series gets reflected in a higher price for the risky asset and a much lower price/dividend ratio of 37 which compares to 23 in the data.

The next figures and table briefly examine some of the higher frequency (weekly) properties of the time series. Table 4 shows that the data and multiple gain simulation generate some negative skew, and large kurtosis. For the small gain runs there is very little evidence for any deviation from Gaussian returns. The skewness is zero, and the kurtosis is only 3.4. Figures 5 and 6 display histograms for weekly returns for the simulations and the data, respectively. Both show features consistent with the table of fat tails, and some negative skewness. Results for the small gain runs are consistent with the table in that they are independent and close to Gaussian. For brevity they will no longer be included in the results.

Figures 7 and 8 continue with two other well known stylized facts of financial time series and agent-based markets. In both cases monthly data is used to stay consistent with the long range theme of the paper, and for comparisons with results presented in later sections. Monthly return autocorrelations in figure 7 show very little evidence for any linear structure in either the data or the simulations as should be expected.²⁹ The volatility correlations in figure 8 are a little different from what is often used. The monthly volatility series are built as realized volatility values using the underlying higher frequency series. In the case of the data, these are daily returns, and for the simulations weekly returns. These measures are generally more accurate than simple squared or absolute returns. In both cases the figure displays the large and highly persistent nature of volatility. Positive autocorrelations go out for nearly two years, and show a slow drop off which is characteristic of a long memory process.

The features presented so far are not new. They simply show that like many other agent-based models, this one is able to generate returns which follow the basic stylized facts in distributions and persistence. Most of these are relatively short horizon, but it should be noted that the long memory aspect of volatility demonstrates that it spans many time scales from short to long. The next section moves on to more details on longer horizon aspects of the data. Many of these have not previously been reported or tested in agent-

²⁹For the simulations much of the short range behavior is driven by the short range arbitrage types who work hard to eliminate correlated returns.

based models.

3.3 Nonlinear features

This section turns to more long range features, and compares simulation results with those from the long range merged data set. All of the simulations use the last 100,000 weeks (approximately 1900 years) from a 200,000 week simulation, and the data go from 1886 through 2012. In most cases a nonparametric regression will be estimated to see if there are any interesting nonlinear features in the model, and if there is any empirical support for these.

The first set of results are not particularly new. Once it is clear that prices are deviating from some fundamental for significant periods of time, then predictive regressions based on these deviations have to have some value.³⁰ Table 5 presents simple linear regressions at two different forecasting horizons, one quarter, and one year. As is common in this literature the future excess returns are regressed on the current log price/dividend ratio.³¹ For the data, the results are consistent with values from the literature. The coefficients are negative and significant indicating the presence of mean reversion. Also, for the data the estimated R-squared is only modest with a value of 0.01 for quarterly forecasts, and rising to 0.07 for annual excess return forecasts. The simulations provide qualitative alignment with these values, but show larger amounts of predictability. The coefficients are highly significant which is probably due both to a cleaner mean reversion, and also a much longer sample. The quarterly and annual R-squared's are now 0.05, and 0.20. Even though these values are larger than the R-squared's estimated on the data, it is interesting that with a simulated large sample, there is still a large amount of imprecision in long range forecasts.

The next set of results uses a nonparametric kernel regression to see if there are any nonlinear connections in either model or the data. The method used is relatively standard and straightforward. A Gaussian kernel will be applied in a simple bivariate regression of y on x . The key parameter in this setup is the bandwidth. The bandwidth will be chosen using randomized cross validation. The sample is split randomly into estimation and testing samples using an (80,20) split. The model fit is estimated with mean absolute error in the test partition with fitted values coming from the estimation subsample. The new feature used here will be to estimate the optimal bandwidth using the lengthy simulation series, and apply

³⁰ See Campbell & Shiller (1988) for the early work. Cochrane (2008) is a good demonstration of the predictability of returns and cash flows and how they are connected. However, the out of sample use of these predictors has often been questioned, Goyal & Welch (2003) and Fisher & Statman (2006). Also, see Pettenuzzo, Timmerman & Valkanov (2013) for some positive evidence predictability using a Bayesian approach. Lettau & Ludvigson (2010) is a very complete survey on this subject.

³¹ Estimation is done using overlapping data in both cases. For the simulations, given the large amount of data, one could shift to nonoverlapping samples without changing the results.

this same bandwidth to the much shorter U.S. data series. This is essentially using the simulation as a kind of model specification detection tool, but then reestimating parameters for the data.

Figure 9 shows the results of the nonparametric regression for the model and data for the one year excess return forecast. The conditional returns are surprisingly similar in both cases, and also very near linear. It is probably unlikely that much improvement would come from moving to a nonlinear forecasting model in either case. It is also comforting that even though the model generates slightly different regression coefficients it is very close on this more detailed specification test.

Regressing long run returns is very standard in the finance literature. Figure 2 suggests another strong relationship in the model. Return volatility appears to move opposite to the price/dividend ratio, with volatility low at high prices, and high when the market falls to new low levels. This is consistent with a general notion of overconfidence, or under assessment of risk, at the top and over assessment of risk at a market bottom. To test this more carefully the within month volatility is compared to the end of month price/dividend ratio. This is essentially looking at the price/dividend relative to its recently completed volatility level. This is obviously a completely endogenous relationship, and not much should be assumed from this from a predictive or causality perspective.³²

Simple OLS results are presented in table 6. Results are given for both the merged data and the model. Also, it is common to look at logged volatility which has a generally near Gaussian symmetric distribution, so the regression is run on both raw and logged volatility series. In all case there is a negative relationship indicating that volatility is generally low near the top of a P/D run up. For the data, the R-squared's are very small. Only 0.01 for the variance, and increasing to 0.12 for the log volatility series. The simulation again increases these to 0.28 and 0.34 respectively. This result is clearly much stronger in the simulated series than in the actual, but it is moving in the same direction.

The nature and shape of this relationship is again explored with a kernel regression in figure 10. This shows more nonlinear features than the previous kernel regression. For the simulation, the relationship is strong and monotonic with almost a factor of ten change in conditional volatility from the high to the low in the market. There is some weak indication for a slight nonlinear kink in this relationship. Over the low P/D the data generally follow this pattern. However, for extremely high P/D ratios the sign changes and the conditional variance actually rises. This is an interesting deviation from the model and needs to be further explored.³³

³² This is really no reason regression could not be done in the other direction. These should be viewed as a form of internal moment conditions more than a simple or nonparametric regression.

³³ It should be noted that at the extremes of the kernel regression the data becomes relatively sparse and confidence bands would

Another important feature of financial returns series is the asymmetric impact of lagged returns on future volatility. Originally discovered in Black (1976) and Christie (1982), it is often referred to as the “leverage effect” because it can be connected to debt equity ratios and market valuation. The basic feature is that past price decreases lead to a greater impact on future variances than price increases. In other words, volatility tends to rise in falling markets.³⁴ Some form of this can be seen graphically for the simulation in 2, but how well does this line up with the data? Figure 11 shows the kernel regression of future monthly variance on lagged monthly return. Two dramatic features emerge in the figure. First, the simulations and the data are remarkably similar here. Second, they both show the strong asymmetry between returns and future volatility. In both cases a five percent fall in the market is connected with almost twice the level of conditional future variance. This nonparametric curve is referred to as the “news impact curve” in Engle & Ng (1993), where similar patterns are observed for equity returns.

Modern finance has been dominated by the construction of multi factor models. One of these factors, momentum, stands out as being a little different from the others. It is very simple, and not terribly well founded in any economic model of risk. It is based on the medium range (about half a year) persistence of returns.³⁵ Momentum is obviously tied to other trend following strategies such as moving averages, but its systematic use of the cross section is a much better way to implement this.³⁶ Standard momentum tests construct portfolios based on lagged six month returns, then they skip a month, and record results over the next six months. Often this is referred to as a $(6 - 1 - 6)$ strategy. Without a cross section the results are presented purely in a time series context by looking at future returns conditioned on the previous six month returns (subject to skipping a month). Figure 12 compares the simulations to the actual data. Both display a very interesting non monotonic relationship. Over the positive region of lagged returns they both show a consistent positive connection with future returns as the momentum effect predicts. However, moving into negative returns one observes a reversal in both simulation and data. Large falls in the market eventually show up as conditional reversals (increased conditional returns). This again feels consistent with figure 2 where the upward movements in the market are long and drawn out, generating a large amount of possible momentum. Falls in the market are relatively short lived and quickly reverse.

This suggests a pattern where persistence and trends are stronger on the longer upswings in the market

expand, so results at very high P/D ratios should be viewed with some caution.

³⁴Hasanhodzic & Lo (2011) dramatically demonstrate that this empirical feature probably has nothing to do with leverage. This also drives asymmetric GARCH models such as Nelson (1991) and Glosten, Jagannathan & Runkle (1993).

³⁵ See Chan, Jegadeesh & Lakonishok (1996) or Jegadeesh & Titman (1993) for early references. See Carhart (1997) for the evidence that momentum is connected across the cross section in interesting ways, and is occasionally used as a fourth factor in standard asset pricing models.

³⁶ See Brock, Lakonishok & LeBaron (1992) for time series evidence of moving average rules which are related to momentum.

when volatility is lower. It is possible that this feature may show up in other ways. Previously it has been shown that for some financial series the conditional autocorrelations change in relation to past volatility levels.³⁷ When volatility is low conditional return autocorrelations are low, and when volatility is high they fall to near zero, or even negative values for various equity and foreign exchange time series. Figure 13 estimates these connections at much longer horizons than those used previously, and using the kernel regressions used in the previous figures. The return variance estimated over the past three months is used for a volatility measure, and the return autocorrelation is estimated over the following three months. These values are estimated with daily data for the returns, and with weekly data for the simulations.³⁸ The figure shows the clear downward pattern for both series. It is stronger for the U.S. data set which goes from about 0.15 to near zero from the highest levels of volatility. The simulation presents a more modest reduction from 0.12 to 0.08. The feature appears to be present in the simulated model, but at a much less significant level.

There may be many causes for this more muted impact. First, the use of weekly data has to have some impact on these values. Second, and probably most important, the actual data is based on an index which can have many interesting lead lag patterns across different stocks.³⁹ The simulation is based on a single stock, with one key strategy designed to eliminate short range autocorrelations. For the CRSP-VW index there is no such single minded trader present.

Another interesting puzzle in long range returns is the lack of solid time series evidence for a consistent estimate for a positive connection between risk and future excess returns. This issue shows up in models such as the GARCH-M which often show very weak and unstable estimates on parameter related to conditional expected returns. It has recently reappeared in an interesting paper by Rossi & Timmermann (2010) where the authors use some advanced machine learning technology to demonstrate this risk return relationship may be nonlinear. Figure 14 examines this relationship for both simulation and data. Excess returns from the future six months are conditioned on the previous three months of volatility information. The estimation is done using the actual variance rather than the logged variance since the theoretical relationship should be between variance and return for most models, and not logged variance. This does pose a problem since distribution is highly skewed, and for large variances, the samples used are getting small. The graph shows features similar to those reported by Rossi & Timmermann (2010) with a strongly

³⁷See LeBaron (1992).

³⁸The very long range nature of this test is done both to keep with the spirit of this paper, and also because estimating correlations with weekly data requires a reasonable sample size. Three months still only yields 12 weeks to estimate the conditional correlation for the simulations.

³⁹Non synchronous trading is one of the most important of these.

nonmonotonic relation for risk and return. At the lower level of conditional variances risk and return are positively related, as they should be, but this relationship reverses sign for larger volatility levels. The positive connection should be almost obvious for the agent-based model. Here, large volatility over the recent quarter will impact conditional variances estimates, leading to stronger discounts in asset prices and reduced P/D ratios. Naturally, this is connected to higher conditional excess returns going forward. The change in sign in this relationship is still more complex for both data and model to explain. The best possible explanation (at least for the simulations) is that when volatility is at its highest levels, the market is actually falling. The simulated market rarely falls in one period, and takes a few weeks to finally adjust. So the highest volatility level is an indicator that a crash is still in progress, not that it is over, and recovery is imminent. The two figures, while qualitatively similar, are quite different in magnitudes. These extreme volatility levels start at about 0.015 for the data, but not until 0.02 for the model. Left tail behavior may be generally thicker, and more extreme for the simulation relative to the data. Given the sparsity of data in these tails, missing this break point should not be a major concern for a first pass at this test.

Volatility in financial time series displays several interesting features. This paper has presented results on long memory, and also return asymmetries along with connections to correlations and volatility. In their early work on volatility the team from Olsen and Associates documented patterns moving from high to lower frequency volatility measures.⁴⁰

These results are based on different volatility measures. Define a high frequency volatility estimate using within month daily log returns using the following equation,

$$x_t^H = \sum_{j=1}^m |r_{t,j}|.$$

Now define a coarser volatility estimate for a month with the simple absolute return over the monthly summed return,

$$x_t^L = |(\sum_{j=1}^m r_{t,j})|.$$

The volatility asymmetry result looks at cross correlations between these measures. There often appears to be a much stronger impact from the coarser time scale to the finer time scale, than from the fine scale to the coarse scale. One can interpret this as a mechanism where small wiggles at higher frequencies do not impact long range volatility as much as changes in the lower frequency components do. Figure 15 displays

⁴⁰See Dacorogna et al. (2001) and many references. Lynch & Zumbach (2003) is probably the most detailed exposition on this phenomenon with excellent color pictures.

the cross correlations of these high and low frequency volatility measures. It is important to note again, that volatility is measured using weekly data for the simulations, and a finer daily return series for the data. Both display a remarkably similar pattern with coarser volatility measures tending to lead the finer volatility measure. The solid green lines in the figure simply reflect the leading (positive lag) correlations reflected back across zero for comparison. In both cases the leading correlations are larger than the lagging correlations. This feature suggests something interesting about how volatility is processed across multiple time scales both in the model and in the actual data.

4 Conclusions

This paper has explored a relatively simple computational agent-based model in terms of replicating several long range empirical features for aggregate stock returns. This is done while continuing to generate well known aspects of higher frequency price dynamics. The ability of a single model to qualitatively, and in some cases quantitatively, match up with many empirical facts suggests that this model may be a useful benchmark in terms of heterogeneous agent structures.

The replicated features include some well known facts such as price/dividend predictive regressions and the leverage effect. However, in several cases the empirical facts are not as well known. The data and simulations both show that lagged returns have a very different impact on future returns in rising and falling markets. Most of the traditional momentum like features come from periods of price increases, while decreases are more likely to lead to price reversals. Also, volatility moves in an interesting way with the price/dividend ratio. Conditional variances are generally lowest at the market peaks, and highest in the troughs. Finally, risk and return exhibit a much more complicated picture, than that predicted by economic theory in that the relationship appears to be non monotonic with extremely large volatility levels leading to reductions in conditional returns.

The model generates a set of reliable patterns in its time series that can be summarized with a few key features. Prices make large swings from fundamental values, but they are very asymmetric in their dynamics with long, slow, quiet but persistent movements on the upside, and sharp turbulent dynamics as they fall. This basic dynamic framework qualitatively summarizes most of the results which are shown to line up well between data and the model. The actual data is too short and too noisy to reveal these patterns in much detail, but further analysis to see how well this simple description fits it will be important.

The model is still not perfect as it generates some interesting counterfactuals. The most obvious of

these is that the long range boom and bust cycle for prices and dividends is far more predictable from the standpoint of regression R-squared. In the data, movements in prices and dividends display a more erratic pattern. The agent-based market is not perfectly cyclical, but it does have generally more regular cycles in the price/dividend - return pattern than for the data. Similarly, the volatility regularity with market extremes is much more pronounced for the simulation. It remains to be seen exactly what this is coming from. It could simply be the larger amount of data from the simulated market, or a much more regular process (geometric random walk) for the dividend series than is reasonable in the data.⁴¹

It should be noted that the model does contain several design features which appear to contribute to the dynamics in important ways, and these may be fundamental. They include the explicit modeling of risk and return. This mechanism, along with heterogeneous gain levels on the agents' volatility forecasts is able to drive a significant fraction of the long range movements in the price/dividend ratio. This allows for long swings in prices relative to fundamentals without the need for a large flow of agents from one strategy to the next. Strategy shifting flows can be relatively modest which may be closer to the data.

The empirical methodology for the paper is new. The agent-based model is used as a super long range data generator giving sample lengths which are long enough so that nonparametric kernel density methods are able to lock in on the key features for the simulated data. The simulation based time series are used for bandwidth estimation. This is a form of using the simulation as a model selection tool. Once the model specification is chosen, it can be taken to the actual data. It is a weaker form of validation than exact moment matching, but it might still be useful, and could have much more practical value in helping to sort through various empirical specifications for applications to real data series.

The world of agent-based financial modeling remains a complicated area with many competing model specifications. Much progress has been made theoretically in understanding their dynamics, but less progress has been made in terms of narrowing the space of models down by matching to data. This paper moves in this direction by fitting features at both short and long horizons, and demonstrating that a relatively simple new agent-based model is capable of matching a large range of facts.

⁴¹ Remember that agents are assumed to completely understand the dividend process in this market. Weakening this assumption might be important for this.

References

- Andersen, T. G., Bollerslev, T., Christoffersen, P. F. & Diebold, F. X. (2006), Volatility and correlation forecasting, in G. Elliott, C. W. J. Granger & A. Timmermann, eds, 'Handbook of Economic Forecasting', North Holland, pp. 778–878.
- Anufriev, M. & Dindo, P. (2010), 'Wealth-driven selection in a financial market with heterogeneous agents', *Journal of Economic Behavior and Organization* **73**, 327–358.
- Black, F. (1976), 'Studies of stock price volatility changes', *Proceedings of the American Statistical Association, Business and Economics Statistics Section* pp. 177–181.
- Bollerslev, T., Engle, R. F. & Nelson, D. B. (1995), ARCH models, in R. F. Engle & D. McFadden, eds, 'Handbook of Econometrics', Vol. 4, North-Holland, New York, NY, pp. 2959–3038.
- Boswijk, H. P., Hommes, C. H. & Manzan, S. (2007), 'Behavioral heterogeneity in stock prices', *Journal of Economic Dynamics and Control* **31**(6), 1938–1970.
- Boudoukh, J., Michaely, R., Richardson, M. & Roberts, M. R. (2007), 'On the importance of measuring payout yield: Implications for empirical asset pricing', *Journal of Finance* **58**(2), 877–915.
- Branch, W. A. & Evans, G. W. (2011), 'Learning about risk and return: A simple model of bubbles and crashes', *American Economic Journal: Macroeconomics* **3**, 159–191.
- Brock, W. A., Lakonishok, J. & LeBaron, B. (1992), 'Simple technical trading rules and the stochastic properties of stock returns', *Journal of Finance* **47**, 1731–1764.
- Cagan, P. (1956), The monetary dynamics of hyperinflation, in M. Friedman, ed., 'Studies in the quantity theory of money', University of Chicago Press, Chicago, IL, pp. 25–120.
- Campbell, J. & Shiller, R. (1988), 'Stock prices, earnings, and expected dividends', *Journal of Finance* **63**, 661–676.
- Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1996), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.
- Campbell, J. Y. & Viceira, L. M. (1999), 'Consumption and portfolio decisions when expected returns are time varying', *Quarterly Journal of Economics* **114**, 433–495.

- Campbell, J. Y. & Viceira, L. M. (2002), *Strategic Asset Allocation*, Oxford University Press, Oxford, UK.
- Carhart, M. M. (1997), 'On persistence in mutual fund performance', *Journal of Finance* **52**(1), 57–82.
- Chan, L. K. C., Jegadeesh, N. & Lakonishok, J. (1996), 'Momentum strategies', *Journal of Finance* **51**, 1681–1713.
- Chen, S.-H. (2012), 'Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective', *Journal of Economic Dynamics and Control* **36**, 1–25.
- Chiarella, C., Dieci, R. & He, X.-Z. (2009), Heterogeneity, market mechanisms, and asset price dynamics, in T. Hens & K. R. Schenk-Hoppe, eds, 'Handbook of Financial Markets: Dynamics and Evolution', Elsevier, USA, pp. 277–344.
- Chiarella, C., He, X.-Z., Huang, W. & Zheng, H. (2012), 'Estimating behavioral heterogeneity under regime switching', *Journal of Economic Behavior and Organization* **83**, 446–460.
- Chiarella, C., He, X.-Z. & Zwinkels, R. C. J. (2009), Heterogeneous expectations in asset pricing: Empirical evidence from the s&p 500, Technical report, University of Technology, Sydney, Australia.
- Christie, A. A. (1982), 'The stochastic behavior of common stock variances: Value, leverage and interest rate effects', *Journal of Financial Economics* **10**, 407–432.
- Cochrane, J. (2008), 'The dog that did not bark: A defense of return predictability', *Review of Financial Studies* **21**(4), 1533–1575.
- Cochrane, J. H. (2011), 'Discount rates', *Journal of Finance* **66**(4), 1047–1109.
- Corsi, F. (2009), 'A simple approximate long-memory model of realized volatility', *Journal of Financial Econometrics* **7**(2), 174–196.
- Dacorogna, M. M., Gencay, R., Muller, U. A., Olsen, R. B. & Pictet, O. V. (2001), *An Introduction to High-Frequency Finance*, Academic Press, San Diego, CA.
- Ding, Z., Granger, C. W. & Engle, R. F. (1993), 'A long memory property of stock market returns and a new model', *Journal of Empirical Finance* **1**, 83–106.
- Engle, R. F. & Ng, V. K. (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance* **48**, 1749–1778.

- Evans, G. W. & Honkapohja, S. (2001), *Learning and Expectations in Macroeconomics*, Princeton University Press.
- Fisher, K. L. & Statman, M. (2006), 'Market timing in regressions and reality', *Journal of Financial Research* **29**(3), 293–304.
- Frankel, J. A. & Froot, K. A. (1987), 'Using survey data to test standard propositions regarding exchange rate expectations', *American Economic Review* **77**(1), 133–153.
- French, K. R., Schwert, G. W. & Stambaugh, R. F. (1987), 'Expected stock returns and volatility', *Journal of Financial Economics* **19**, 3–29.
- Friedman, M. (1956), *A Theory of the Consumption Function*, Princeton University Press, Princeton, NJ.
- Gaunersdorfer, A. (2000), 'Endogenous fluctuations in a simple asset pricing model with heterogeneous agents', *Journal of Economic Dynamics and Control* **24**, 799–831.
- Georges, C. (2008), 'Staggered updating in an artificial market', *Journal of Economic Dynamics and Control* **32**, 2809–2825.
- Giovannini, A. & Weil, P. (1989), Risk aversion and intertemporal substitution in the capital asset pricing model, Technical Report 2824, National Bureau of Economic Research, Cambridge, MA.
- Glosten, L. R., Jagannathan, R. & Runkle, D. E. (1993), 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance* **48**, 1779–1801.
- Goyal, A. & Welch, I. (2003), 'Predicting the equity premium with dividend ratios', *Management Science* **49**(5), 639–654.
- Granger, C. W. (1980), 'Long memory relationships and the aggregation of dynamic models', *Journal of Econometrics* **14**, 227–238.
- Hasanhodzic, J. & Lo, A. W. (2011), Black's leverage effect is not due to leverage, Technical report, Boston University.
- Hommes, C. H. (2006), Heterogeneous agent models in economics and finance, in L. Tesfatsion & K. L. Judd, eds, 'Handbook of Computational Economics', Elsevier, pp. 1109–1186.

- Hommes, C. H. (2011), 'The heterogeneous expectations hypothesis: Some evidence from the lab', *Journal of Economic Dynamics and Control* **35**, 1–24.
- Hommes, C. H. & Wagener, F. (2009), Complex evolutionary systems in behavioral finance, in T. Hens & K. R. Schenk-Hoppe, eds, 'Handbook of Financial Markets: Dynamics and Evolution', North-Holland, pp. 217–276.
- Hong, H. & Stein, J. C. (1999), 'A unified theory of underreaction, momentum trading, and overreaction in asset markets', *Journal of Finance* **54**, 2143–2184.
- Honkapohja, S. & Mitra, K. (2006), 'Learning stability in economies with heterogeneous agents', *Review of Economic Dynamics* **9**, 284–309.
- Jegadeesh, N. & Titman, S. (1993), 'Returns to buying winners and selling losers: Implications for stock market efficiency', *Journal of Finance* **48**, 65–91.
- Kouwenberg, R. & Zwinkels, R. C. J. (2010), Chasing trends in the U.S. housing market, Technical report, Erasmus University, Rotterdam, The Netherlands.
- LeBaron, B. (1992), 'Some relations between volatility and serial correlations in stock market returns', *Journal of Business* **65**(2), 199–219.
- LeBaron, B. (2001a), 'Empirical regularities from interacting long and short memory investors in an agent based stock market', *IEEE Transactions on Evolutionary Computation* **5**, 442–455.
- LeBaron, B. (2001b), 'Stochastic volatility as a simple generator of apparent financial power laws and long memory', *Quantitative Finance* **1**, 621–631.
- LeBaron, B. (2002), 'Short-memory traders and their impact on group learning in financial markets', *Proceedings of the National Academy of Science: Colloquium* **99**(Supplement 3), 7201–7206.
- LeBaron, B. (2006), Agent-based computational finance, in L. Tesfatsion & K. L. Judd, eds, 'Handbook of Computational Economics', Elsevier, pp. 1187–1233.
- LeBaron, B. (2012), 'Heterogeneous gain learning and the dynamics of asset prices', *Journal of Economic Behavior and Organization* **83**(3), 424–445.
- LeBaron, B. (2013), The impact of heterogeneous gain learning in an agent-based financial market, Technical report, International Business School, Brandeis University.

- LeBaron, B., Arthur, W. B. & Palmer, R. (1999), 'Time series properties of an artificial stock market', *Journal of Economic Dynamics and Control* **23**, 1487–1516.
- Lettau, M. & Ludvigson, S. (2010), Measuring and modeling variation in the risk-return trade-off, in Y. Ait-Shalia & L. P. Hansen, eds, 'Handbook of Financial Econometrics', Vol. 1, Elsevier, pp. 617–690.
- Levy, M., Levy, H. & Solomon, S. (1994), 'A microscopic model of the stock market: cycles, booms, and crashes', *Economics Letters* **45**, 103–111.
- Lux, T. (2009), Stochastic behavioral asset pricing models and the stylized facts, in T. Hens & K. R. Schenk-Hoppe, eds, 'Handbook of Financial Markets: Dynamics and Evolution', North-Holland, pp. 161–215.
- Lynch, P. & Zumbach, G. (2003), 'Market heterogeneities and the causal structure of volatility', *Quantitative Finance* **3**, 320–331.
- Mitra, K. (2005), 'Is more data better?', *Journal of Economic Behavior and Organization* **56**, 263–272.
- Muth, J. (1960), 'Optimal properties of exponentially weighted forecasts', *Journal of the American Statistical Association* **55**, 299–306.
- Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**, 347–370.
- Nelson, D. B. (1992), 'Filtering and forecasting with misspecified ARCH modelsI: Getting the right variance with the wrong model', *Journal of Econometrics* **52**, 61–90.
- Pastor, L. & Stambaugh, R. F. (2009), 'Predictive systems: Living with imperfect predictors', *Journal of Finance* **64**, 1583–1628.
- Pettenuzzo, D., Timmerman, A. & Valkanov, R. (2013), Forecasting stock returns with economic constraints, Technical report, Department of Economics, Brandeis University.
- Phelps, E. S. (1967), 'Phillips curves, expectations of inflation and optimal unemployment over time', *Economica* **34**(135), 254–281.
- Rossi, A. & Timmermann, A. (2010), What is the shape of the risk-return relation?, Technical report, UC - San Diego.
- Sargent, T. (1999), *The Conquest of American Inflation*, Princeton University Press, Princeton, NJ.

- Schwert, G. W. (1990), 'Indexes of United States stock prices from 1802-1987', *Journal of Business* **63**, 399–426.
- Shiller, R. J. (2000), *Irrational Exuberance*, Princeton University Press, Princeton, NJ.
- Taylor, M. & Allen, H. (1992), 'The use of technical analysis in the foreign exchange market', *Journal of International Money and Finance* **11**(3), 304–14.
- Thurner, S., Dockner, E. J. & Gaunersdorfer, A. (2002), Asset price dynamics in a model of investors operating on different time horizons, Technical report, University of Vienna.
- Zschischang, E. & Lux, T. (2001), 'Some new results on the Levy, Levy, and Solomon microscopic stock market model', *Physica A* **291**, 563–573.

Table 1: **U.S. Income Series**

| | Dividends | Adjusted dividends | Earnings | GDP | Consumption |
|--------|-----------|--------------------|----------|------|-------------|
| Growth | 1.06 | 1.77 | 1.91 | 3.10 | 2.18 |
| Std | 0.09 | 0.12 | 0.31 | 0.03 | 0.02 |

All data are U.S. real annualized series. Mean growth rates are in percentages. The financial series come from the restricted Shiller data set, and cover 1886-2012. Real GDP and Consumption (nondurables) are from Fred quarterly data, 1947-2012, seasonally adjusted.

Table 2: *Parameter Definitions*

| Parameter | Value |
|------------------------|-----------------------|
| d_g | 0.018 |
| σ_d | 0.12 |
| r_f | 0 |
| γ | 3.5 |
| λ | 0.0005 |
| I | 16000 |
| J | 4000 |
| g_j | [100,20,10,5,2] years |
| g_L | 200 years |
| g_u | [100,20,10,5,2] years |
| L | 5 percent/year |
| $[\alpha_L, \alpha_H]$ | [0.01,0.99] |
| σ_ϵ | 0.01 |
| M_{AR} | 3 |
| h_j | [0.025,0.10] |

The annual standard deviation of dividend growth is set to the level from real dividends in Shiller's annual long range data set adjusted for share repurchases.

Table 3: *Annual Excess Return Summary Statistics*

| | Merged U.S. 1886-2012 | Multiple Gain | Small Gain |
|-------------------|--------------------------|------------------|---------------|
| Mean($R_{e,t}$) | 7.3 | 9.9 | 5.3 |
| Std.($R_{e,t}$) | 19.1 | 22.3 | 8.4 |
| Mean(P_t/D_t) | 22.5 | 20.1 | 36.7 |
| Std(P_t/D_t) | 6.6 | 5.3 | 1.9 |

Table 4: *Weekly Return Summary Statistics*

| Series | CRSP (1926-2012) | Multiple gain | Small gain |
|----------|------------------|---------------|------------|
| Skewness | -0.5 | -0.3 | -0.0 |
| Kurtosis | 8.4 | 10.9 | 3.4 |

$$R_{t+1} - R_{f,t+1} = \alpha + \beta \log(P_t/D_t)$$

Table 5: **Price/Dividend Regressions**

| Series | β | R^2 |
|-------------------------|----------|-------|
| Merged U.S. (quarterly) | -0.04 | 0.01 |
| | (-4.56) | |
| Merged U.S. (annual) | -0.19 | 0.07 |
| | (-21.11) | |
| Simulation (quarterly) | -0.08 | 0.05 |
| | (-38.00) | |
| Simulation (annual) | -0.32 | 0.20 |
| | (-79.25) | |

$$\sigma_t^2 = \alpha + \beta \log(P_t/D_t)$$

$$\log(\sigma_t^2) = \alpha + \beta \log(P_t/D_t)$$

Table 6: **Price/Dividend and Volatility**

| Series | β | R^2 |
|--------------------------------|-----------------------|-------|
| Merged U.S. | $-0.18 \cdot 10^{-2}$ | 0.01 |
| | (-4.50) | |
| Merged U.S. ($\log(\sigma)$) | -0.39 | 0.12 |
| | (-4.33) | |
| Simulation | $-1.84 \cdot 10^{-2}$ | 0.28 |
| | (-92) | |
| Simulation ($\log(\sigma)$) | -2.62 | 0.34 |
| | (-131) | |

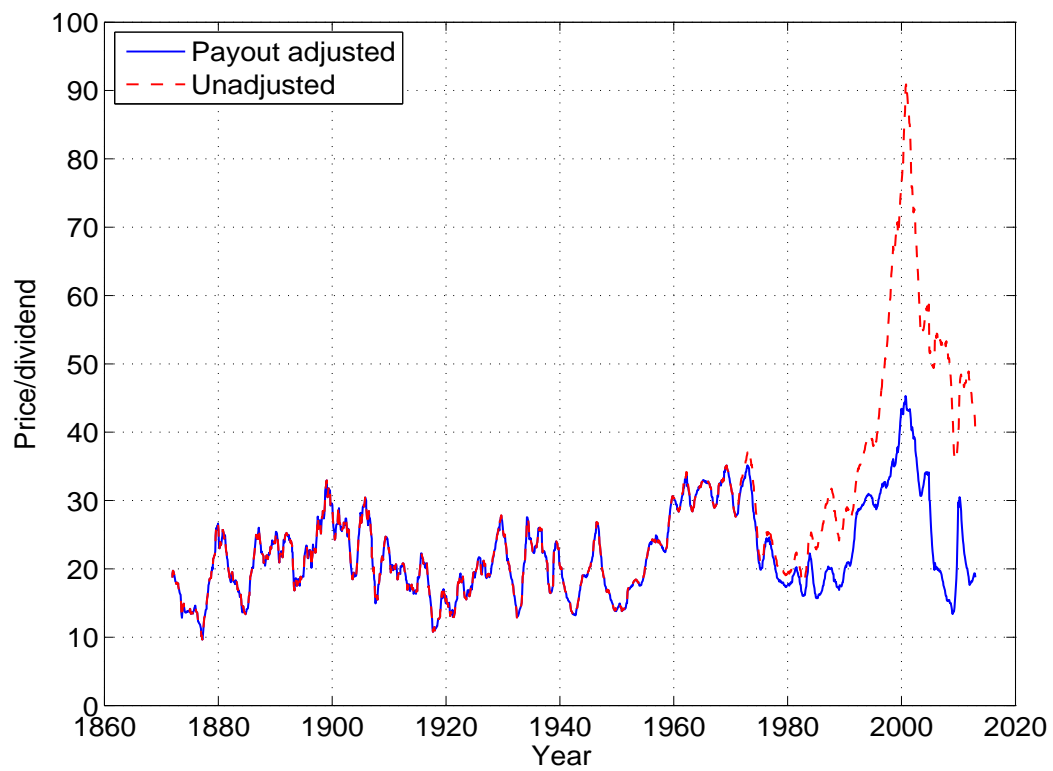


Figure 1: Price/Dividend ratio

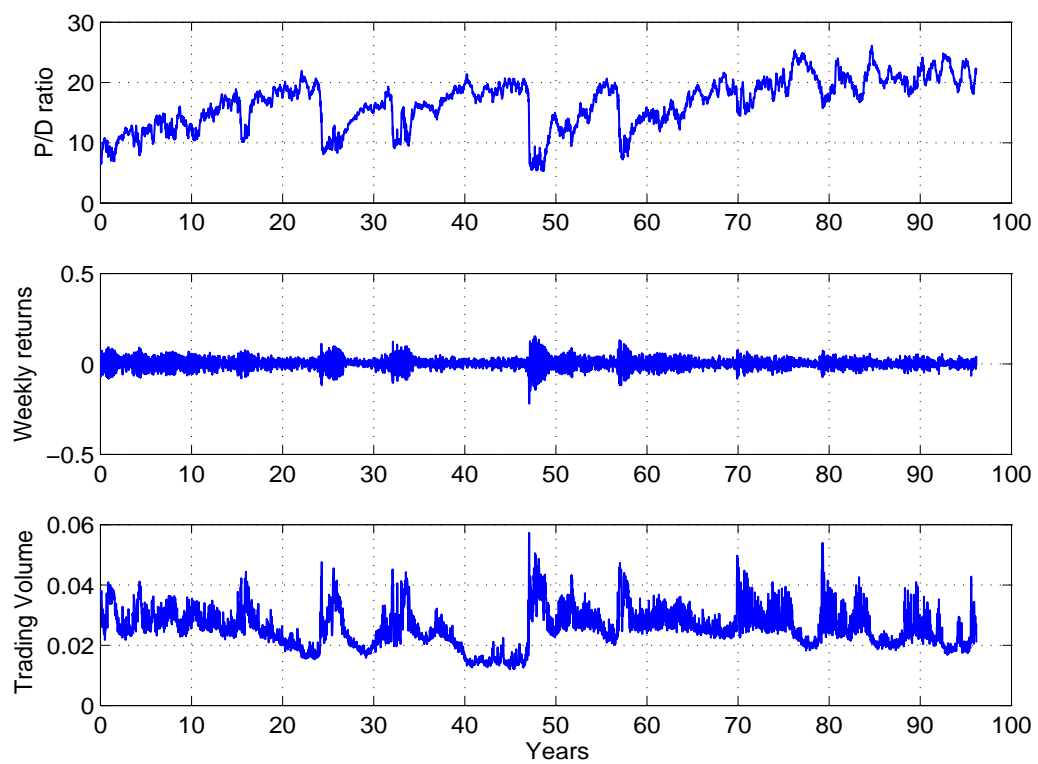


Figure 2: **Benchmark (multiple gain) simulation**

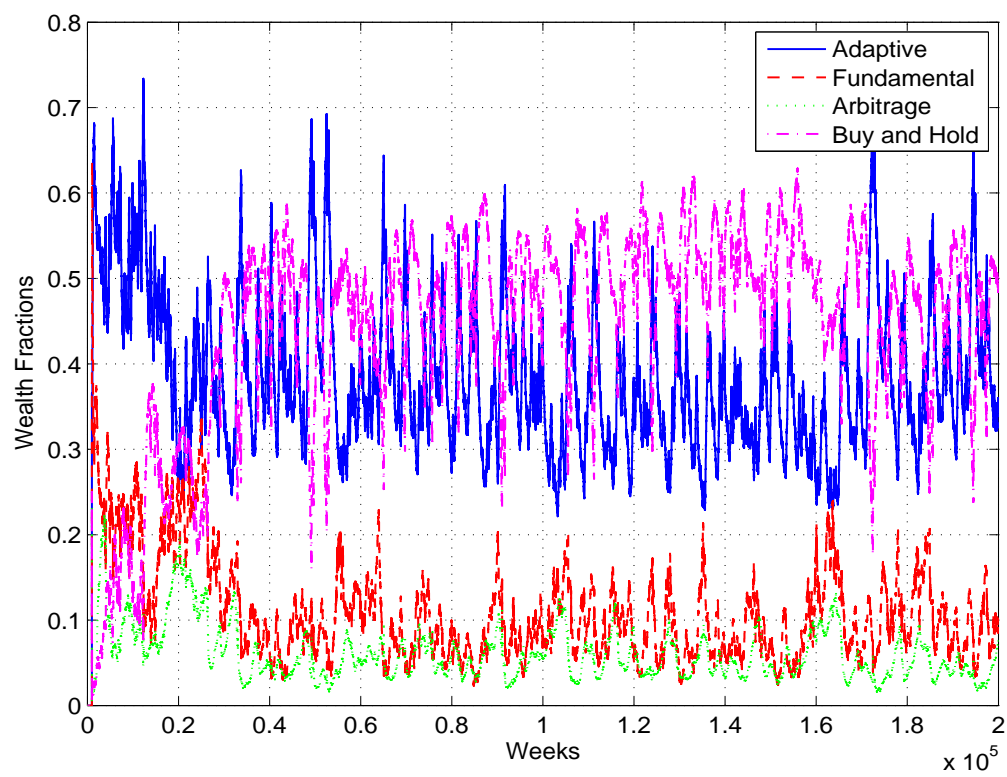


Figure 3: **Wealth Fractions**

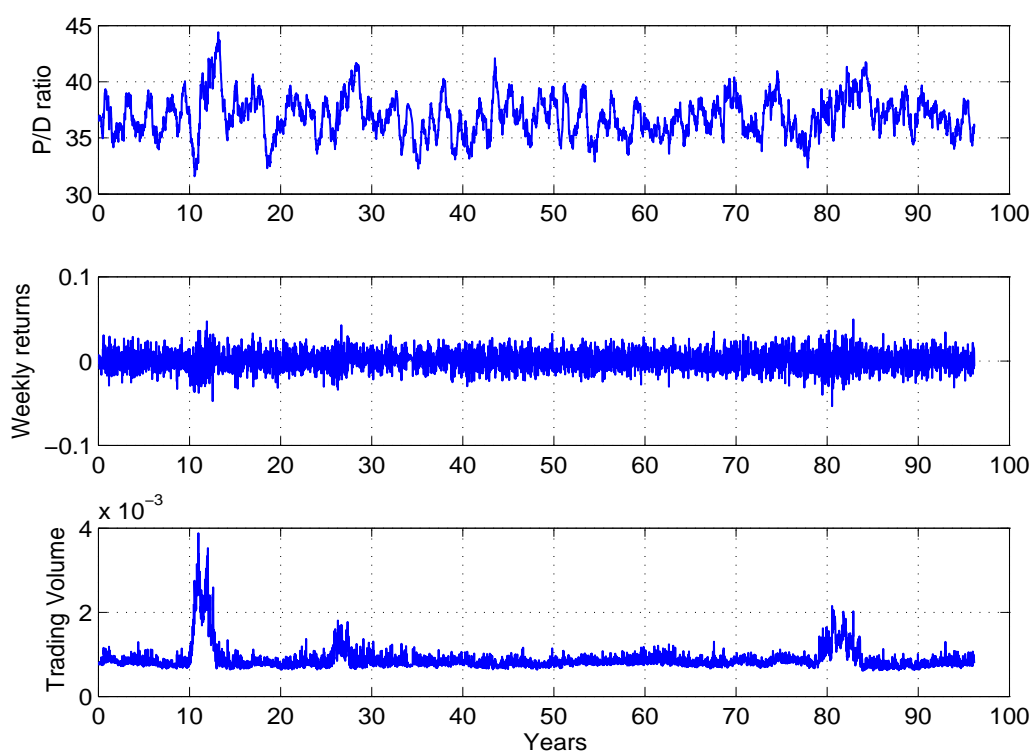


Figure 4: **Benchmark (small gain) simulation**

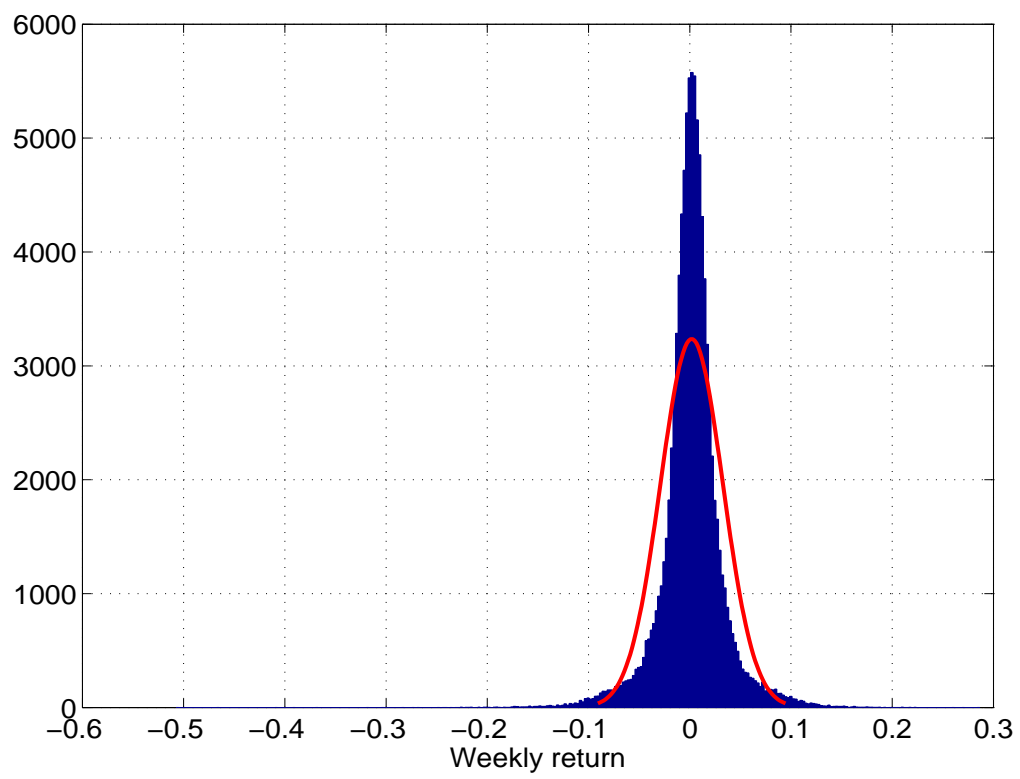


Figure 5: **Weekly returns: Multiple gain simulation**

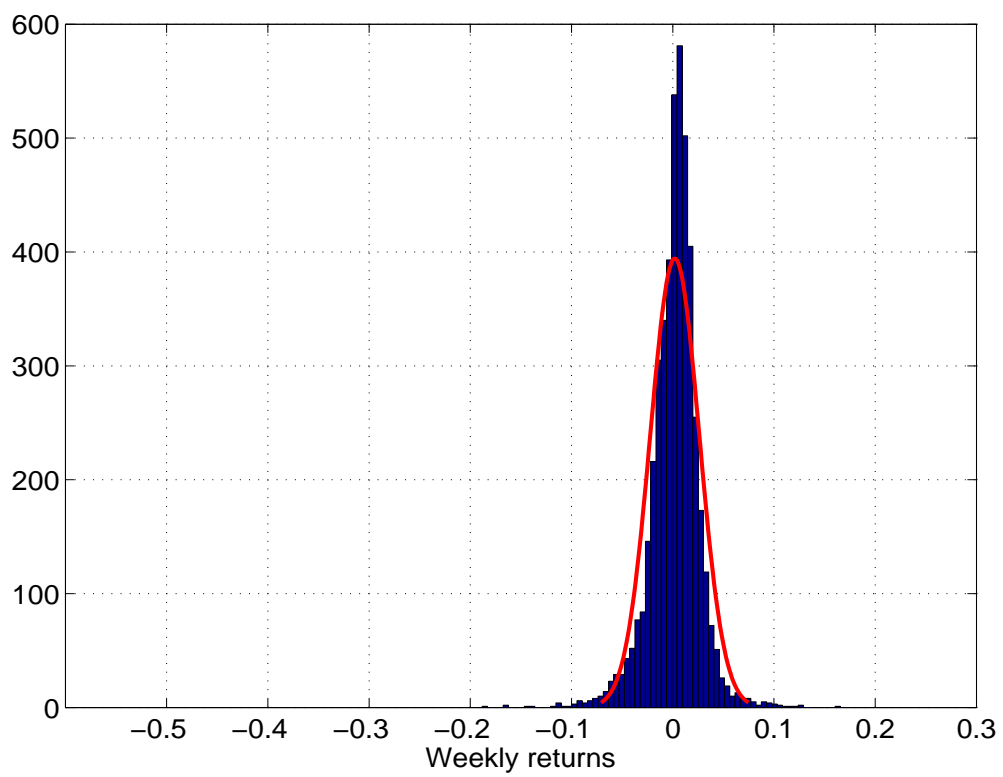


Figure 6: **Weekly returns: CRSP 1926-2012**

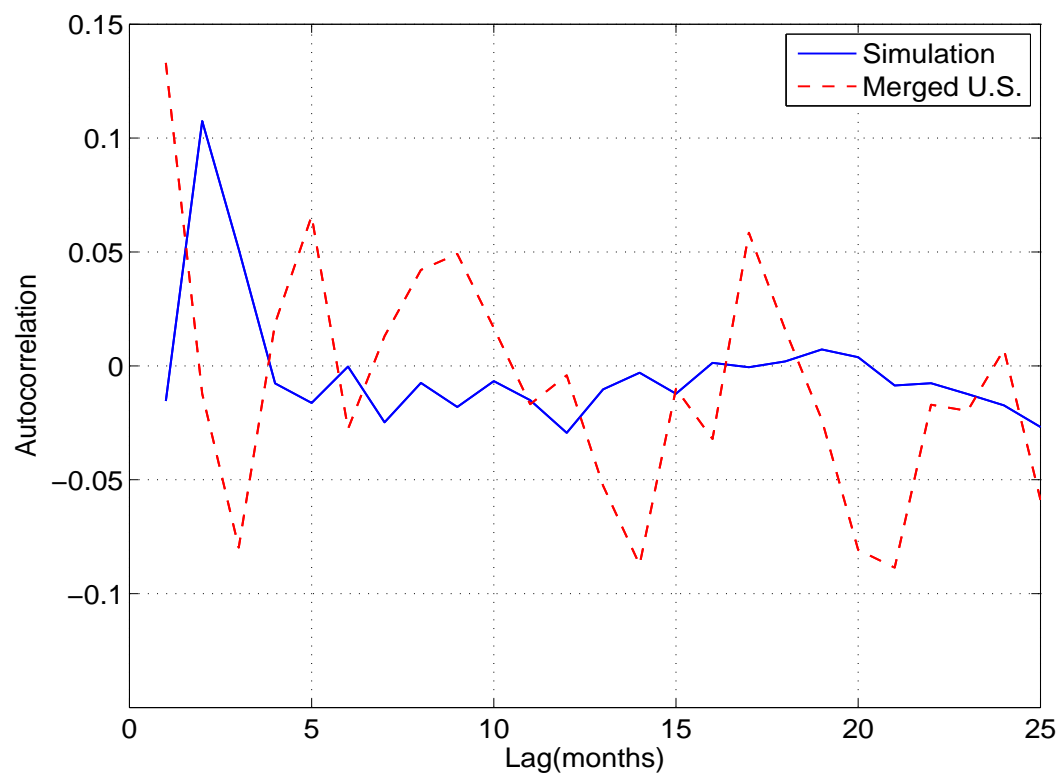


Figure 7: **Monthly return autocorrelations**

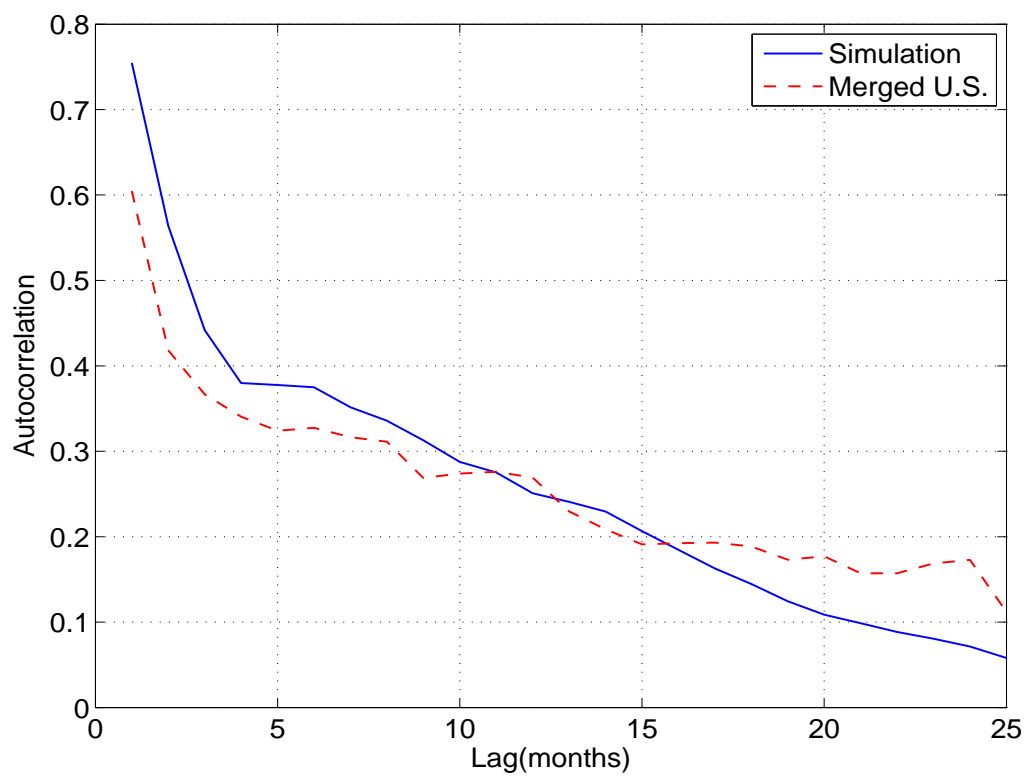


Figure 8: **Monthly volatility autocorrelations**

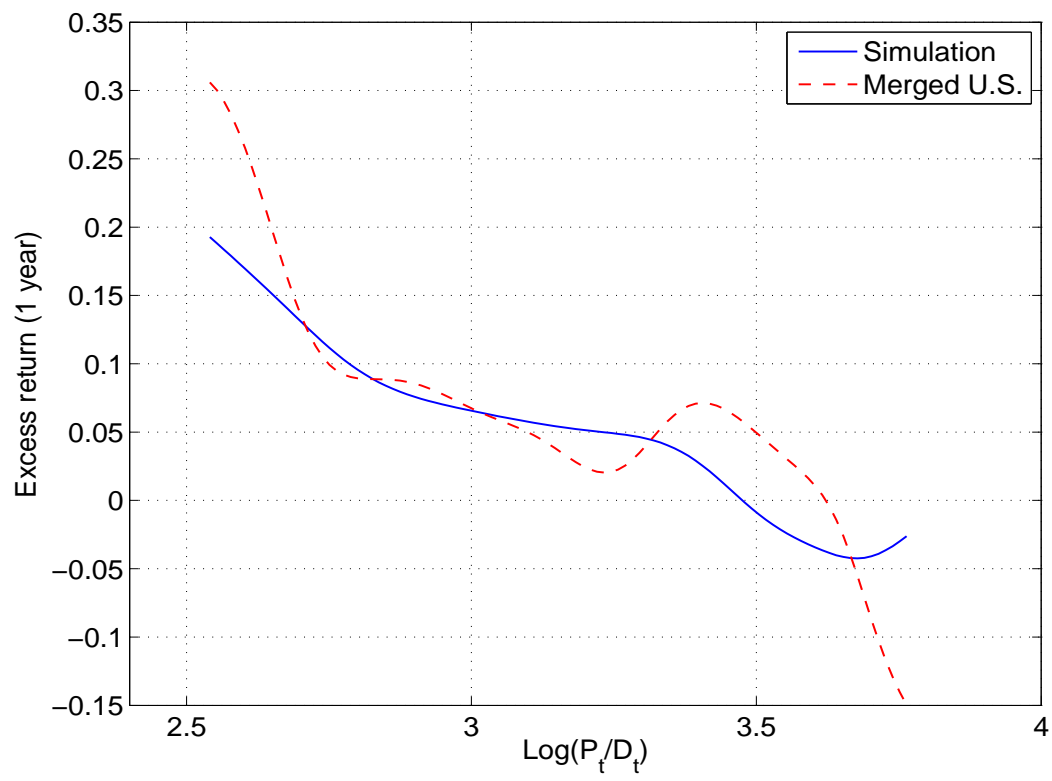


Figure 9: **Kernel regression: $\log(P_t/D_t)$ versus ER_{t+1}**

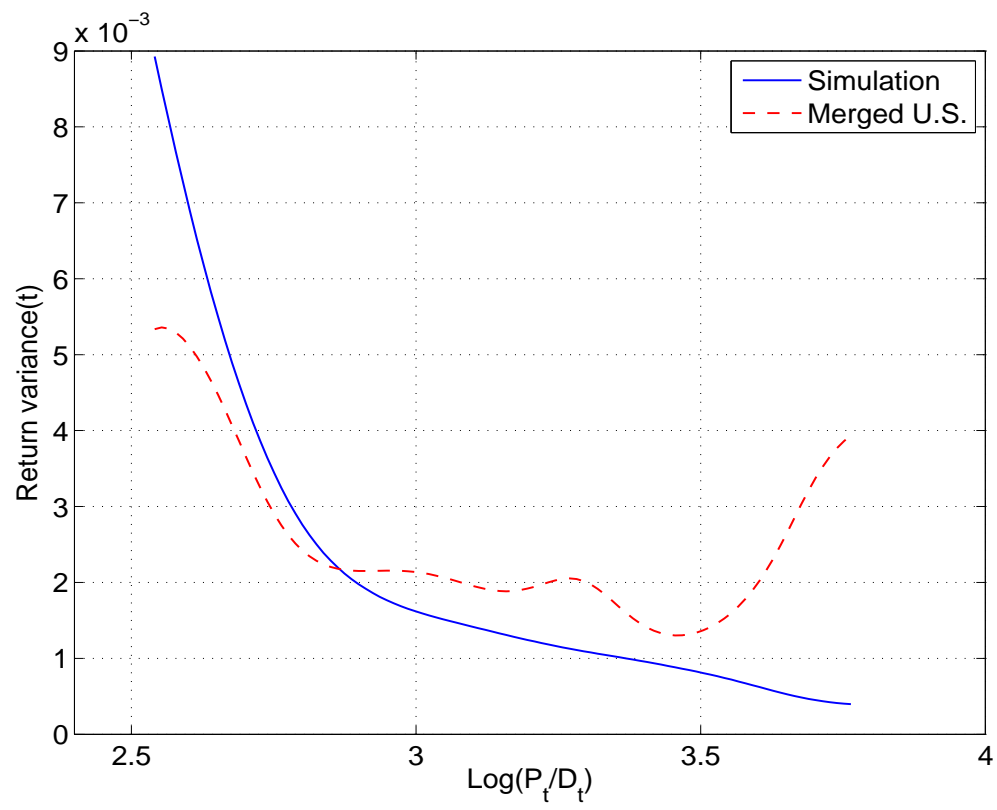


Figure 10: **Kernel regression: $\log(P_t/D_t)$ versus σ_t^2**

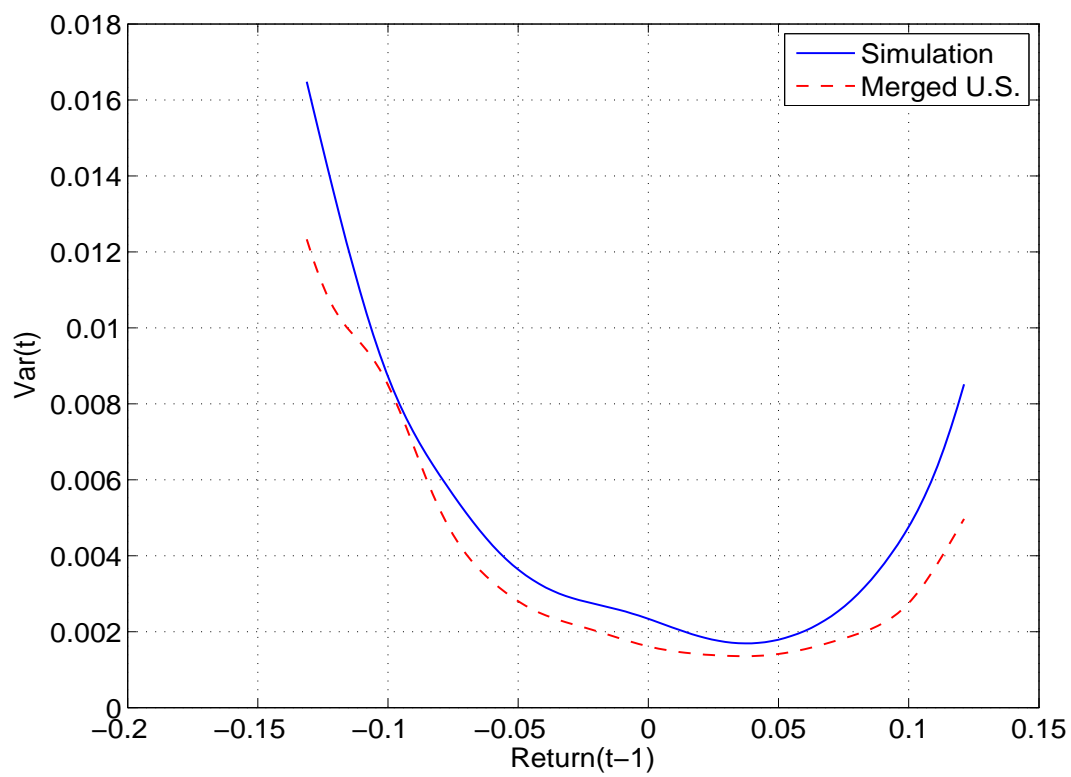


Figure 11: **News impact curve:** R_{t-1} and σ_t^2

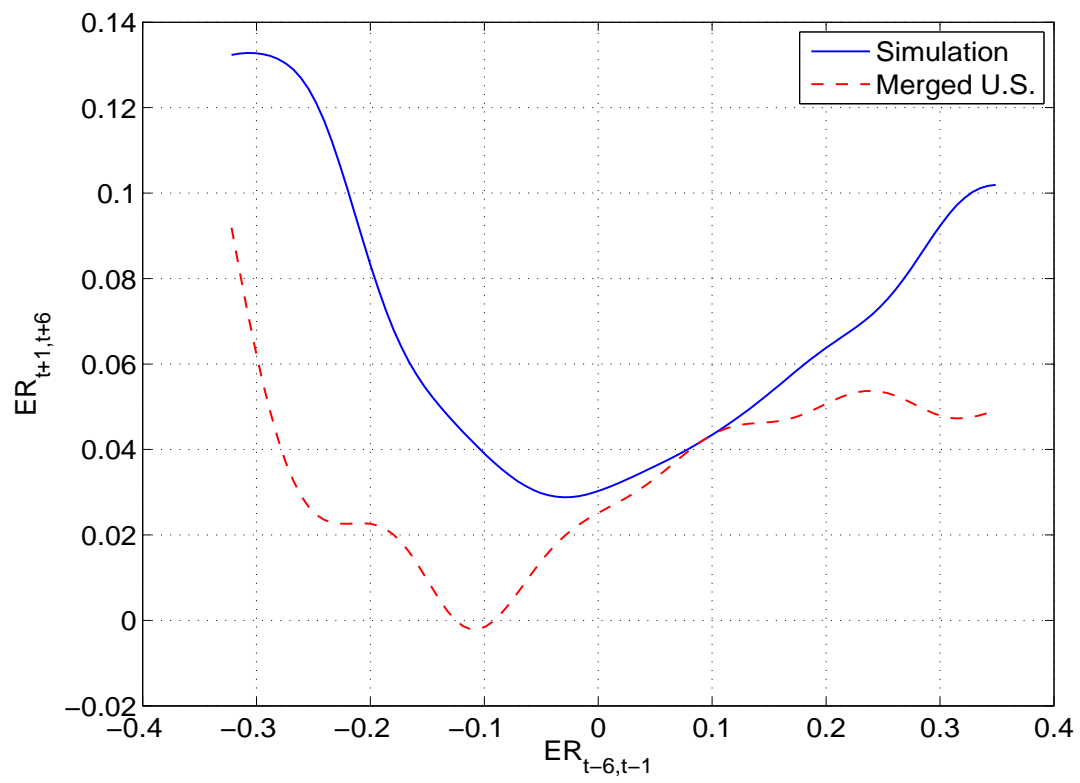


Figure 12: **Momentum**: 6 month - 1 month - 6 month forecast

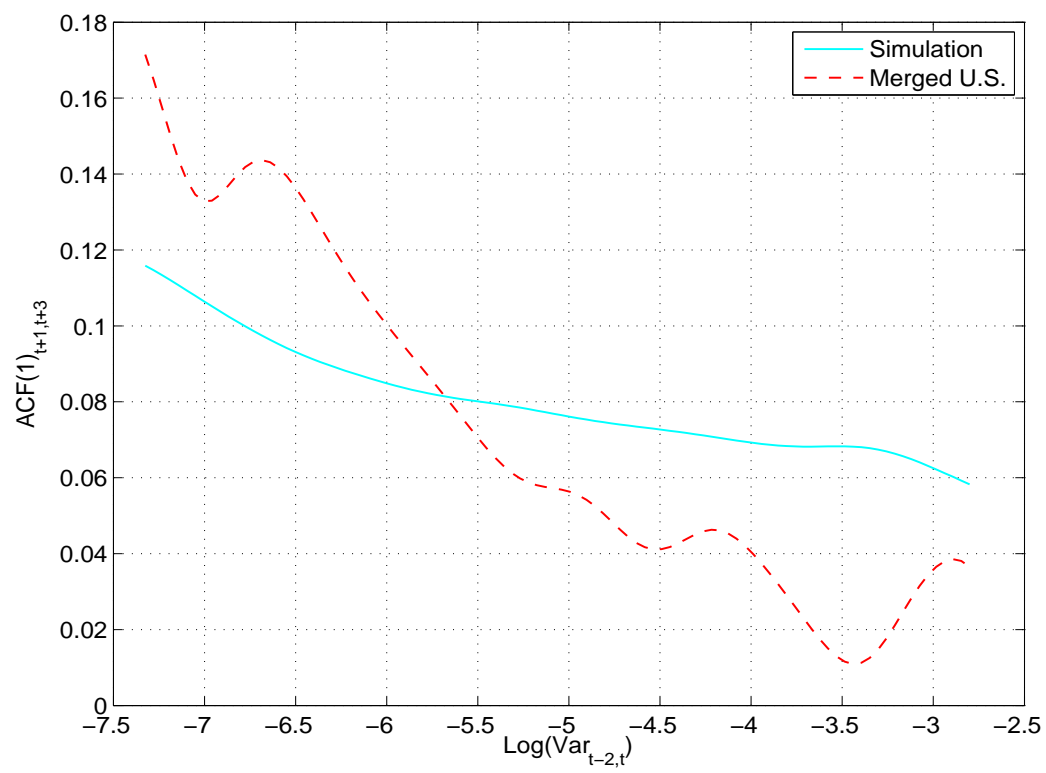


Figure 13: **Conditional autocorrelations**

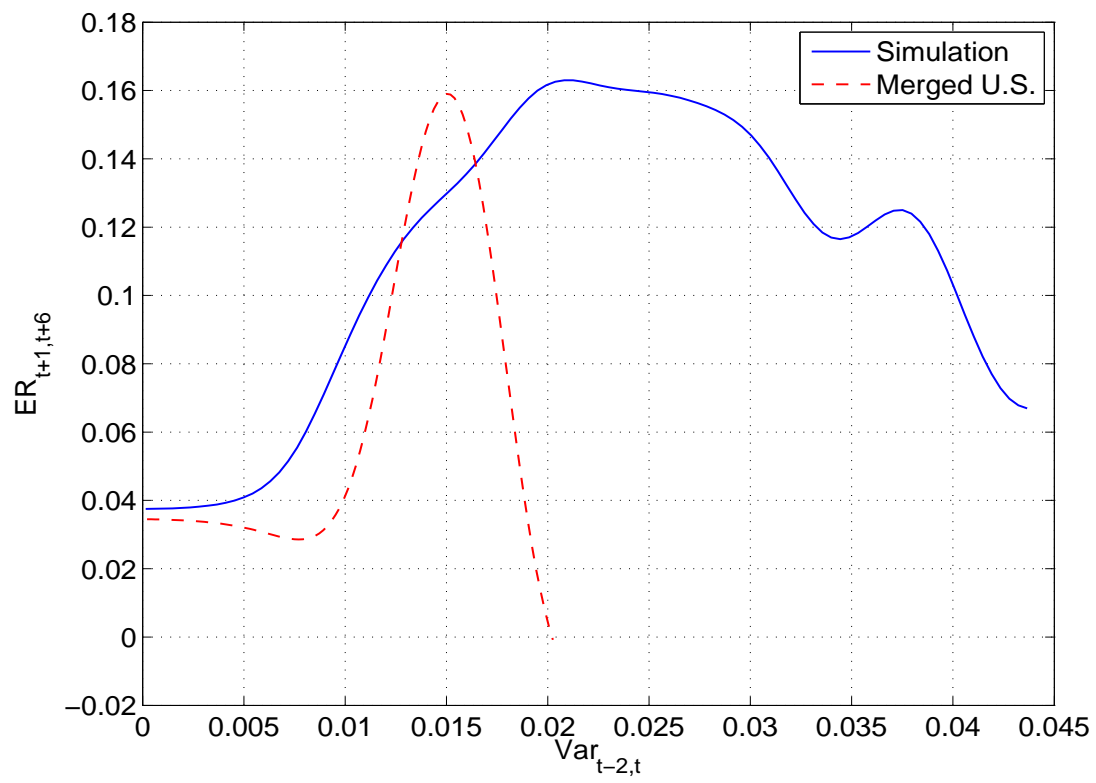


Figure 14: Risk and Expected Returns

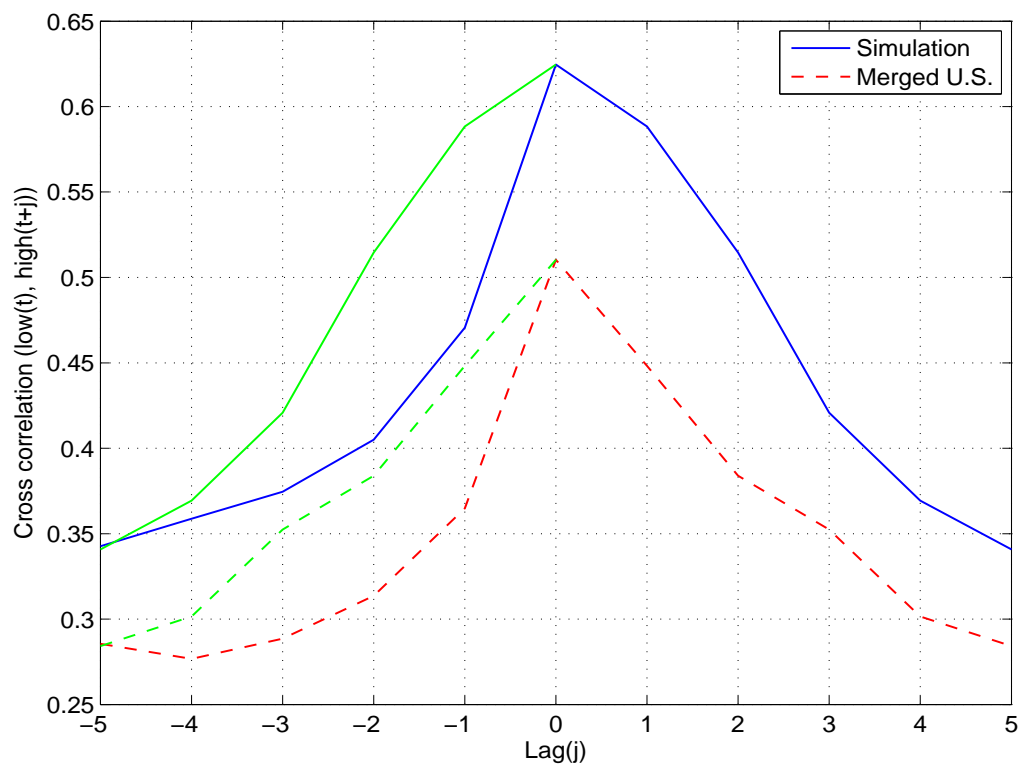


Figure 15: **Volatility asymmetry** : Course versus fine volatility measures. Green lines display reflected leading estimates