

# The Optimality of Short Term Rules of Thumb at Long Horizons for an Agent-based Financial Market

Blake LeBaron \*  
Brandeis University

February 2017

## Abstract

Agent-based financial markets have been extremely successful at matching features of financial markets at relatively short horizons, but they are often silent about long range features that cross into the area of macro/financial models. This is an area where they are often criticized by many in mainstream macroeconomics. This paper takes a relatively simple market which is capable of matching all the usual short horizon facts, and examines comovements in consumption and asset returns at longer horizons. This is done through the perspective of traditional intertemporal preferences. The key test is to see how close the simple ad hoc (but myopically optimizing) rules come to more traditional models for macroeconomic consumers. It is calibrated to exogenous inputs of both dividend and labor income. When generating high frequency data the model is relatively consistent with intertemporal optimization. However, at longer horizons it fails to hit the theoretical optimization values, but it fails in ways that are very close to the failures of representative agent consumption based models.

Keywords: Learning, agent-based modeling, asset pricing

---

\*International Business School, Brandeis University, 415 South Street, Mailstop 32, Waltham, MA 02453 - 2728, [blebaron@brandeis.edu](mailto:blebaron@brandeis.edu), [www.brandeis.edu/~blebaron](http://www.brandeis.edu/~blebaron).

# 1 Introduction

Heterogeneous agent models with adaptive populations of learning agents have proved successful at matching many empirical features of financial data.<sup>1</sup> However, their own internal dynamic complexity has typically restricted them to very simple myopic preferences for individuals in the model. This simplification questions their usefulness. Could their results be coming from the fact that agents are making basic intertemporal mistakes that would unlikely be made by actual people? Could the simple ad hoc rules of thumb that are assumed be a useful approximation to intertemporal optimization? This paper is a first attempt to try to deal with these questions and build a bridge between most agent-based models, and more traditional models in macro finance.

In its comparisons with more traditional macro finance research this paper will turn to lower frequency macro features, and analysis of first order condition violations from a cross section of heterogeneous agents.<sup>2</sup> Violation of standard intertemporal Euler equations is important, since it indicates that there is some distance between these models and most of standard macro modeling. Eventually, this distance needs to be justified either empirically experimentally.<sup>3</sup>

The consumer modeling used here draws heavily from Campbell & Viceira (2002). Their framework is important because it incorporates both intertemporal and myopic decision making for agents. Their log linearized solutions for standard multi-period consumers can be simplified to a myopic agent problem subject to various restrictive assumptions. Therefore, in using their solutions there is some justification for the agents used here, and how they may need to be modified in the future.

The model also builds from various lines of research. In the first case, agents attempt to learn about both risk and return as in LeBaron, Arthur & Palmer (1999) and Branch & Evans (2011). These have proved to be an important aspect in generating interesting financial dynamics, and is intuitively plausible given the amount of predictability in risk measures for financial time series. The model also continues with the heterogeneous gain framework from LeBaron (2012).<sup>4</sup> Finally, this paper may be one of the first to bring

---

<sup>1</sup> 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.

<sup>2</sup> Many simple agent-based models have recently turned to longer range data. Several examples of this are Boswijk, Hommes & Manzan (2007), Kouwenberg & Zwinkels (2010), Chiarella, He & Zwinkels (2014), and Chiarella, He, Huang & Zheng (2012).

<sup>3</sup> A very early approach which looks at the cost of deviations from Euler equations is Cochrane (1989). Though the present work is very different, it shares some of the spirit of this earlier paper which shows that some rule of thumb behaviors might be near optimal.

<sup>4</sup> 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. This same environment is present here.

in both labor and asset income into a heterogeneous agent learning framework. Agents receive labor income as well as dividend income from their equity holdings. The payout to dividends versus labor income dynamics is calibrated to U.S. data on labor income share dynamics. Also, cross-sectional inequality is replicated by matching some of the features in Piketty & Saez (2003) upper decile labor shares over time. In both of these cases, later data is ignored where these measures begin to move dramatically in the 2000's. This nonstationary regime shift is beyond the scope of this model.

Section 2 overviews the model structure. Section 3 summarizes some basic empirical features. Section 4 turns to the specific Euler equation tests at both high and low frequencies, and section 5 concludes.

## 2 Model structure and calibration

### 2.1 Labor and dividend income

The market contains two exogenous macro income streams which are calibrated to U.S. macro data. First, aggregate dividend flows are given (in logs) by,<sup>5</sup>

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

and the aggregate dividend is therefore,

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

Aggregate labor income is assumed to be cointegrated with the dividend process as in

$$l_t = d_t + y_t \quad (3)$$

$$y_t = \rho y_{t-1} + \epsilon_{y,t} \quad (4)$$

$$L_t = e^{l_t} \quad (5)$$

where  $y_t$  represents a stationary idiosyncratic component to total labor income. It has been calibrated to long range labor/capital income ratios in the U.S, but ignoring the drop in this ratio occurring after 1990. In the model the long range ratio is one given  $E(y_t) = 0$ . Labor income for each individual  $i$  in period  $t$  is then given by,

$$L_{t,i} = \omega_i L_t \quad (6)$$

$$\omega_i \sim e^{\epsilon_L} \quad (7)$$

$$\epsilon_L \sim N(0, \sigma_L), \quad (8)$$

---

<sup>5</sup>Lower case letters represent log values.

where the log normal labor share given by  $\omega_i$  is designed to simulate skewness in labor income. The value for  $\sigma_L$  is set to 0.83 which generates a top decile labor income share of about 1/3 which aligns with results across the data in Piketty & Saez (2003), but does not capture increases in this value in recent years. For both labor income shares, and labor income inequality there is no attempt to capture any potentially nonstationary aspects of the data. This would add a level of complexity out of the bounds of the current paper. Also, both these values are assumed exogenous to the model.

## 2.2 Assets

The market consists of only two assets. First, there is a risky asset paying the previously defined dividend each period. Time periods in the model are incremented in units of weeks. The price of the stock is given by  $P_t$ .

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

$$R_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}}, \quad (9)$$

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,

$$\sum_{i=1}^I S_{t,i} = 1. \quad (10)$$

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 standard intertemporal budget constraint holds for each agent  $i$ ,

$$W_{t,i} = P_t S_{t,i} + B_{t,i} + C_{t,i} = (P_t + D_t) S_{t-1,i} + (1 + R_f) B_{t-1,i} + L_{t,i}, \quad (11)$$

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

## 2.3 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,

$$\max_{\alpha_{t,i}} \frac{E_t^i W_{t+1,i}^{1-\gamma}}{1-\gamma}, \quad (12)$$

$$st. \quad W_{t+1,i} = (1 + R_{t+1,i}^p)(W_{t,i} - C_{t,i}), \quad (13)$$

$$R_{t+1,i}^p = \alpha_{t,i} R_{t+1} + (1 - \alpha_{t,i}) R_f. \quad (14)$$

$\alpha_{t,i}$  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,

$$\alpha_{t,i} = \frac{E_t^i(r_{t+1}) - r_f + \frac{1}{2}\sigma_{t,i}^2}{\gamma\sigma_{t,i}^2} + \epsilon_{t,i}, \quad (15)$$

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 a normally distributed individual shock designed to make sure that there is some small amount of heterogeneity to keep trade operating.<sup>6</sup>

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 would be important, but adds significant model complexity. The constraints in  $\alpha_{t,i}$  are imposed through a smoothed logistic function which remaps the raw portfolio weights into the range. The function is set to be linear in the inner range of  $\alpha$ , and then nonlinear with the first derivative matched in the extreme versions of  $\alpha$ . A plot of this is given in figure 2.

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

$$W_{t+1,i} = (1 + R_{t+1}^p)(1 - \lambda)W_{t,i} + L_{t,i}. \quad (16)$$

This also gives the current period budget constraint,

$$P_t S_{t,i} + B_{t,i} = (1 - \lambda)((P_t + D_t)S_{t-1,i} + (1 + R_f)B_{t-1,i}) + L_{t,i}. \quad (17)$$

This simplified portfolio strategy will be used throughout. 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.<sup>7</sup>

<sup>6</sup>The derivation of this follows Campbell & Viceira (2002). It involves taking a Taylor series approximation for the log portfolio return. The shock,  $\epsilon_{t,i}$ , is distributed  $N(0, \sigma_\epsilon^2)$ .

<sup>7</sup>See 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.

## 2.4 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 equipped with the minimum gain level, denoted  $g_L$ .

$$\bar{r}_t = (1 - g_L)\bar{r}_{t-1} + g_L r_t \quad (18)$$

$$\overline{pd}_t = (1 - g_L)\overline{pd}_{t-1} + g_L pd_{t-1} \quad (19)$$

$$\bar{\sigma}_t^2 = (1 - g_L)\bar{\sigma}_{t-1}^2 + g_L (r_t - \bar{r}_t)^2 \quad (20)$$

$$\bar{\sigma}_{pd,t}^2 = (1 - g_L)\bar{\sigma}_{pd,t-1}^2 + g_L (pd_t - \overline{pd}_t)^2 \quad (21)$$

$$\overline{\Delta p}_t = (1 - g_L)\overline{\Delta p}_{t-1} + g_L \Delta p_t \quad (22)$$

$$\Delta p_t = p_t - p_{t-1} \quad (23)$$

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,

$$f_t^j = f_{t-1}^j + g_j(r_t - f_{t-1}^j). \quad (24)$$

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, momentum, trend following technical trading, and Kalman filtering.<sup>8</sup> 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

---

<sup>8</sup>See 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.

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.

The second forecasting rule is based on a classic fundamental strategy. It is based on the log linear approximation for returns,<sup>9</sup>

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

$$\rho = \frac{1}{1 + e^{-(\overline{p}-\overline{d})}} \quad k = -\log(\rho) - (1 - \rho) \log\left(\frac{1}{\rho} - 1\right). \quad (26)$$

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,  $pd_t$ .

$$f_t^j = E_{t,j}(\Delta p_{t+1}) = \overline{\Delta p_t} + \beta_t^j(pd_t - \overline{pd_t}). \quad (27)$$

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

$$E_{t,j}(p_{t+1}) = f_t^j + p_t \quad (28)$$

$$E_{t,j}(d_{t+1}) = d_t \quad (29)$$

$$E_{t,j}r_{t+1} \approx k + \rho E_{t,j}p_{t+1} + (1 - \rho)E_{t,j}d_{t+1} - p_t. \quad (30)$$

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

$$f_t^j = \bar{r}_t + \sum_{i=0}^{M_{AR}-1} \beta_{t,i}^j(r_{t-i} - \bar{r}_t) \quad (31)$$

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 range arbitrage strategy.”<sup>10</sup>

The previous two rules will be estimated each period using recursive least squares. There are many examples of this for financial market learning.<sup>11</sup> The key difference is

<sup>9</sup> See Campbell, Lo & MacKinlay (1996) for a textbook exposition.

<sup>10</sup> 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).

<sup>11</sup> See Evans & Honkapohja (2001) for many examples, and also very extensive descriptions of recursive least squares

that this model will stress heterogeneity in the learning algorithms with wealth shifting across many different rules, each using a different gain parameter in its online updating.<sup>12</sup>

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.

## 2.5 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,

$$\begin{aligned}\beta_{t+1}^j &= \beta_t^j + \frac{g_j}{\bar{\sigma}_{pd,t}^2} (pd_{t-1} - \bar{p}d_{t-1}) u_{t,j} \\ u_{t,j} &= ((p_t - p_{t-1}) - f_{j,t-1})\end{aligned}\tag{32}$$

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

For the lagged return regression the update follows,

$$\begin{aligned}\beta_{t+1,i}^j &= \beta_{t,i}^j + \frac{g_j}{\bar{\sigma}_{r,t}^2} (r_{t-i} - \bar{r}_{t-i}) u_{t,j}, \\ u_{t,j} &= (r_t - f_t^j)\end{aligned}\tag{33}$$

where  $g_j$  is again the critical gain parameter, and it varies across forecast rules.<sup>13</sup>

---

learning methods.

<sup>12</sup>Another recent model stressing heterogeneity in an OLS learning environment is Georges (2008) in which OLS learning rules are updated asynchronously.

<sup>13</sup>This 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,  $\beta_t^j$ .



## 2.6 Variance forecasts

The optimal portfolio choice demands a forecast of the conditional variance as well as the conditional mean.<sup>14</sup> The variance forecasts will be generated from adaptive expectations as in,

$$\hat{\sigma}_{t+1,j}^2 = \hat{\sigma}_{t,j}^2 + g_{j,\sigma}(e_{t,j}^2 - \hat{\sigma}_{t,j}^2) \quad (34)$$

$$e_{t,j}^2 = (r_t - f_{t-1}^j)^2, \quad (35)$$

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.<sup>15</sup> 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.<sup>16</sup> 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 0.5. 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.

## 2.7 Market clearing

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

$$1 = \sum_{i=1}^I Z_{t,i}(P_t). \quad (36)$$

---

<sup>14</sup> Several other papers have explored the dynamics of risk and return forecasting. This includes Branch & Evans (2011) and Gaunersdorfer (2000). In LeBaron (2001) 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.

<sup>15</sup> See Nelson (1992) for early work on this topic.

<sup>16</sup> See Bollerslev, Engle & Nelson (1995) or Andersen, Bollerslev, Christoffersen & Diebold (2006) for surveys of the large literature on volatility modeling.

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

$$P_t Z_{t,i} = (1 - \lambda) \alpha_{t,i}(P_t) W_{t,i}, \quad (37)$$

$$Z_{t,i}(P_t) = (1 - \lambda) \alpha_{t,i}(P_t) \frac{(P_t + D_t) S_{t-1,i} + B_{t-1,i}}{P_t}. \quad (38)$$

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  $\alpha_{t,i}$ .<sup>17</sup> 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.<sup>18</sup>

## 2.8 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. Half lives can be thought of using the simple exponential forecast mechanism with

$$f_{t+1}^j = (1 - g_j) f_t^j + g_j e_{t+1}. \quad (39)$$

This easily maps to the simple exponential forecast rule,

$$f_t = \sum_{k=1}^{\infty} (1 - g_j)^k e_{t-k}. \quad (40)$$

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

$$\frac{1}{2} = (1 - g_j)^{m_h}, \quad (41)$$

<sup>17</sup> A binary search is used to find the market clearing price using starting information from  $P_{t-1}$ .

<sup>18</sup> 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.

or

$$g_j = 1 - 2^{-1/m_h}. \quad (42)$$

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 10 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.

## 2.9 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,

$$\hat{U}_{t,j} = \hat{U}_{t-1,j} + g_u^i (U_{t,j} - \hat{U}_{t-1,j}), \quad (43)$$

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

$$U_{t,j} = \frac{1}{1-\gamma} (1 + R_{t,j}^p)^{(1-\gamma)} \quad (44)$$

with  $R_{t,j}^p$  the portfolio holdings of rule  $j$  at time  $t$ .

The design of the model allows for varying gain levels across agents through different values of  $g_u^i$ . However, this additional level of complexity complicates the model, so gain levels are assumed constant across agents in terms of utility estimation. They are set to the central gain level in the middle of the discrete distribution of gain levels, at 16 years. A range of fixed values from 10 through 50 would generate similar results.

The final component for 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 features

#### 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 2015, 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.<sup>19</sup> For the period from 1871 to 1925 the data set constructed by Shiller is used.<sup>20</sup> 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.<sup>21</sup> For most results the range of prices, returns and volatility will cover 1886-2015 with a splice into the standard CRSP series in 1926.

The real dividend series in the Shiller data set is used for calibration. Figure 9 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.<sup>22</sup>

---

<sup>19</sup>This method is very commonly used with intraday data, but was originally used in French, Schwert & Stambaugh (1987) at the daily frequency.

<sup>20</sup>This is available at his website. It is the main input for Shiller (2000), and many papers.

<sup>21</sup>Schwert (1990) gives extensive details on the construction of this series which is available at his website.

<sup>22</sup>The 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.

### 3.2 Basic stylized facts: model and data

Table 1 shows the parameters used for all model simulations. Also, table 2 displays annualized half lives for the range of 10 gains used in the learning model. These are estimated to be uniform in log2 space. The homogeneous agents are assumed to use the discrete gain level 7 for fitness/utility estimation,  $g_i^u$ , in equation 43 which corresponds to a half life of about 16.<sup>23</sup>

Figure 3 summarizes the basic empirical features that the model produces both for short and longer term returns information.<sup>24</sup> The three panels display the price dividend ratio, weekly returns, and trading volume respectively. All visually demonstrate features common to most financial data. The price dividend ratio moves through a wide range of values, taking big dramatic swings. Returns appear to demonstrate both large values, and pockets of clustered volatility. These pockets of high volatility are often accompanied by large trading volume.

Figure 4 looks at the weekly returns series in more detail. The top panel repeats the return time series. In the second panel deviations from normality are displayed for the unconditional return distribution. The bottom panel shows the autocorrelations for both returns and absolute returns. The returns are nearly uncorrelated, and the absolute returns are positively correlated with the correlation persisting out many weeks. Both these features are well known from most financial time series.

Table 3 summarizes return features at the annual horizon, and compares with returns from a long run U.S. data set. The simulations use 50,000 weeks of simulated data, or roughly 1,000 years. The model matches the mean and standard deviation well. There is some stronger evidence for mean reversion in the annual simulation data as well as more kurtosis and skewness in the longer horizon data. Table 4 compares the dynamics of the price dividend ratios in both the data, and the model.<sup>25</sup> Here the alignment both in terms of mean, variability, and persistence is very close.<sup>26</sup> There is some evidence for less persistence in the simulation which may be related to the faster speed of return mean reversion demonstrated in the previous table.

Table 5 gives a brief comparison of higher frequency returns in the model and in the data. In general, they are in very close alignment indicating success in the calibration exercise at many horizons. The model and data show fat tailed return distributions with very little autocorrelation. The fat tailed distributions are displayed both by excess kurtosis and using a quantile ratio statistic looking at the 99 versus 95 quantile range ratio.

---

<sup>23</sup>Results are generally not sensitive to a range of longer half lives from about 10-100.

<sup>24</sup>The results are not shown here, but this model will converge to a tractable rational expectations equilibrium if the gain distribution is restricted to only long half lives,  $> 50$  years. This is an important benchmark test for the model, showing that the set of rules is capable of learning the equilibrium, and the researcher has not prebiased the results against this possibility.

<sup>25</sup>For the data the price dividend ratios are adjusted for share repurchases.

<sup>26</sup>The properties of the price/dividend ratio in the model are closely tied to the consumption/wealth parameter,  $\lambda$ .

This value gets larger as the extreme returns move out. It would be 1.3 for a Gaussian distribution. One unusual feature for the simulations that is consistent with table 3 is that the longer horizon returns still look a little unusual. They are not converging to normality as quickly as the data. The simulated monthly return still shows a kurtosis of 35 which is large relative to the estimated value of 10 in the data.<sup>27</sup> In general, the model shows a very good alignment with the general features we know from actual data.

## 4 Intertemporal optimality

Replicating many empirical features is easy for simple single period preferences. However, the myopic preferences that are used open up many questions about the usefulness of these results. While behavior is, by design, consistent with optimizing behavior and portfolio choice over single periods, it may deviate from this over longer horizons. The following sections push the basic model further to see how well it performs when considered from the perspective of an intertemporal optimizing agent.

### 4.1 Simple consumption

Data generated thus far has shown reasonable market dynamics. The consumer is following simple rules of thumb designed to replicate an intertemporal optimizing agent, but it is not clear how well it is doing at his. Could the agents be behaving far off the intertemporal optimum, calling into question their behavior in the model?

This section analyzes the baseline case where consumption is a constant fraction of wealth,

$$C_{t,i} = \lambda W_{t,i}.$$

Even though agents follow this basic “rule-of-thumb” for their consumption decision, it will be tested against standard first order conditions,

$$1/\beta = E\left(\frac{u'(c_{t+1})}{u'(c_t)}(1 + r_f)\right) \quad \text{FOC1} \quad (45)$$

$$0 = E\left(\frac{u'(c_{t+1})}{u'(c_t)}(r_t - r_f)\right). \quad \text{FOC2} \quad (46)$$

Rolling samples of both first order conditions are estimated with equation 43 that replicates the exponential moving average that the agents are using for utility for the Euler equation estimates. A plot of the right side of the first condition,(FOC1), is given in figure, 6. This figure plots both the mean across all  $I$  agents and a two standard deviation

---

<sup>27</sup>This may be related to the length of the simulation relative to the actual data. With nearly 1000 years of data the possibility for more very extreme moves, or periods like the great depression will be more likely.

band estimated from the contemporaneous, and changing, cross sectional standard deviation. The value does move around through time, and can move above and below 1.<sup>28</sup> Since the risk free rate in this model is constant the dynamics are driven by consumption dynamics in the model. Large increases must be driven by large falls in consumption which drops suddenly, and then continue falling.

The sample values for this model are in table 6. The first line of the table reports the first order condition for the risk free rate (FOC1). The mean value over time is slightly larger than 1, with a large amount of variability over time. The cross sectional variability is slightly less, as reported in the table with a mean standard deviation of 0.0008. All these features are visible in figure 6. The mean value pegs the free parameter of the time rate of discounting for the consumer's preferences. It gives an annualized value of  $\beta = (1/1.0009)^{52} = 0.96$  which is reasonable.

Figure 7 shows the time series for the excess return first order condition (FOC2) along with the cross sectional dispersion represented by 2 standard deviation bands as in the previous figure. FOC2 stays very close to zero as it should with relatively tight bands, except during periods after extreme moves of the market. During these periods, dispersion appears to be wide and the value of the FOC drops negative. These are likely periods where consumption (wealth) is low, and conditional returns are high. This is obviously reverse of what we should expect in standard consumption based asset pricing models. Also, the possibility that this part of the model (choice between risky and risk free assets) works well except for a few extreme periods would appear to replicate features seen in the data.<sup>29</sup>

Table 6 reports a summary of the figure in line 2. The mean is near zero, and generally displays only a few extreme deviations. The variability does not suggest much of a strong deviation from the optimality conditions in this case. However, the right column of this table presents the std. error for the time series mean estimate.<sup>30</sup> Given the 50,000 week sample, the standard error is very small, indicating the estimated FOC is significantly different from its zero target.

This simple model of consumption choice would appear to be relatively effective at replicating a reasonable long term intertemporal decision maker. Unfortunately, there are some problems in terms of lining up with reality. Agents are consuming a constant fraction of wealth, and wealth is moving directly with asset prices which are extremely volatile. This volatility naturally drives variability in consumption, creating a highly variable risk measure for the consumption based asset pricing conditions. Table 7 demonstrates how severe this problem is. The first row displays the growth rate and standard deviation of the growth rate for quarterly (13 weeks) consumption. The U.S. data for real

<sup>28</sup>Values below one would be inconsistent with a discount rate,  $\beta < 1$ .

<sup>29</sup>Evidence for this can be seen in Lettau & Ludvigson (2009).

<sup>30</sup>Standard errors are adjusted with a Newey/West procedure with 25 lags since the series is highly persistent.

percapita consumption from 1947-2015 is displayed on the bottom row.<sup>31</sup> The growth rates are similar at about two percent, but the variability in the simulated consumption growth is almost a factor of 10 different from the actual data. In the model consumption responds to the highly variable stock market wealth, giving a variable consumption risk measure, but counter factual in terms of real data.

## 4.2 Smoothed consumption

The variable consumption level is a challenge for the simple ad hoc rule for agent behavior. One possible solution is to try a consumption rule which responds to longer term changes in wealth. This naturally makes sense given that the model is simulating at a weekly frequency. Now replace wealth in the consumption rule with an adaptive filter as was done for utility,

$$\bar{W}_{t,i} = \bar{W}_{t-1,i} + g_u(W_{t,i} - \bar{W}_{t-1,i}). \quad (47)$$

Consumption will be the same constant fraction of wealth,  $\lambda$ , as before, but wealth is replaced with the long term average,  $\bar{W}_{t,i}$ . Row 2 of table 7 shows that the new consumption rule now better matches both the growth and variability of the U.S. aggregate series.

Figure 8 repeats the previous time series of the estimated first order condition for the risk free rate with this new consumption rule. The figure again displays the mean and bands set to twice the standard deviation from the cross section over all agents. The overall value for the first order condition now appears to be comfortably less than 1, although the bands indicate that some small fraction of consumers might occasionally be above 1. Being less than one for this value is problematic as it implies a value of  $\beta$  greater than one, or reverse time discounting. This problem could be reduced in this model by simply increasing the real interest rate, but that is not really fair since the interest rate is set to a reasonable low value, one percent, to line up with the data.

A summary of this figure is given in table 6, line 3. The mean of the weekly series is a small level below 1, but still has relatively large cross sectional and time series variability. The overall sample mean yields an annualized discount of  $\beta = 1.04$  which is out of range for any reasonable dynamic optimizing consumer. It is important to note that this tension of demanding either a higher real interest rate, or strangely counter factual preference discount is exactly what is seen in financial data as the risk free rate puzzle.

Figure 9 repeats a plot for the excess return FOC2 now using the slow moving consumption series. The figure is again a combination of the mean and cross sectional variability, but in this case, the cross-sectional variability is so small that it is difficult to detect.<sup>32</sup> This statistic which should be zero if agents were at their intertemporal FOC

<sup>31</sup>Source is FRED at the Federal Reserve Bank of St. Louis.

<sup>32</sup> This is an interesting feature in itself, showing that there appears to be little heterogeneity in the cross section for this measure.



appears to be positive most of the time. It only takes some occasional dips below zero which might be driven by extreme moves in the market. For this first order condition a positive value suggests that at the simulated return levels in the market the intertemporally optimizing agent would prefer to shift their portfolio from the risk free to the risky asset. They perceive the risk/return tradeoff shown in the market as being favorable at the going rates. This is exactly the problem of the equity premium puzzle from U.S. data.

The figure is again summarized by the last row in table 6. The sample mean is again very small with an estimate of  $14 \times 10^{-4}$ . The estimated standard error for the time series mean is again very small at  $1.93 \times 10^{-5}$  indicating the mean estimate is significantly different from zero.

### 4.3 Quarterly results

This section again uses the weekly frequency simulations, but they are aggregated up to the quarterly frequency for comparison to post war U.S. quarterly data. The myopic model was only designed to be optimizing over the one week horizon, so it is stretch to test what will happen at the quarterly frequency. The last 25,000 weeks are now converted into nearly 500 years of quarterly consumption and return data. The former is only stored for aggregate consumption, and not the individual agents as was done before. Model simulations only use the slow consumption adjustment from equation 47.

Figure 10 shows the (FOC1) value for the U.S from 1947-2015. It is smoothed with an 8 quarter moving average. The value is less than one most of the time which is consistent with the risk free rate puzzle. It only appears to cross through one during major recessions when consumption is falling. Figure 11 repeats this plot for the simulation data. A similar pattern shows up with the values mostly below one, with only a few crossings from below.

The first two lines in table 8 show the estimated values for risk free first order condition. Both the data and the simulation are close with values of 0.987 and 0.988 respectively. They both imply incorrect discount rates of  $\beta = 1.05$ . The standard errors show that both means are significantly less than 1 which is problematic for the standard intertemporal optimizing agent in either case.

Figure 12 repeats the U.S. results for the excess return FOC2. This is again an 8 quarterly rolling mean. Values are above zero, but there are a few periods where they turn negative during recessions. The results in figure 13 show similar features in the simulation data over a much longer time horizon. The estimated values are generally positive, but with a few deviations into negative values.

Summary results are presented in table 8. The alignment between the model and the data in their deviations from the first order conditions is again very close. For the U.S. the mean value is 0.017 and for the simulation it is 0.015. Even the standard errors are close

with estimates of 0.080 and 0.114 respectively. Standard errors on the estimated means show them to be significantly larger than zero. This positive sign is an equity premium related phenomenon. In both cases optimizing agents would increase their utility by shifting portfolios toward the risky asset. In other words, they do not perceive a risk in terms of how it moves with the simple quarterly marginal rate of substitution that would justify its high return in either the model or the data.

## 5 Conclusions

This paper has looked at a simple multi-agent learning model with heterogeneous gain learning, risky dividend and labor income. Although its structure aligns with much of the myopic portfolio choice mechanisms used in a large part of the agent-based literature, it is analyzed from the perspective of an intertemporal optimizing agent who happens to follow these simple rules of thumb for behavior. The rules of thumb are crafted to have strong connections to simplified versions of full optimizing behavior as inspired by Campbell & Viceira (2002). The model replicates most of the high frequency properties of asset returns at higher frequencies (weekly), and does well at longer horizons as well.

The model is put to the test of replicating standard Euler equations that are well known in the macro/finance literature. On this dimension the results are mixed. With a simple rule yielding consumption as a constant fraction of wealth, the results look reasonable in terms of Euler equations. This is an interesting test of whether the myopic rules are a good approximation to a more complex decision making process. However, the market is effective at generating extensive asset price volatility. Given that agents hold a large fraction of wealth in equity, their weekly levels of consumption are extremely volatile. Even at a quarterly frequency consumption growth rates are nearly 10 times levels observed in U.S. data.

The model is augmented by forcing consumption to be a constant fraction of a longer range average of wealth. This is an intuitive mechanism to slow down the reaction of consumption to changes in stock market wealth. Doing this brings consumption down to levels of variability close to U.S. data. However, now with calmer consumption, the Euler equations are violated. It is interesting that the violations are very similar to those seen in data. The rule of thumb mechanism appears to be making the same mistakes as the aggregate U.S. representative agent. Some of this is related to getting the basic moments correct on asset returns, and consumption growth, but there must be some deeper matching of covariants at work.

This work is only a first attempt to connect simpler agent-based models to the world of more traditional macro models. Further work is necessary. In particular, it will be important to quantify the magnitudes of the Euler equation deviations. Are there big,

economically significant, mistakes being made by the simple agents? If so, then the entire exercise may be called into question. It is also possible that the errors are economically small, and within the bounds of reasonable adjustment costs to consumption.<sup>33</sup> However, it is possible that the errors might also suggest slightly more realistic rules of thumb for the agents. They also may tell us which empirical puzzles cannot be solved in the scope of a simple heterogeneous learning model. It is possible that we may need to bring in both the dynamics of learning along with more specific institutional details that impact consumption choices at the individual level to give a clear model of financial and macroeconomic dynamics.

---

<sup>33</sup>This is the message in Cochrane (1989) in evaluating deviations from rationality.

## 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.
- 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.
- 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. 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.
- 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. (2014), 'Heterogeneous expectations in asset pricing: Empirical evidence from the s&p 500', *Journal of Economic Behavior and Organization* **105**, 1–16.
- Cochrane, J. (1989), 'The sensitivity of tests of the intertemporal allocation of consumption to near-rational alternatives', *American Economic Review* **79**, 319–337.
- Evans, G. W. & Honkapohja, S. (2001), *Learning and Expectations in Macroeconomics*, Princeton University Press.
- 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.
- 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.
- Kouwenberg, R. & Zwinkels, R. C. J. (2010), Chasing trends in the U.S. housing market, Technical report, Erasmus Univeristy, Rotterdam, The Netherlands.
- LeBaron, B. (2001), '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. (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., 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. (2009), 'Euler equation errors', *Review of Economic Dynamics* **12**, 255–283.
- 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.
- 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. (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.
- Phelps, E. S. (1967), 'Phillips curves, expectations of inflation and optimal unemployment over time', *Economica* **34**(135), 254–281.
- Piketty, T. & Saez, E. (2003), 'Income inequality in the United States, 1913-1998', *Quarterly Journal of Economics* **118**(1), 1–39.
- 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: *Parameter Definitions*

Parameter	Value
$d_g$	0.02
$\sigma_d$	0.07
$r_f$	0.01
$\gamma$	3.5
$\lambda$	0.0005
$I$	16000
$J$	4000
$g_u$	16 years
$L$	5 percent/year
$\sigma_\epsilon$	0.01
$M_{AR}$	3
$h_j$	[0.025, 0.10]

The annual standard deviation of dividend growth is set to the level estimated in post war U.S. data estimated from the implied dividend growth using the CRSP value weighted index.



Table 2: *Half life gain distribution in years*

	Half life
1	0.48
2	0.86
3	1.56
4	2.81
5	5.06
6	9.12
7	16.44
8	29.62
9	53.36
10	96.15

Half lives are distributed uniformly in  $\log_2$  space over the range of 26 to 5000 weeks.

Table 3: *Annual Excess Return Summary Statistics*

	Mean	Std.	Skewness	Kurtosis	ACF(1)
U.S. Data (1885-2015)	0.069	0.198	-0.19	2.87	-0.01
Simulation (1924 years)	0.081	0.217	-1.32	7.43	-0.21

Table 4: *Price/dividend ratios*

	Mean	Std.	ACF(1)
U.S. Data (1885-2015)	23.2	6.50	0.75
Simulation (1924 years)	22.6	4.17	0.52

Table 5: *Higher Frequency Excess Return Summary Statistics*

	Mean*100	Std.*100	Skewness	Kurtosis	ACF(1)	Q-ratio
U.S. Daily Data (1885-2015)	0.23	0.99	-0.16	18.9	0.05	1.84
U.S. Weekly Data (1885-2015)	0.12	2.34	-0.46	10.1	0.01	1.74
Simulation Weekly	0.15	2.94	-0.51	20.5	-0.03	1.83
U.S. Monthly Data (1885-2015)	0.53	5.09	0.12	10.1	0.09	1.65
Simulation Monthly	0.60	5.66	-2.51	35.9	0.05	1.99

Q-ratio is a robust measure of fat tailed shapes. It estimated the 99 percent quantile range divided by the 95 percent quantile range. For a Gaussian distribution this value is 1.3. Wider extreme moves in the data increase this ratio. This statistic is added because the existence of 4th moments may be in question for some of these series.

Table 6: *First order condition summaries*

	Mean	Std	Cross sec. std	Annual $\beta$	std(Mean)
FOC 1: Constant wealth/consumption	1.0009	0.0017	0.0008	0.96	6.90e-5
FOC 2: Constant wealth/consumption	-1.43e-4	10e-4	3.5e-4		4.25e-5
FOC 1: Slow consumption	0.9992	0.0006	0.0007	1.04	2.79e-5
FOC 2: Slow consumption	14e-4	6.65e-4	3.000e-6		2.69e-5

$$\text{FOC1: } z_t = \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (1 + r_f))$$

$$\text{FOC2: } z_t = \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (r_t - r_f))$$

Table 7: Quarterly consumption summaries

	Growth	Growth Std.
Wealth/consumption	0.0044	0.0581
Slow consumption	0.0046	0.0053
US per capita real C (1947-2015)	0.0052	0.0082

Table 8: Quarterly first order conditions

	Mean	Std	Std(mean)	Implied annual $\beta$
FOC1: U.S. 1947-2015	0.987	0.026	0.002	1.05
FOC1: Sim/slow consumption (480 years)	0.988	0.015	3.47e-4	1.05
FOC2: U.S. 1947-2015	0.017	0.080	0.005	
FOC2: Slow consumption (480 years)	0.015	0.114	0.002	

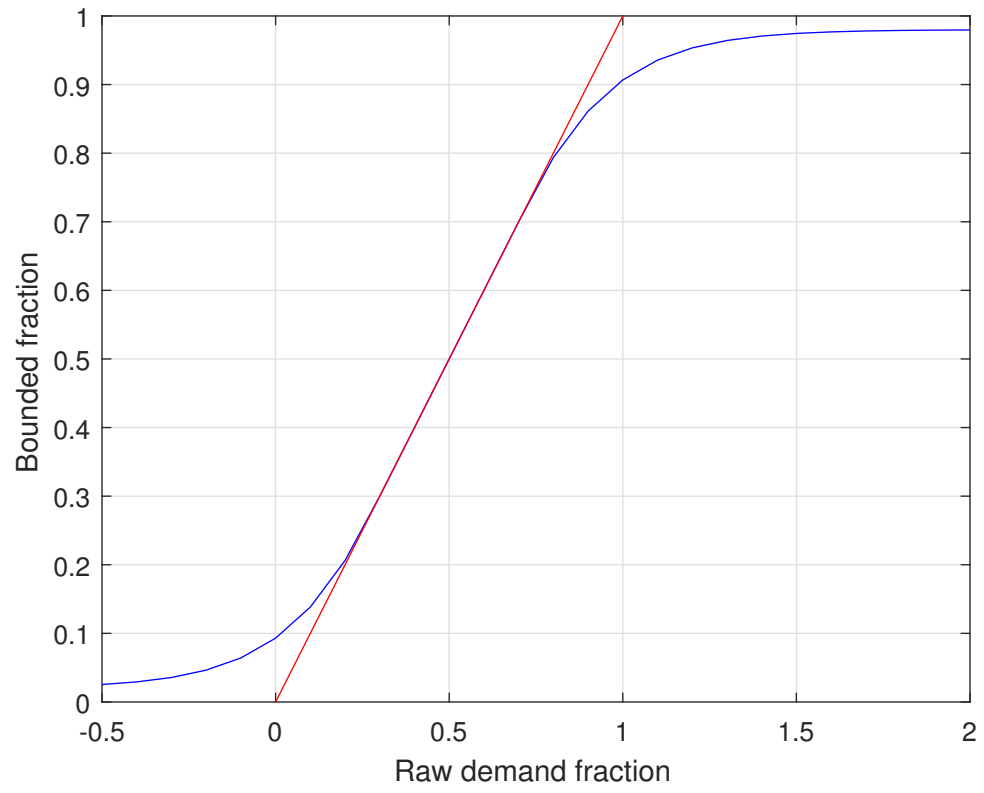


Figure 1: **Bounded share fractions**



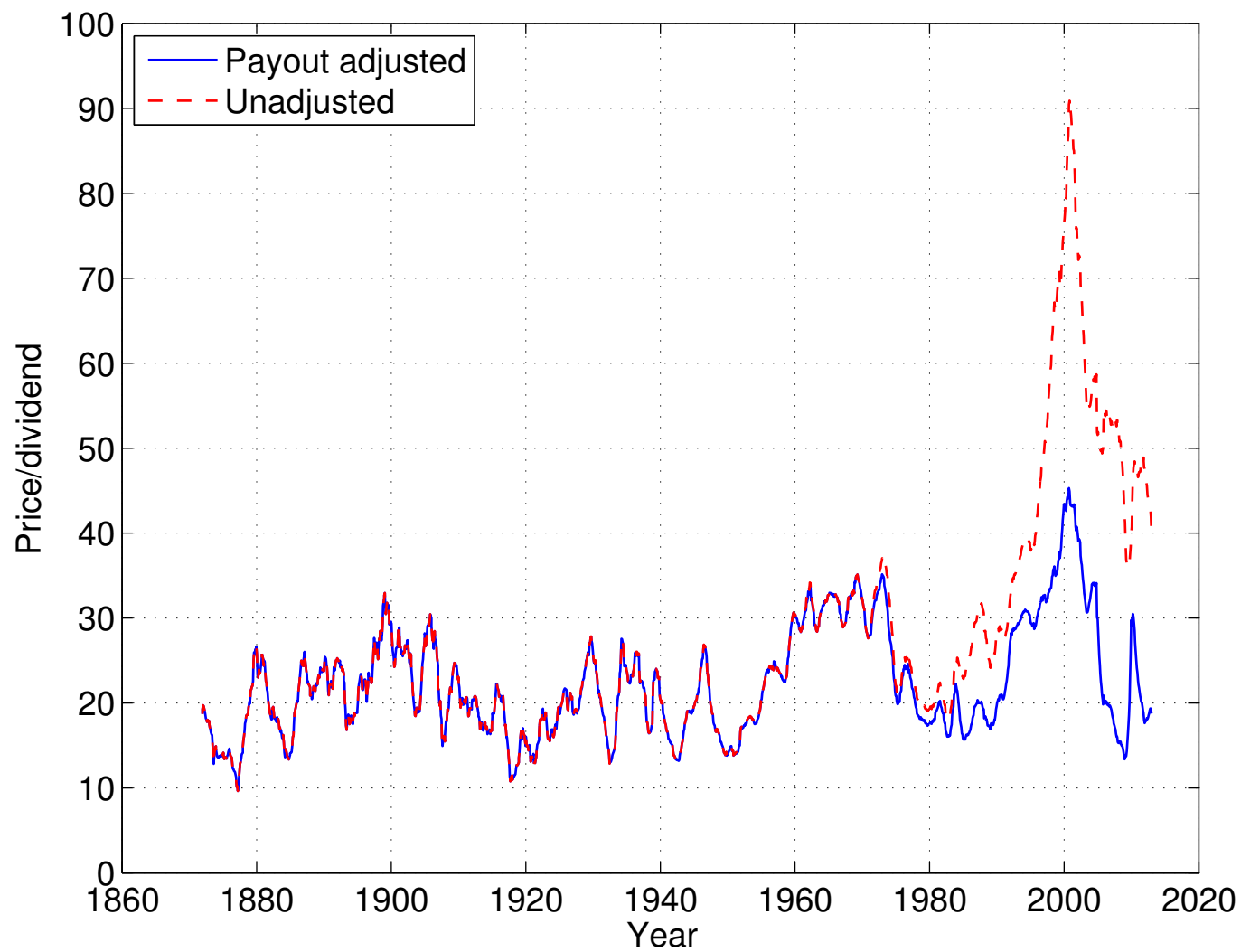


Figure 2: Price/dividend ratios

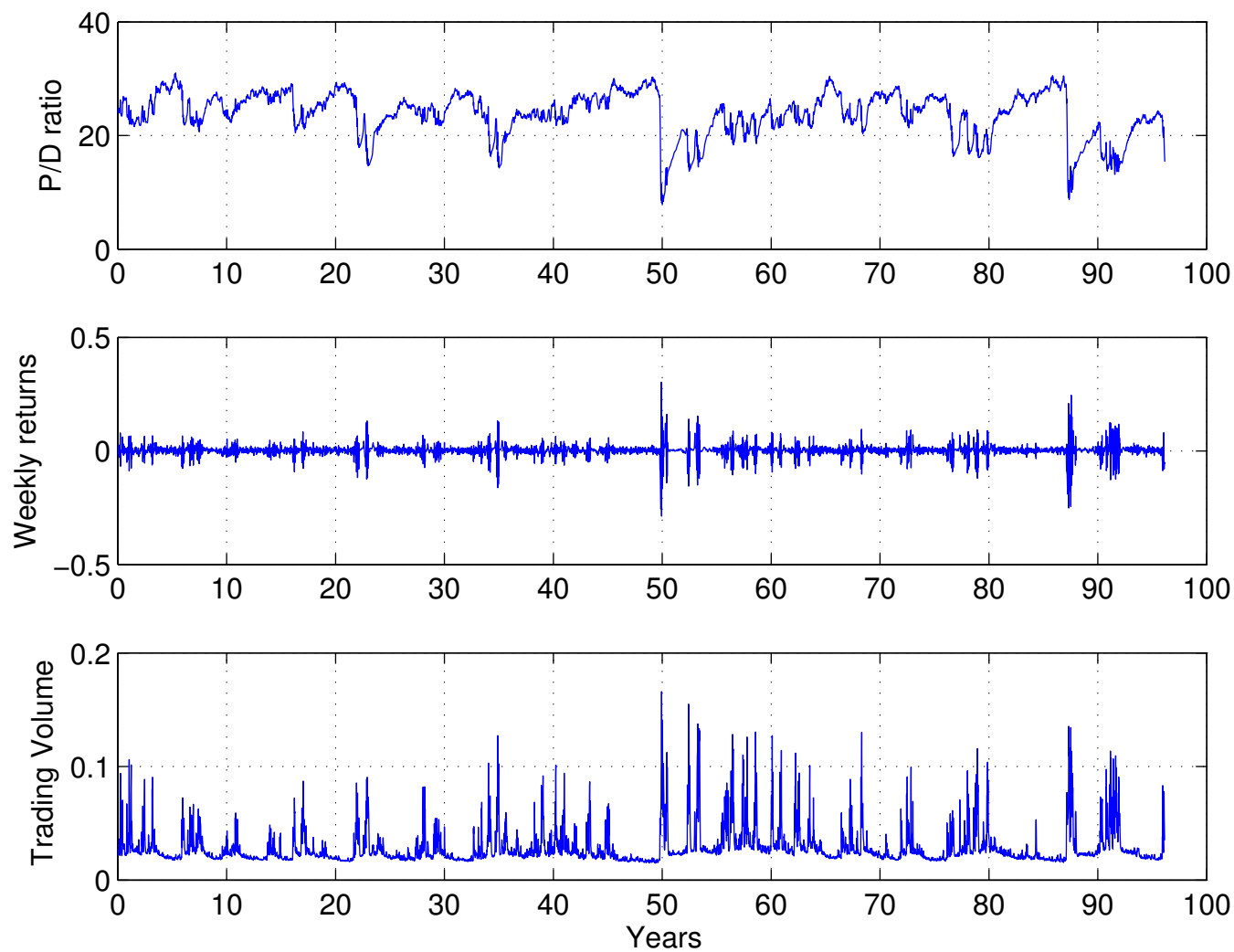


Figure 3: **Base simulation: Price/dividend, returns, volume**

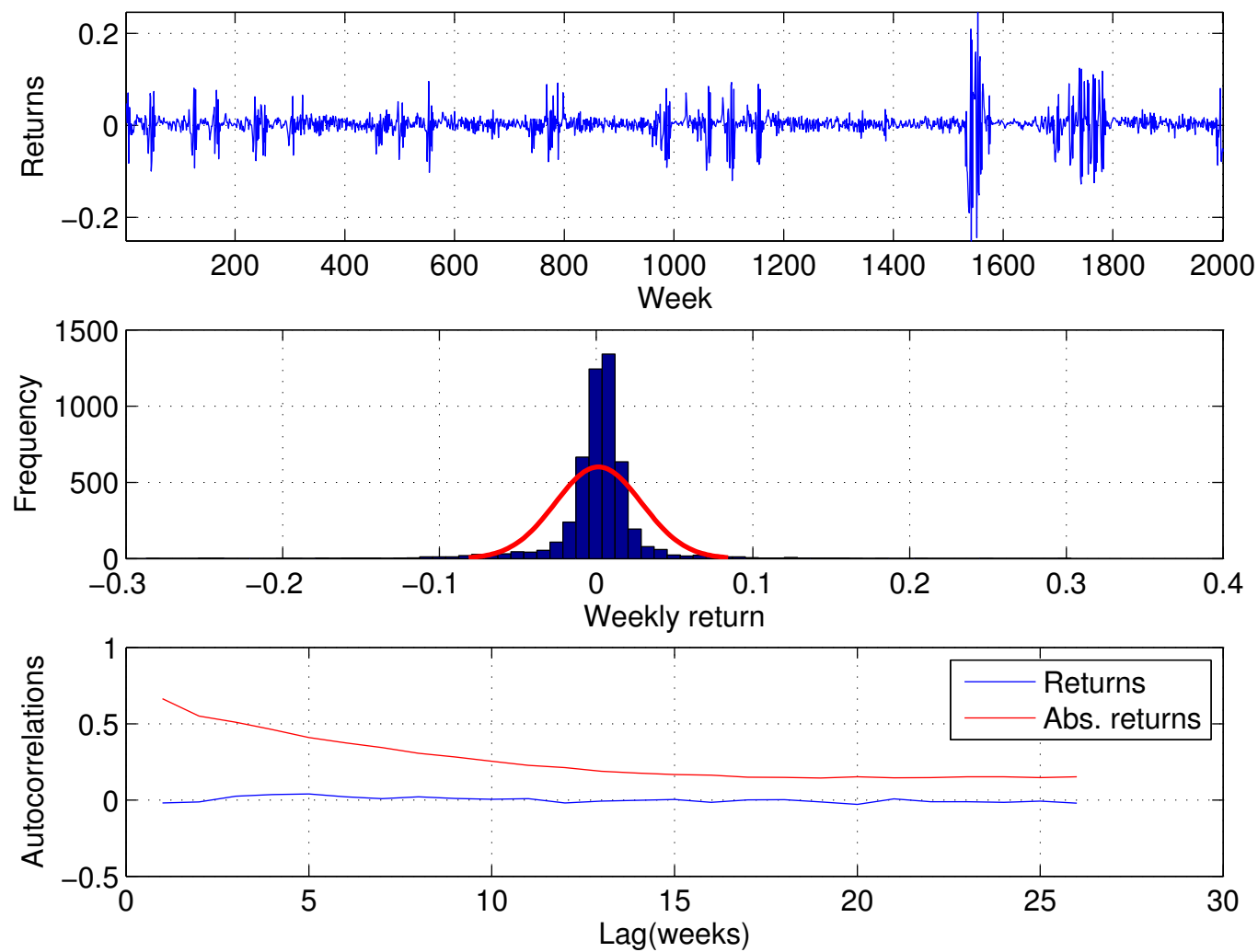


Figure 4: Weekly return features

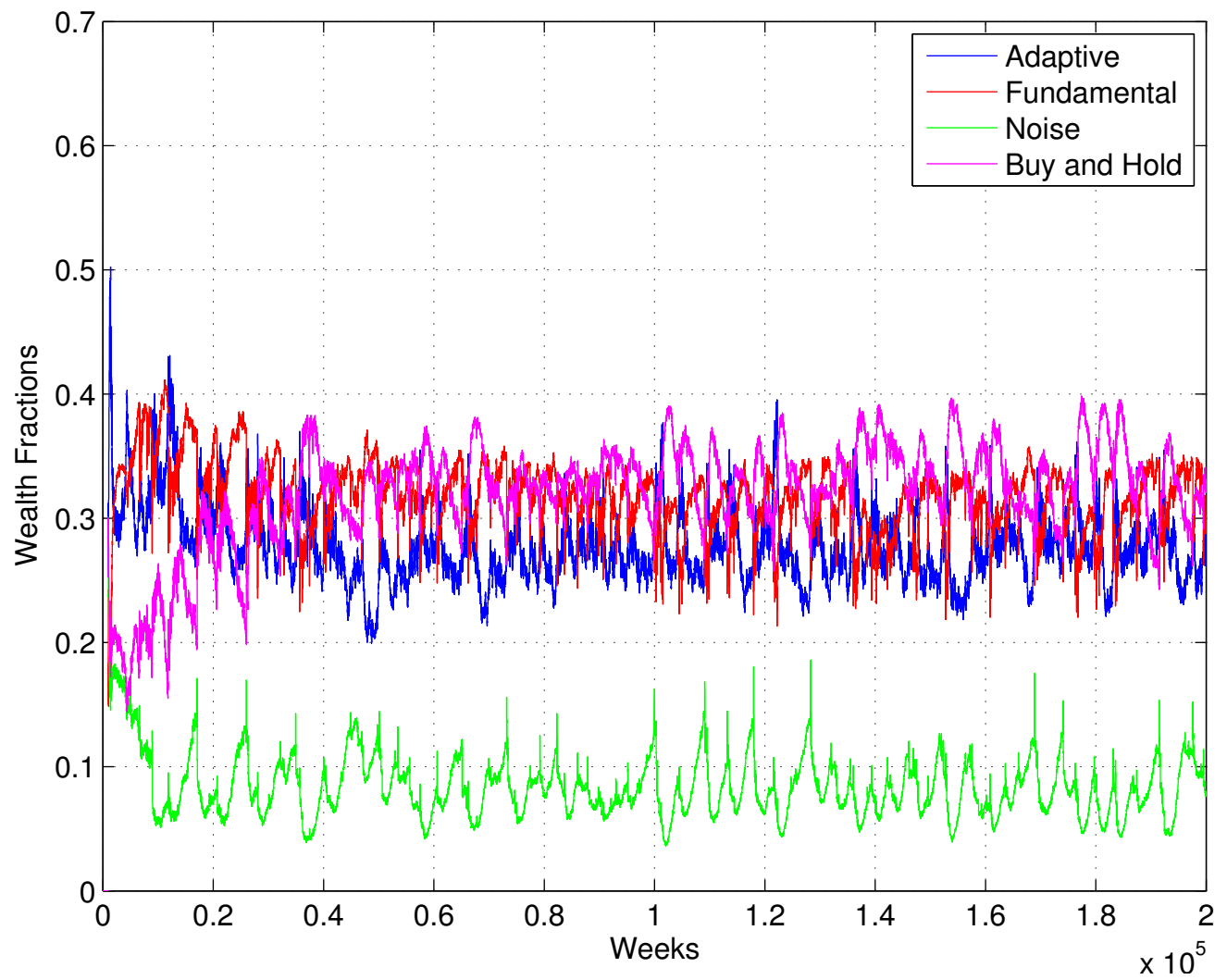


Figure 5: **Basic plot**

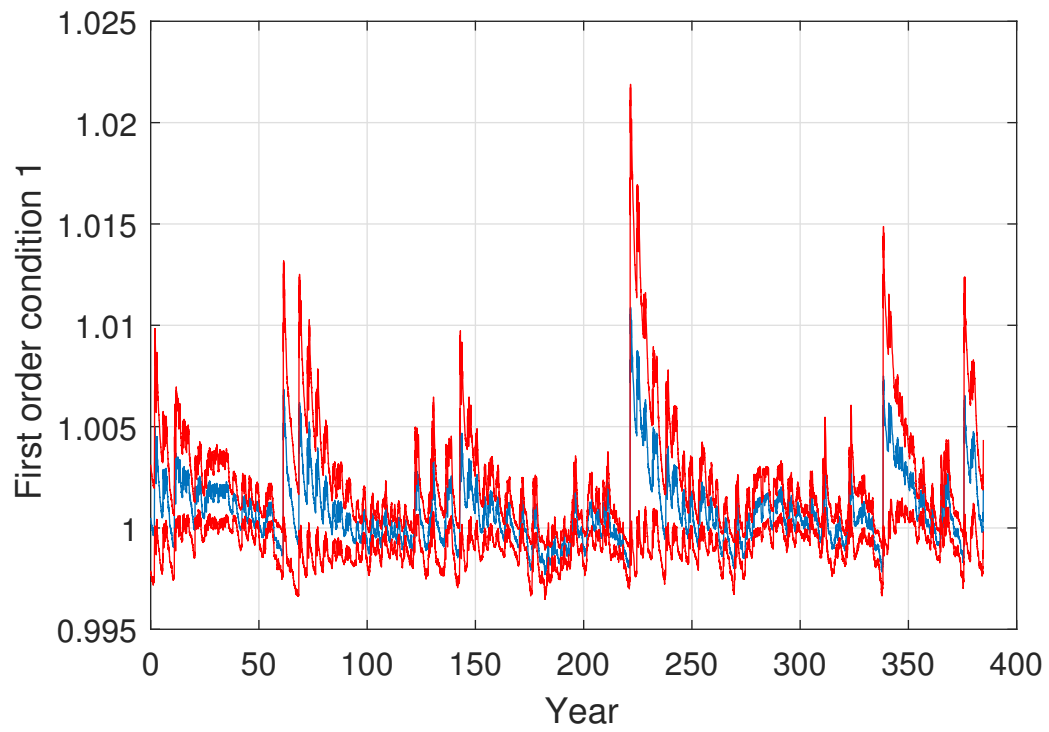


Figure 6: **Weekly simulation: First order condition 1**

$$z_t = \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (1 + r_f)$$

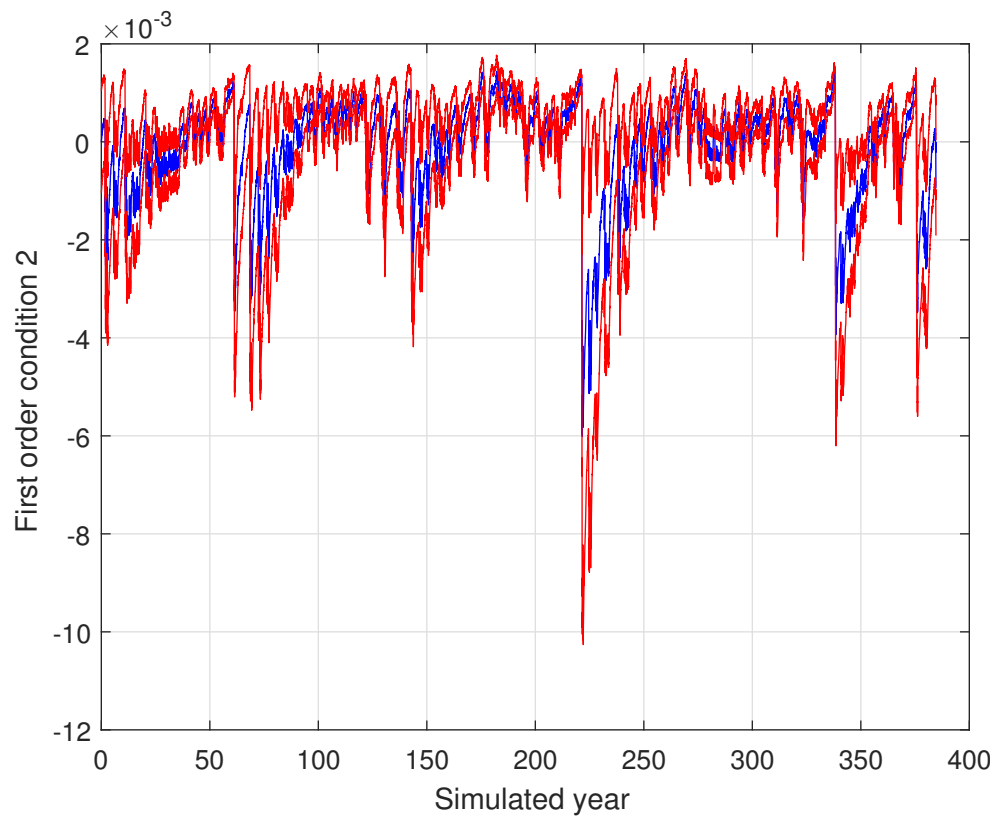


Figure 7: **Weekly simulation: First order condition 2**

$$z_t = \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (r_t - r_f)$$

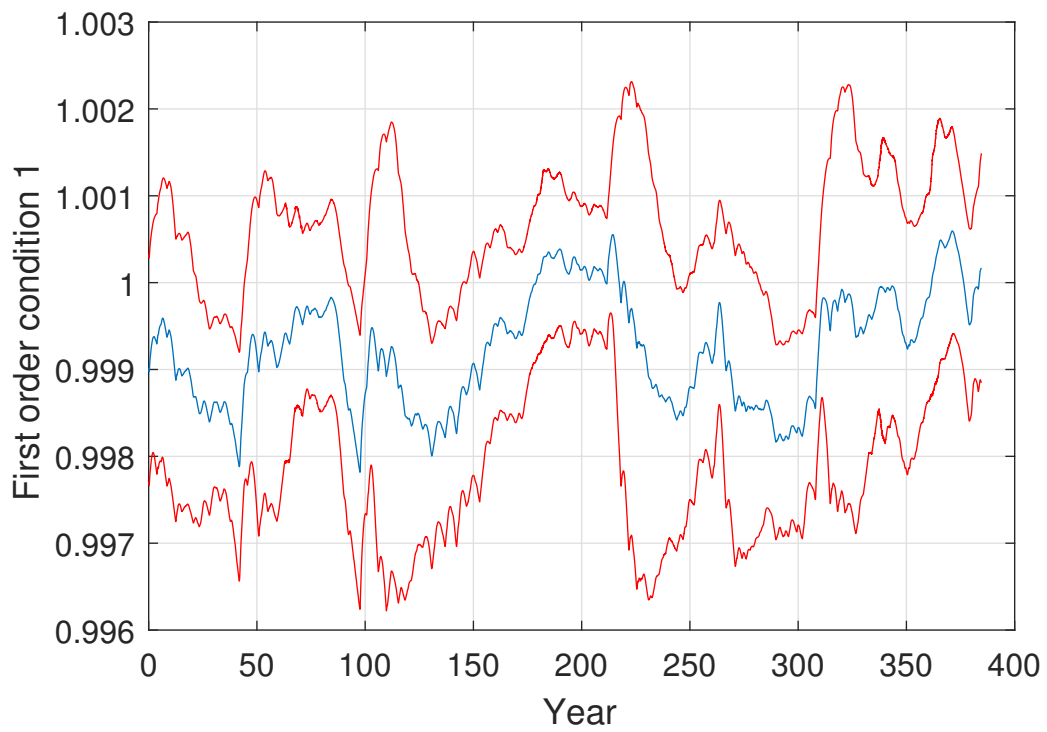


Figure 8: **Slow consumption: First order condition 1**

$$z_t = \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (1 + r_f)$$

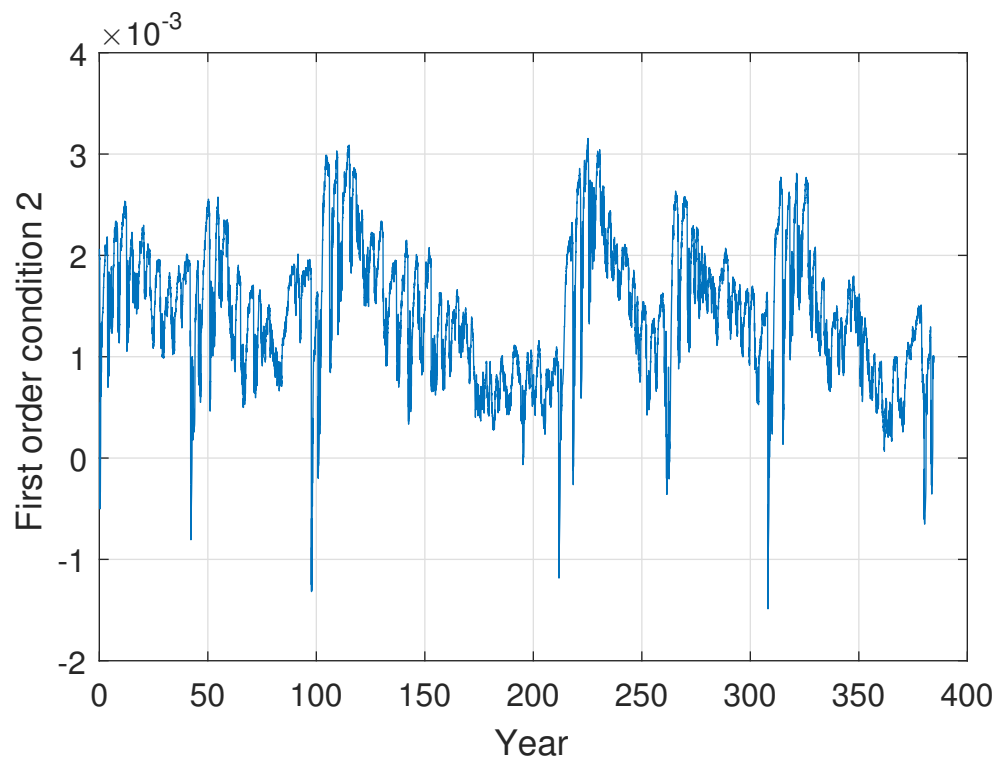


Figure 9: **Slow consumption: First order condition 2**

$$z_t = \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (r_t - r_f)$$



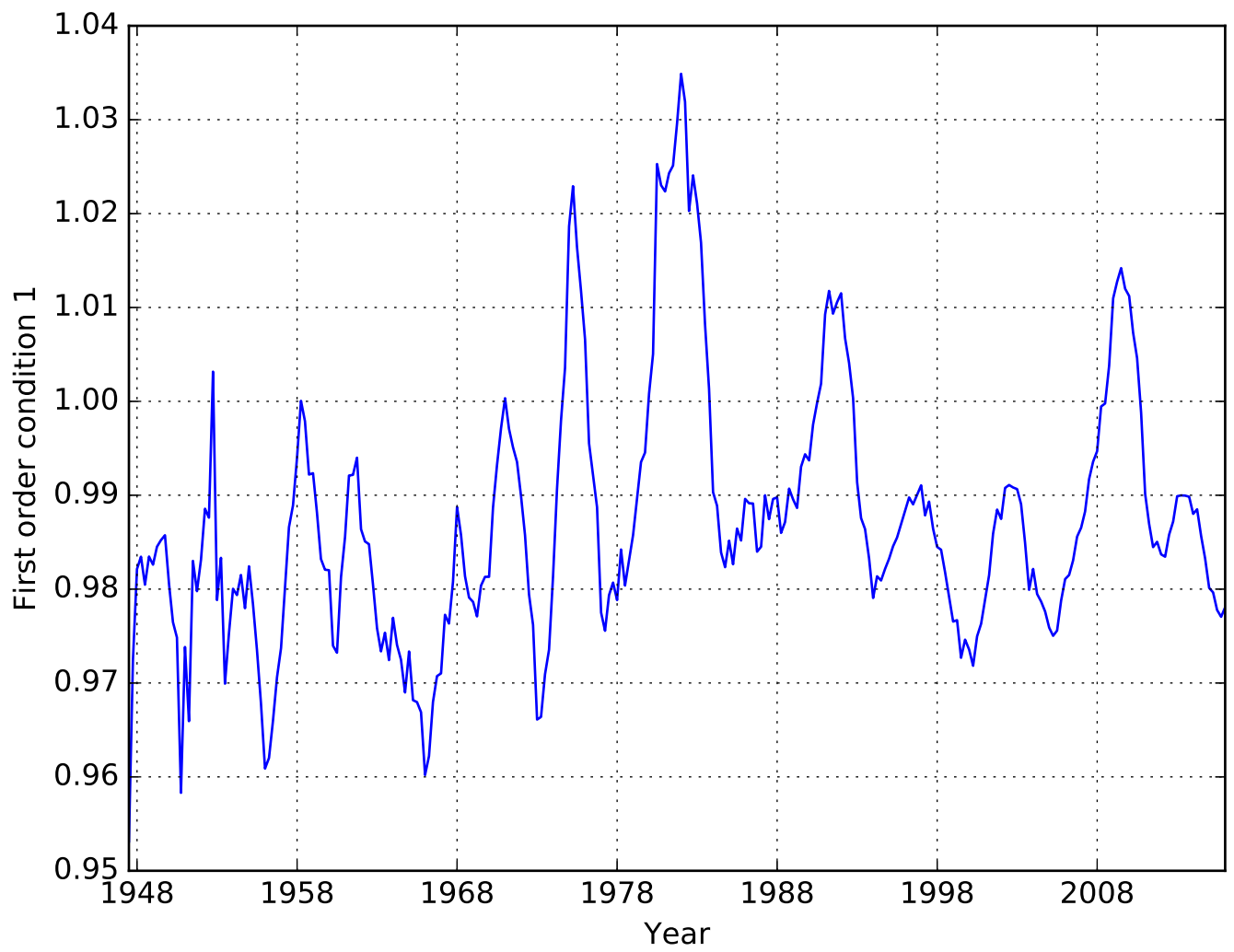


Figure 10: **FOC1: US Quarterly**

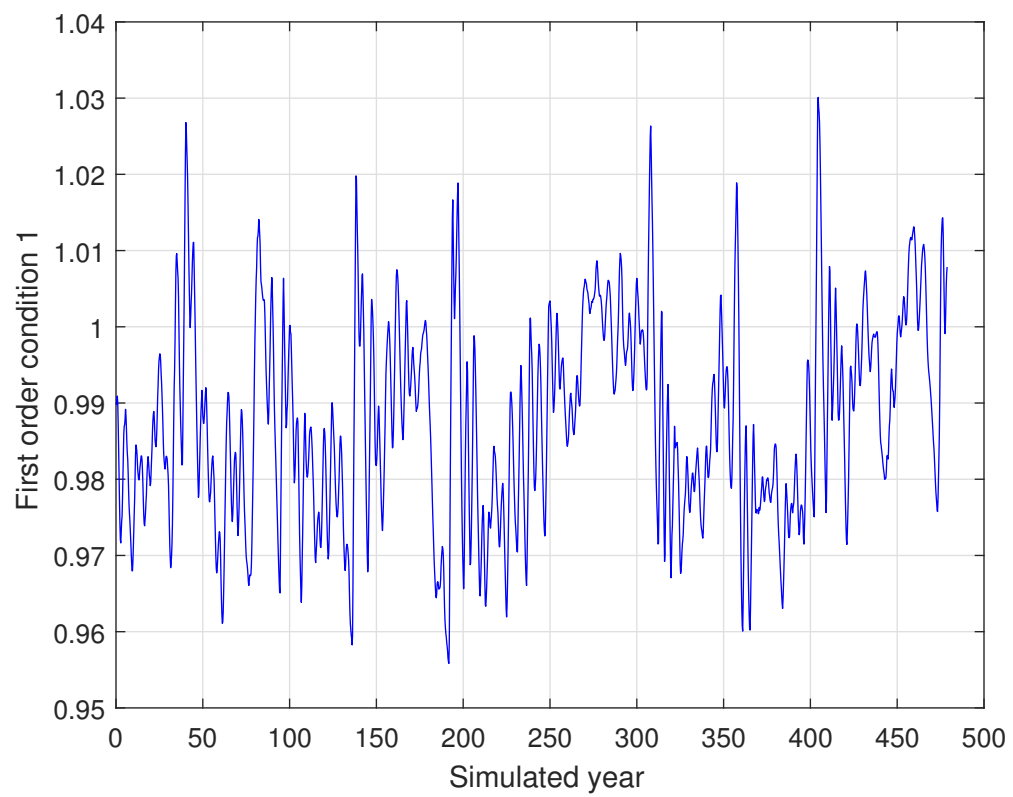


Figure 11: **FOC1: Sim quarterly**

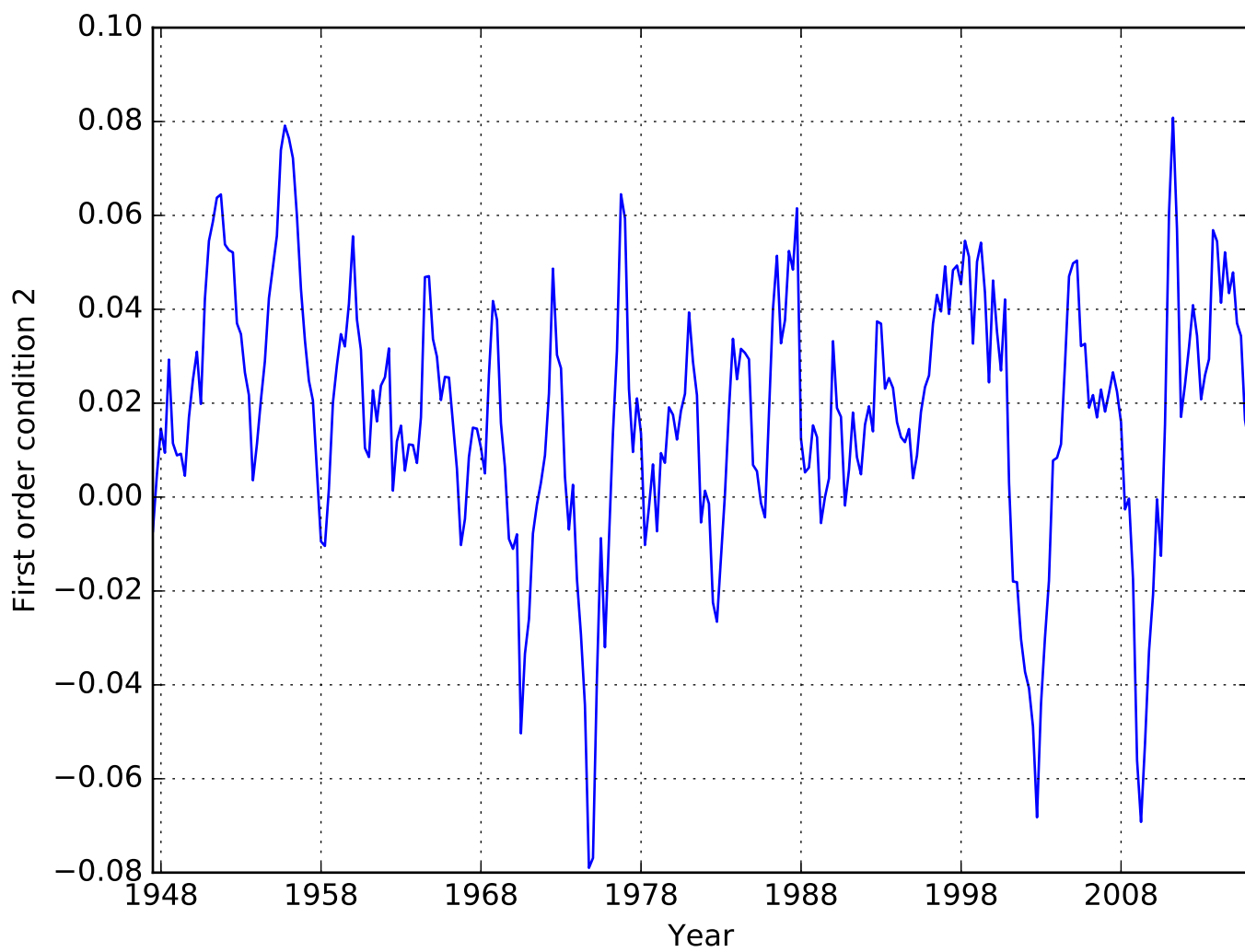


Figure 12: FOC2: US Quarterly

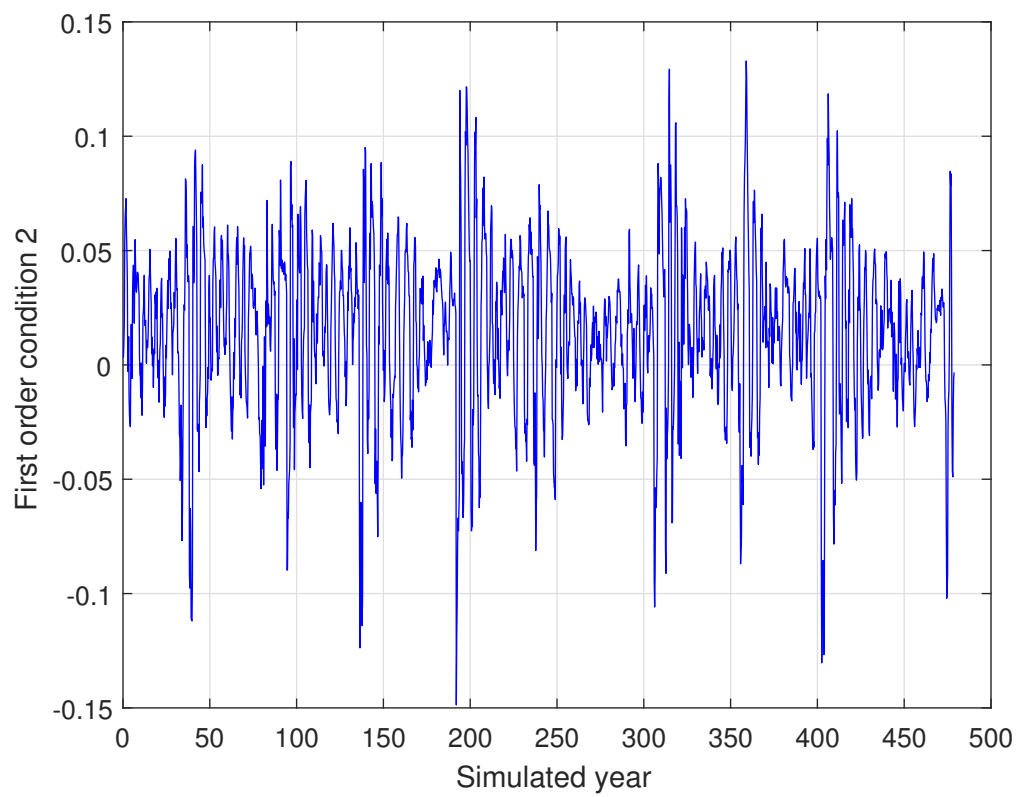


Figure 13: **FOC1: Sim quarterly**