Discussion of:

Heterogeneous Gain Learning and the Dynamics of Asset Prices

by *Blake LeBaron*

Kent Daniel[†]

[†]Columbia University, Graduate School of Business

Conference on Heterogeneous Expectations and Economic Stability
11 Feb, 2011

Outline

- Empirical Evidence on Market Predictability
- Model Setup
- Findings
- Questions

What we know about predicting the market:

1. There are large swings in market volatility:
   - ARCH & its variants work pretty well.
   - Volatility is negatively related to past returns.
2. Dividend growth isn't predictable.
3. No momentum, but negative serial correlation at 3-5 years.
4. d/p ratios forecast future returns, but not strongly:
   - Campbell and Shiller (1988), Fama and French (1988a), others.
5. Past returns and business-cycle indicators (*e.g.* TERM & DEF) also forecast future returns:
   - Fama and French (1988b, 1989)
6. Return predictability + zero dividend predictability $\Rightarrow$ high price volatility relative to dividends:
   - Shiller (1981), Cochrane (1991)

What we know about predicting the market:

1. There are large swings in market volatility:
   - ARCH & its variants work pretty well.
   - Volatility is negatively related to past returns.
2. Dividend growth isn't predictable.
3. No momentum, but negative serial correlation at 3-5 years.
4. d/p ratios forecast future returns, but not strongly:
   - Campbell and Shiller (1988), Fama and French (1988a), others.
5. Past returns and business-cycle indicators (*e.g.* TERM & DEF) also forecast future returns:
   - Fama and French (1988b, 1989)
6. Return predictability $+$ zero dividend predictability $\Rightarrow$ high price volatility relative to dividends:
   - Shiller (1981), Cochrane (1991)

What we know about predicting the market:

1. There are large swings in market volatility:
   - ARCH & its variants work pretty well.
   - Volatility is negatively related to past returns.
2. Dividend growth isn't predictable.
3. No momentum, but negative serial correlation at 3-5 years.
4. d/p ratios forecast future returns, but not strongly:
   - Campbell and Shiller (1988), Fama and French (1988a), others.
5. Past returns and business-cycle indicators (*e.g.* TERM & DEF) also forecast future returns:
   - Fama and French (1988b, 1989)
6. Return predictability + zero dividend predictability $\Rightarrow$ high price volatility relative to dividends:
   - Shiller (1981), Cochrane (1991)

What we know about predicting the market:

1. There are large swings in market volatility:
   - ARCH & its variants work pretty well.
   - Volatility is negatively related to past returns.
2. Dividend growth isn't predictable.
3. No momentum, but negative serial correlation at 3-5 years.
4. d/p ratios forecast future returns, but not strongly:
   - Campbell and Shiller (1988), Fama and French (1988a), others.
5. Past returns and business-cycle indicators (*e.g.* TERM & DEF) also forecast future returns:
   - Fama and French (1988b, 1989)
6. Return predictability + zero dividend predictability $\Rightarrow$ high price volatility relative to dividends:
   - Shiller (1981), Cochrane (1991)

What we know about predicting the market:

1. There are large swings in market volatility:
   - ARCH & its variants work pretty well.
   - Volatility is negatively related to past returns.
2. Dividend growth isn't predictable.
3. No momentum, but negative serial correlation at 3-5 years.
4. d/p ratios forecast future returns, but not strongly:
   - Campbell and Shiller (1988), Fama and French (1988a), others.
5. Past returns and business-cycle indicators (*e.g.* TERM & DEF) also forecast future returns:
   - Fama and French (1988b, 1989)
6. Return predictability + zero dividend predictability $\Rightarrow$ high price volatility relative to dividends:
   - Shiller (1981), Cochrane (1991)

What we know about predicting the market:

1. There are large swings in market volatility:
   - ARCH & its variants work pretty well.
   - Volatility is negatively related to past returns.
2. Dividend growth isn't predictable.
3. No momentum, but negative serial correlation at 3-5 years.
4. d/p ratios forecast future returns, but not strongly:
   - Campbell and Shiller (1988), Fama and French (1988a), others.
5. Past returns and business-cycle indicators (*e.g.* TERM & DEF) also forecast future returns:
   - Fama and French (1988b, 1989)
6. Return predictability $+$ zero dividend predictability $\Rightarrow$ high price volatility relative to dividends:
   - Shiller (1981), Cochrane (1991)

Predicting the Market

From Cochrane (2008), "The Dog that Did Not Bark: A Defense of Market Predictability:"

Table 1
Forecasting regressions

Regression	b	t	$R^2(\%)$	$\sigma(bx)(\%)$
$R_{t+1} = a + b(D_t/P_t) + \varepsilon_{t+1}$	3.39	2.28	5.8	4.9
$R_{t+1} - R_t^f = a + b(D_t/P_t) + \varepsilon_{t+1}$	3.83	2.61	7.4	5.6
$D_{t+1}/D_t = a + b(D_t/P_t) + \varepsilon_{t+1}$	0.07	0.06	0.0001	0.001
$r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{t+1}^r$	0.097	1.92	4.0	4.0
$\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{t+1}^{dp}$	0.008	0.18	0.00	0.003

R_{t+1} is the real return, deflated by the CPI, D_{t+1}/D_t is real dividend growth, and D_t/P_t is the dividend-price ratio of the CRSP value-weighted portfolio. R_{t+1}^f is the real return on 3-month Treasury-Bills. Small letters are logs of corresponding capital letters. Annual data, 1926–2004. $\sigma(bx)$ gives the standard deviation of the fitted value of the regression.

Predicting the Market

- The monthly R^2s on the previous page rise to 30-60% at long-horizons, depending on horizon and time-period
  - Fama and French (1988a)
- The OLS t-statistics (on previous page) are biased:
  - heteroskedasticity and Stambaugh (1986).
- Out of sample, d/p based forecasts are ineffective:
  - Goyal and Welch (1999), and others
- However, as Cochrane (2008) emphasizes, if *both* dividends and returns are unpredictable, then the d/p ratio must be constant.
  - Cochrane argues that tests reject dividend predictability at levels that require return predictability.

Predicting the Market

- The monthly R^2s on the previous page rise to 30-60% at long-horizons, depending on horizon and time-period
  - Fama and French (1988a)
- The OLS t-statistics (on previous page) are biased:
  - heteroskedasticity and Stambaugh (1986).
- Out of sample, d/p based forecasts are ineffective:
  - Goyal and Welch (1999), and others
- However, as Cochrane (2008) emphasizes, if *both* dividends and returns are unpredictable, then the d/p ratio must be constant.
  - Cochrane argues that tests reject dividend predictability at levels that require return predictability.

Predicting the Market

- The monthly R^2s on the previous page rise to 30-60% at long-horizons, depending on horizon and time-period
  - Fama and French (1988a)
- The OLS t-statistics (on previous page) are biased:
  - heteroskedasticity and Stambaugh (1986).
- Out of sample, d/p based forecasts are ineffective:
  - Goyal and Welch (1999), and others
- However, as Cochrane (2008) emphasizes, if *both* dividends and returns are unpredictable, then the d/p ratio must be constant.
  - Cochrane argues that tests reject dividend predictability at levels that require return predictability.

Predicting the Market

- The monthly R^2s on the previous page rise to 30-60% at long-horizons, depending on horizon and time-period
  - Fama and French (1988a)
- The OLS t-statistics (on previous page) are biased:
  - heteroskedasticity and Stambaugh (1986).
- Out of sample, d/p based forecasts are ineffective:
  - Goyal and Welch (1999), and others
- However, as Cochrane (2008) emphasizes, if *both* dividends and returns are unpredictable, then the d/p ratio must be constant.
  - Cochrane argues that tests reject dividend predictability at levels that require return predictability.

Predicting the Cross-Section

In contrast to the evidence on the predictability of the market, there is large predictability in the cross-section:

Portfolio Weights (%)						
Mkt	SMB	HML	UMD	ISU	ACR	S.R.*
100.0	–	–	–	–	–	0.31
75.1	24.9	–	–	–	–	0.32
28.2	14.6	57.2	–	–	–	0.80
21.1	10.2	41.9	26.8	–	–	1.18
18.8	15.3	13.9	9.6	42.4	–	1.55
17.4	14.5	12.3	8.2	36.7	11.0	1.60

- This table shows realized Sharpe-Ratios of (*ex-post* optimal) strategy combinations
- Note that this SR $\Rightarrow$ extremely high pricing kernel volatility ($\sigma_m > 160\%/\text{year}$)

Setting

- Assets:
  1. Risk-free asset, perfectly elastic supply.
  2. Single risky asset in fixed (unit) supply.
     - Dividend process is Geometric Brownian Motion
- Agents:
  - Select portfolio each period to maximize expected utility over next period's wealth.
    - Holding of risky asset restricted to $\alpha_L \leq \alpha_{t,i} \leq \alpha_H$.
  - Myopic power utility, $\gamma = 3.5$
  - Consume constant fraction of wealth/period.
  - Forecast mean and variance of market returns using a variety *ad-hoc* rules.
  - Given CRRA setup, market power moves with their wealth.
  - In some simulations, agents also slowly learn rule performance, and switch to profitable rules.

Setting

- Assets:
  1. Risk-free asset, perfectly elastic supply.
  2. Single risky asset in fixed (unit) supply.
     - Dividend process is Geometric Brownian Motion
- Agents:
  - Select portfolio each period to maximize expected utility over next period's wealth.
    - Holding of risky asset restricted to $\alpha_L \leq \alpha_{t,i} \leq \alpha_H$.
  - Myopic power utility, $\gamma = 3.5$
  - Consume constant fraction of wealth/period.
  - Forecast mean and variance of market returns using a variety *ad-hoc* rules.
  - Given CRRA setup, market power moves with their wealth.
  - In some simulations, agents also slowly learn rule performance, and switch to profitable rules.

Setting

- Economy:
  - The economy is run over 200,000 weeks.
    - $\approx$ 4,000 years.
  - 16,000 agents
  - 4,000 forecast rules
    - Variety of rules and gain parameters.

Forecast Rules – *Mean Returns*

1. *Fundamental Strategy*: (Value Investor)
$$f_t^j = \bar{r}_t + \beta_t^j \left(pd_t - \overline{pd}\right)$$
2. *Adaptive Linear Forecast:* (Momentum Trader)
$$f_t^j = f_{t-1}^j + g_j \left(r_t - f_{t-1}^j\right)$$
3. *Linear Regression*: (Technical Analyst, "Noise Trader")
$$f_t^j = \bar{r} + \sum_{i=0}^{2 \text{ weeks}} \beta_{t,i}^j (r_{t-i} - \bar{r})$$
4. *Benchmark*: (Passive Investor)
$$f_t^j = \bar{r}$$

Forecast Rules – *Mean Returns*

1. *Fundamental Strategy*: (Value Investor)

$$f_t^j = \bar{r}_t + \beta_t^j \left(pd_t - \overline{pd}\right)$$

2. *Adaptive Linear Forecast:* (Momentum Trader)

$$f_t^j = f_{t-1}^j + g_j \left(r_t - f_{t-1}^j\right)$$

3. *Linear Regression*: (Technical Analyst, “Noise Trader”)

$$f_t^j = \bar{r} + \sum_{i=0}^{2 \text{ weeks}} \beta_{t,i}^j (r_{t-i} - \bar{r})$$

4. *Benchmark*: (Passive Investor)

$$f_t^j = \bar{r}$$

Forecast Rules – *Mean Returns*

1. *Fundamental Strategy*: (Value Investor)

$$f_t^j = \bar{r}_t + \beta_t^j \left(pd_t - \overline{pd} \right)$$

2. *Adaptive Linear Forecast:* (Momentum Trader)

$$f_t^j = f_{t-1}^j + g_j \left(r_t - f_{t-1}^j \right)$$

3. *Linear Regression*: (Technical Analyst, "Noise Trader")

$$f_t^j = \bar{r} + \sum_{i=0}^{2 \text{ weeks}} \beta_{t,i}^j (r_{t-i} - \bar{r})$$

4. *Benchmark*: (Passive Investor)

$$f_t^j = \bar{r}$$

Forecast Rules – *Mean Returns*

1. *Fundamental Strategy*: (Value Investor)

$$f_t^j = \bar{r}_t + \beta_t^j \left(pd_t - \overline{pd} \right)$$

2. *Adaptive Linear Forecast:* (Momentum Trader)

$$f_t^j = f_{t-1}^j + g_j \left(r_t - f_{t-1}^j \right)$$

3. *Linear Regression*: (Technical Analyst, "Noise Trader")

$$f_t^j = \bar{r} + \sum_{i=0}^{2 \text{ weeks}} \beta_{t,i}^j (r_{t-i} - \bar{r})$$

4. *Benchmark*: (Passive Investor)

$$f_t^j = \bar{r}$$

Forecast Rules

- The parameters of the expected-return forecast rules are updated based on a set of gain parameters
  - These take values that correspond to half-lives ranging from 1 to 50 years
    - $g \in \{1, 2.5, 7, 18, 50\}$ years
- Variances are forecast using an GARCH(1,1)-like model.
  - Agents update variance estimates with a probabilistic lag.
    - This introduces some heterogeneity, preventing really severe market swings.

Forecast Rules

- The parameters of the expected-return forecast rules are updated based on a set of gain parameters
  - These take values that correspond to half-lives ranging from 1 to 50 years
    - $g \in \{1, 2.5, 7, 18, 50\}$ years
- Variances are forecast using an GARCH(1,1)-like model.
  - Agents update variance estimates with a probabilistic lag.
    - This introduces some heterogeneity, preventing really severe market swings.

Related Literature

- Hong and Stein (1999) propose a related model with two sets of agents:
  1. *Newswatchers* obtain signals about future cash flows, but with a stochastic delay.
  2. *Momentum traders* trade based on a limited history of prices and do not observe fundamental information.
- Depending on their recent past performance, the groups gain or lose capital.
- As *Newswatchers* become domainant, momentum strategies become profitable, and *vice-versa*
- A combination of short-horizon (high-gain) learning and shocks result in swings in profitability and dominance of strategies and traders.

Related Literature

- Hong and Stein (1999) propose a related model with two sets of agents:
  1. *Newswatchers* obtain signals about future cash flows, but with a stochastic delay.
  2. *Momentum traders* trade based on a limited history of prices and do not observe fundamental information.
- Depending on their recent past performance, the groups gain or lose capital.
- As *Newswatchers* become domainant, momentum strategies become profitable, and *vice-versa*
- A combination of short-horizon (high-gain) learning and shocks result in swings in profitability and dominance of strategies and traders.

Related Literature

- Hong and Stein (1999) propose a related model with two sets of agents:
  1. *Newswatchers* obtain signals about future cash flows, but with a stochastic delay.
  2. *Momentum traders* trade based on a limited history of prices and do not observe fundamental information.
- Depending on their recent past performance, the groups gain or lose capital.
- As *Newswatchers* become domainant, momentum strategies become profitable, and *vice-versa*
- A combination of short-horizon (high-gain) learning and shocks result in swings in profitability and dominance of strategies and traders.

Related Literature

- Hong and Stein (1999) propose a related model with two sets of agents:
  1. *Newswatchers* obtain signals about future cash flows, but with a stochastic delay.
  2. *Momentum traders* trade based on a limited history of prices and do not observe fundamental information.
- Depending on their recent past performance, the groups gain or lose capital.
- As *Newswatchers* become domainant, momentum strategies become profitable, and *vice-versa*
- A combination of short-horizon (high-gain) learning and shocks result in swings in profitability and dominance of strategies and traders.

Convergence to Equilibrium

- An initial simulation is run with slow learning
  - learning half-lives are 40-50 years.

"The objective is to look at a poplulation of agents restricted to only using long time series in their decision making and learning dynamics."

- The model converges to REE-like behavior
  - d/p ratios and market volatility are steady.

Returns

However, with heterogeneity (*e.g.*, high- and low-gain agents) in the economy, we see:

- Volatility Clustering
- Crashes
- Skewness

Wealth Dynamics

- "Buy-and-hold controls almost 50% of the wealth."

  "It is interesting that there is still enough wealth controlled by the dynamic strategies to have an impact on pricing, even though they are only about 30 percent of the market."

  - Not really that surprising.
- "Adaptive strategies [Momentum & Technical] strategies are 2nd ..., followed by Fundamental, followed by a very small fraction of the noise traders."
  - This is surprising!
  - Here, the Warren Buffets of the world don't, over time, do well.

Wealth Dynamics

- "Buy-and-hold controls almost 50% of the wealth."

  "It is interesting that there is still enough wealth controlled by the dynamic strategies to have an impact on pricing, even though they are only about 30 percent of the market."

  - Not really that surprising.
- "Adaptive strategies [Momentum & Technical] strategies are 2nd ..., followed by Fundamental, followed by a very small fraction of the noise traders."
  - This is surprising!
  - Here, the Warren Buffets of the world don't, over time, do well.

Wealth Dynamics

- “Buy-and-hold controls almost 50% of the wealth.”

  “It is interesting that there is still enough wealth controlled by the dynamic strategies to have an impact on pricing, even though they are only about 30 percent of the market.”

  - Not really that surprising.
- “Adaptive strategies [Momentum & Technical] strategies are 2nd ..., followed by Fundamental, followed by a very small fraction of the noise traders.”
  - This is surprising!
  - Here, the Warren Buffets of the world don’t, over time, do well.

Slow, Fundamental Traders

- In the Hong/Stein model, neither group alone creates an efficient market
  - Here, it seems like the slow value investors would come pretty close.
- Blake spends considerable time discussing why the fundamental traders don't do better around market crashes.
  - *They often don't buy the risky asset following a crash, and don't sell out as the bubble is forming.*

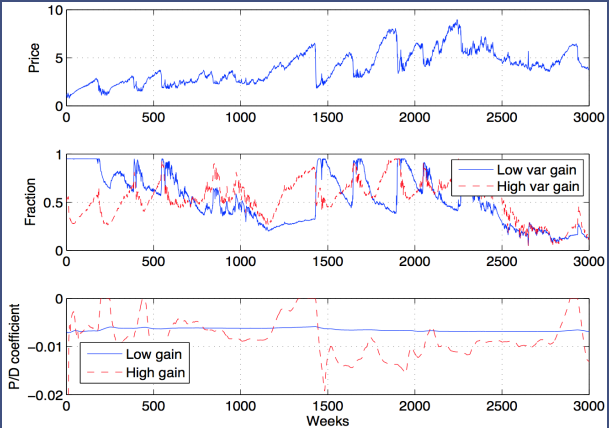
Why the fundamental traders don't succeed.

- The two reasons Blake gives are:
  1. Market crashes increase the estimated volatility for the high gain forecasters, which keeps the fundamental traders out.
     - But the low gain fundamental to pretty well.
  2. "...the estimated models will weaken as the market goes into a bubble period ... They have lost faith in the basic fundamental forecast." (p. 17)

Why the fundamental traders don't succeed.

- The two reasons Blake gives are:
  1. Market crashes increase the estimated volatility for the high gain forecasters, which keeps the fundamental traders out.
     - But the low gain fundamental to pretty well.
  2. "...the estimated models will weaken as the market goes into a bubble period ... They have lost faith in the basic fundamental forecast." (p. 17)

Strategy Fractions By Gain Level (Figure 12)



Questions:

- Long-term Survival
- Meta-Rules
  - Could combinations of rules (e.g., a combination of value and momenutm w/ low gain) dominate?
  - If not, what are the reasons for the exclusion of these rules
- Crowds
  - There are 16,000 agents in all simulations – How important is this?
  - One really interesting area is crowd dynamics. Are there any?
- Robustness of Findings
  - Quite a few little details seem to be important in getting to the results.
- Identification/Estimation?

Questions:

- Long-term Survival
- Meta-Rules
  - Could combinations of rules (e.g., a combination of value and momenutm w/ low gain) dominate?
  - If not, what are the reasons for the exclusion of these rules
- Crowds
  - There are 16,000 agents in all simulations – How important is this?
  - One really interesting area is crowd dynamics. Are there any?
- Robustness of Findings
  - Quite a few little details seem to be important in getting to the results.
- Identification/Estimation?

Questions:

- Long-term Survival
- Meta-Rules
  - Could combinations of rules (e.g., a combination of value and momenutm w/ low gain) dominate?
  - If not, what are the reasons for the exclusion of these rules
- Crowds
  - There are 16,000 agents in all simulations – How important is this?
  - One really interesting area is crowd dynamics. Are there any?
- Robustness of Findings
  - Quite a few little details seem to be important in getting to the results.
- Identification/Estimation?

Questions:

- Long-term Survival
- Meta-Rules
  - Could combinations of rules (e.g., a combination of value and momenutm w/ low gain) dominate?
  - If not, what are the reasons for the exclusion of these rules
- Crowds
  - There are 16,000 agents in all simulations – How important is this?
  - One really interesting area is crowd dynamics. Are there any?
- Robustness of Findings
  - Quite a few little details seem to be important in getting to the results.
- Identification/Estimation?

Questions:

- Long-term Survival
- Meta-Rules
  - Could combinations of rules (e.g., a combination of value and momenutm w/ low gain) dominate?
  - If not, what are the reasons for the exclusion of these rules
- Crowds
  - There are 16,000 agents in all simulations – How important is this?
  - One really interesting area is crowd dynamics. Are there any?
- Robustness of Findings
  - Quite a few little details seem to be important in getting to the results.
- Identification/Estimation?

References I

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Cochrane, John H., 1991, Volatility tests and efficient markets: Review essay, *Journal of Monetary Economics* 27, 463–485.

———, 2008, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* 21, 1533–1575.

Fama, Eugene F., and Kenneth R. French, 1988a, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

———, 1988b, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

———, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Goyal, Amit, and Ivo Welch, 1999, The myth of predictability: Does the dividend yield orecast the equity premium?, Anderson School, UCLA.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

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

References II

Stambaugh, Robert F., 1986, Discussion - does the stock market rationally reflect fundamental values?, *Journal of Finance* 41, 601–602.