Excess Volatility: Beyond Discount Rates

Stefano Giglio       Bryan Kelly

University of Chicago Booth School of Business
Internal inconsistency in the price behavior of short and long maturity claims

Prices on long end of curve dramatically more variable than justified by behavior of short end

This form of excess volatility is

Irreconcilable with standard models of prices
Not explainable with variation in discount rates

Evident in all asset classes we study:

- equity volatility
- sovereign, corporate credit risk
- interest rates
- inflation
- currency volatility
- commodities
Representative Result: S&P 500 Variance Swaps

Introduction

Factors=2, $R^2=99.6\%$
This Talk

Introduction

1. Introduction
2. Term Structure Methodology
   ▶ Detailed example
   ▶ Model, estimation, testing
3. Empirical Findings
   ▶ Excess volatility in many asset classes
4. Potential Explanations
   ▶ Missing factors
   ▶ Long-range dependence
   ▶ Non-linearities
   ▶ Mispricing
     ▶ Trading strategy
     ▶ Extrapolative expectations
5. Robustness
6. Conclusions
Term Structure Methodology

(Detailed One-factor Example)
Term Structure: Securities

Consider asset class that pays a per period cash flow of $x_{t+j}$

- Term structure: Prices of claims to $x_{t+j}$ over next $j = 1, 2, \ldots$ months
- Definition of securities*
  - Price of an $n$-month forward claim is
    \[ f_{t,n} = E_t^Q[x_{t+n}] \]
  - Price of a cumulative claim is
    \[ p_{t,n} = E_t^Q[x_{t+1} + \ldots + x_{t+n}] \]
  - Treat exponential claims identically (with additional assumptions)

* Assume risk free rate is zero and constant for this example
Restrictions Implied by One-factor Model

One-factor Example Model

Under the pricing measure, $Q$, cash flows evolve as

$$x_t = \rho x_{t-1} + \epsilon_t$$

- Model $Q$ process for $x$, instead of modeling $M$ and $x$, to study $p = E[Mx]$

- “Affine-$Q$” dynamics imply $f_{t,n} = E_t^Q [x_{t+n}]$ obeys

$$f_{t,1} = \rho x_t, \quad f_{t,2} = \rho^2 x_t, \quad \ldots, \quad f_{t,n} = \rho^n x_t$$

- Cross-equation restrictions
  1. Strict linear factor structure
  2. Loadings on $x_t$ are geometrically decaying function of $\rho$

- Given $f_{t,1}$ and an estimate of $\rho$, the behavior of the entire term structure is pinned down
Restrictions in Terms of Price Volatility

- Equivalent restrictions for cumulative claims ($p_{t,j} = E_t^Q[x_{t+1} + ... + x_{t+j}]$)

  \[ p_{t,1} = \rho x_t, \quad p_{t,2} = (\rho + \rho^2)x_t, \quad \ldots, \quad p_{t,n} = (\rho + \ldots + \rho^n)x_t \]

- Can be recast in terms of price volatility

  \[ \sigma(p_{t,1}) = \rho \sigma(x_t), \quad \sigma(p_{t,2}) = (\rho + \rho^2)\sigma(x_t), \quad \ldots, \quad \sigma(p_{t,n}) = (\rho + \ldots + \rho^n)\sigma(x_t) \]
Restrictions in Terms of Price Volatility

- Equivalent restrictions for cumulative claims ($p_{t,j} = E_t^Q[x_{t+1} + \ldots + x_{t+j}]$)

  
  \[ p_{t,1} = \rho x_t, \quad p_{t,2} = (\rho + \rho^2)x_t, \quad \ldots, \quad p_{t,n} = (\rho + \ldots + \rho^n)x_t \]

- Can be recast in terms of price volatility

  \[ \sigma(p_{t,1}) = \rho \sigma(x_t), \quad \sigma(p_{t,2}) = (\rho + \rho^2)\sigma(x_t), \quad \ldots, \quad \sigma(p_{t,n}) = (\rho + \ldots + \rho^n)\sigma(x_t) \]
Excess Volatility Hypothesis

Hunch

Prices of long maturity claims move around “too much” (Shiller 1979, Stein 1989)

- How much is “too much”?
- Affine-$\mathbb{Q}$ model ($f_{t,j} = \rho^j x_t$) makes specific prediction

\[
\sigma(f_{t,120}) = \rho^{120} \sigma(x_t) = \rho^{120-1} \frac{\rho \sigma(x_t)}{\sigma(f_{t,1})}
\]

- A precise notion of “too much”: $\sigma(f_{t,120})/ (\rho^{120-1} \sigma(f_{t,1})) = 1$

- Key input: Estimate of $\rho$
Our Excess Volatility Test

How to operationalize this test? Key input: Estimate of $\rho$

1. $p_{t,1}$ stands in for factor
2. Regress $p_{t,2}$ on $p_{t,1}$. Estimate is $\hat{b}_2 = \frac{\text{Cov}(\rho x_t, [\rho + \rho^2] x_t)}{\text{Var}(\rho x_t)} = 1 + \hat{\rho}$
   ► Reveals $\mathbb{Q}$-persistence of $x_t$

<table>
<thead>
<tr>
<th>Restricted</th>
<th>Unrestricted</th>
</tr>
</thead>
<tbody>
<tr>
<td>► Restricted coefficient is $\tilde{b}_n(\hat{\rho}) = \frac{\rho + \ldots + \rho^n}{\hat{\rho}}$</td>
<td>► Regress $p_{t,n}$ on $p_{t,1}$, coeff $\hat{b}_n$</td>
</tr>
<tr>
<td>► Fitted price is $\tilde{p}_{t,n} = \tilde{b}<em>n p</em>{t,1}$</td>
<td>► Fitted price is $\hat{p}_{t,n} = \hat{b}<em>n p</em>{t,1}$</td>
</tr>
<tr>
<td>► Restricted volatility is $\sigma(\tilde{p}_{t,n}) = \tilde{b}<em>n \sigma(p</em>{t,1})$</td>
<td>► High $R^2$ regression</td>
</tr>
<tr>
<td></td>
<td>► Unrestricted volatility is $\sigma(\hat{p}_{t,n}) = \hat{b}<em>n \sigma(p</em>{t,1})$</td>
</tr>
</tbody>
</table>

$$VR_n = \frac{\sigma^2(\hat{p}_{t,n})}{\sigma^2(\tilde{p}_{t,n})} = \left(\frac{\hat{b}_n}{\tilde{b}_n}\right)^2$$
Our Excess Volatility Test: Visualization

\[ \sigma(p_n) \]

Maturity \( (n) \)

0 2 4 6 8 10
Our Excess Volatility Test: Visualization

\[ \sigma(p_n) \]

\[
\left. \left( \rho^Q + \rho^Q x \right) \sigma_x \right\} \Rightarrow \hat{\rho}^Q
\]

Maturity \((n)\)

0 2 4 6 8 10
Our Excess Volatility Test: Visualization

\[ \sigma(p_n) = \rho_Q \sigma_x + (\rho_Q^2) \sigma_x \]

\[ \sum_{j=1}^{n} (\hat{\rho}_Q^j) \sigma_x \rightarrow \hat{\rho}_Q \]
Long Versus Short, or Short Versus Long

There are many ways to test cross-equation restrictions

Reasons for our formulation

1. Long end overreacts or short end underreacts?
   ▶ Estimate long/test short asymptotically equivalent
   ▶ Long-end is overreacting also to cash flows, \( f_{t,n}/\rho^n = a + bx_t + e \)

2. Tests using full term structure (e.g. likelihood ratio) lack power to detect these particular violations in finite samples (see Table 5)

3. Speaks to long history of variance ratios tests in finance/economics
Our Excess Volatility Test: Preview of Results

S&P 500 Variance Swaps

Factors = 2, $R^2 = 99.6\%$

Unrestricted

Restricted

95% Test
Our Excess Volatility Test: Preview of Results

- Variance ratios for long maturity claims are large, typically $> 2$

- Prices of long lived claims are excessively volatile
  1. Given behavior of short maturity claims
  2. Given model restrictions
Strict linear factor structure

- In unrestricted regressions on factors
  \[ f_{t,1} = b_1x_t, \quad f_{t,2} = b_2x_t, \quad \ldots, \quad f_{t,n} = b_nx_t \]

- 1–3 common factors (e.g. PC’s) explain >99% of variation in panel

Loadings on \( x_t \) are geometrically decaying function of \( \rho \)

- Restricted regression coefficients are strongly rejected on long end
  \[ f_{t,1} = \rho x_t, \quad f_{t,2} = \rho^2 x_t, \quad \ldots, \quad f_{t,n} = \rho^n x_t \]

- Restricted coefficients are too small compared to unrestricted \( b_n \)
- This is a symptom of overreaction of long-lived claims
Term Structure Methodology

(General Version)
General Methodology for Excess Volatility Test

- General setting is for $K$-factor affine-$\mathbb{Q}$ model

\[ x_t = \delta' \begin{bmatrix} H_t \end{bmatrix}_K, \quad H_t = \rho H_{t-1} + \epsilon_t, \quad \rho = \begin{bmatrix} \rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \rho_K \end{bmatrix} \]

- Vast majority of structural and reduced-form asset pricing models are in this family

- Main adjustment versus one-factor setting,
  - Estimate $\rho$ from regression of $p_{t,K+1}$ on $p_{t,1}, \ldots, p_{t,K}$
  - $K$ coefficients map directly to persistence parameters $\rho_1, \ldots, \rho_K$

- Variance ratio test becomes

\[ VR_n = \frac{\hat{b}_n' \Sigma_K \hat{b}_n}{\tilde{b}_n' \Sigma_K \tilde{b}_n}, \quad \text{where} \quad \Sigma_K = Var([p_{t,1}, \ldots, p_{t,1}]) \]
General Methodology for Excess Volatility Test

- Further generalizations
  - Exponential-affine claims (prices of the form $E_t^Q[\exp(x_{t+n})]$)
  - Heteroskedasticity
  - Variable risk-free rates
  - Measurement error

- We allow for all of these in a range of robustness checks
What Might Be Missing?

1. Time-varying risk premia
2. Missing factors
3. Non-linear dynamics of factors under \( \mathbb{Q} \)
4. Long-memory dynamics of factors under \( \mathbb{Q} \)
5. Mispricing

- 1 is ruled out because we are working under \( \mathbb{Q} \) only
- 2–5 are important possibilities that we examine in detail

My talk will spend substantial time on this
Empirical Results
Data

- Equity variance claims (swaps, options)
- Currency variance claims (options)
- Credit default swaps (sovereign, corporate)
- Inflation swaps
- Interest rates (Treasuries)
- Commodity futures

Only look at claims/maturities with verifiable and sufficient liquidity
Main Results

Starting point:

- Set number of factors, $K$, as number of principal components necessary to explain at least 99% of the variance in the panel of all $n$ available maturities

- Analyze additional choices in robustness
Equity Variance Claims (I)

Panel A: IBM IV
Factors=2, $R^2=99.7\%$

Panel B: Apple IV
Factors=2, $R^2=99.4\%$

Panel C: Citigroup IV
Factors=2, $R^2=99.8\%$

Panel D: S&P 500 IV
Factors=2, $R^2=99.5\%$
Equity Variance Claims (II)

Panel A: STOXX 50 IV
Factors=2, $R^2=99.8\%$

Panel B: FTSE 100 IV
Factors=2, $R^2=99.8\%$

Panel C: DAX IV
Factors=2, $R^2=99.9\%$

Panel D: NASDAQ IV
Factors=2, $R^2=99.8\%$
Currency Variance Claims

Panel A: GBP/USD
Factors=2, $R^2=99.8\%$

Panel B: GBP/JPY
Factors=2, $R^2=99.7\%$

Panel C: USD/CHF
Factors=2, $R^2=99.6\%$
Liquidity at reported maturities supported by DTCC data
Commodities

Panel A: Crude Oil Futures
Factors = 3, $R^2 = 99.7\%$

Panel B: Gold Futures
Factors = 2, $R^2 = 99.7\%$
Inflation Swaps

Factors=4, $R^2=99.4\%$

Data used by Fleckenstein, Lustig, Longstaff (2013, 2014)
Interest Rates

- Shiller (1979), Gurkaynak et al. (2005), Hanson and Stein (2015)
- Of all term structures, this displays by far the least excess volatility
Possible Explanations
(Misspecification vs. Mispricing)
Possible Explanations

- The empirical results reject the affine-\(\mathbb{Q}\) specification
- Next question: Why?
  - Alternative model compatible with the data?
  - Mispricing?

Joint Hypothesis Problem, Fama (1991)

Market efficiency per se is not testable. It must be tested jointly with some model of equilibrium, an asset-pricing model. ... As a result, when we find anomalous evidence on the behavior of returns, the way it should be split between market inefficiency or a bad model of market equilibrium is ambiguous.

- While ambiguity is unavoidable, we can explore leading potential explanations
Potential Sources of Violation

1. Missing factors
2. Long-memory $\mathbb{Q}$ dynamics
3. Non-linear $\mathbb{Q}$ dynamics
4. Overreaction and mispricing
Missing Factors?

Model

- To explain the results, one would need a missing factor that is
  1) very persistent, 2) low variance
- This appears quantitatively infeasible (Section 4.1)

Empirics

- Can add factors, increasing $R^2$ even further (robustness Table A6)
- Conclusion: Missing factors unlikely driver
Long Memory

- Excessive vol of long-lived claims intuitively raises question: Are long memory cash flow dynamics behind this?
- Shocks decay more slowly than the geometric rate in affine model
- Investigate ARFIMA family of long-memory processes
- “Fractionally integrated $d$” plus ARMA, $d \in (-0.5, 0.5)$ stationary
Long Memory

- Simulate a wide range of specifications ARFIMA(1, d, 0)
  - $d \in (0, 0.5)$ and AR(1) = 0.25, 0.50, or 0.75
  - $K = 1, 2, \text{ or } 3$ factors in estimation

<table>
<thead>
<tr>
<th>$d$</th>
<th>$K$</th>
<th>AR(1)=0.25</th>
<th>AR(1)=0.50</th>
<th>AR(1)=0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$R^2$</td>
<td>$VR_{12}$</td>
<td>$VR_{24}$</td>
</tr>
<tr>
<td>0.10</td>
<td>1</td>
<td>96.8</td>
<td>2.0</td>
<td>2.9</td>
</tr>
<tr>
<td>0.10</td>
<td>2</td>
<td>100.0</td>
<td>1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>0.10</td>
<td>3</td>
<td>100.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>0.30</td>
<td>1</td>
<td>97.7</td>
<td>2.5</td>
<td>4.8</td>
</tr>
<tr>
<td>0.30</td>
<td>2</td>
<td>100.0</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>0.30</td>
<td>3</td>
<td>100.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>0.49</td>
<td>1</td>
<td>98.7</td>
<td>2.3</td>
<td>4.9</td>
</tr>
<tr>
<td>0.49</td>
<td>2</td>
<td>100.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>0.49</td>
<td>3</td>
<td>100.0</td>
<td>1.0</td>
<td>1.1</td>
</tr>
</tbody>
</table>

- Inconsistent with robustness tests, trading strategy
Non-linear Dynamics

- Evolution of cash flows is state-dependent
- Simulate a large class of non-linear dynamics, STAR models of Granger and Terasvirta (1993)

\[ x_t = \rho x_{t-1} \left( 1 - (1 + e^{-\gamma (x_{t-1} - c)})^{-1} \right) + (1 - \rho) x_{t-1} \left( 1 + e^{-\gamma (x_{t-1} - c)} \right)^{-1} + \epsilon_t. \]

\[ \text{Prob(state=1)} \]

\[ \text{Prob(state=2)} \]
Non-linear Dynamics

<table>
<thead>
<tr>
<th>γ</th>
<th>K</th>
<th>ρ=0.01</th>
<th></th>
<th>ρ=0.10</th>
<th></th>
<th>ρ=0.25</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R²</td>
<td>VR₁₂</td>
<td>R²</td>
<td>VR₁₂</td>
<td>R²</td>
<td>VR₁₂</td>
</tr>
<tr>
<td>0.1</td>
<td>1.0</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>0.1</td>
<td>2.0</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>0.1</td>
<td>3.0</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>0.5</td>
<td>1.0</td>
<td>98.6</td>
<td>1.22</td>
<td>99.9</td>
<td>1.04</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>0.5</td>
<td>2.0</td>
<td>100.0</td>
<td>1.04</td>
<td>100.0</td>
<td>1.01</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>0.5</td>
<td>3.0</td>
<td>100.0</td>
<td>1.01</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>99.8</td>
<td>1.02</td>
<td>99.7</td>
<td>1.05</td>
<td>100.0</td>
<td>1.01</td>
</tr>
<tr>
<td>1.0</td>
<td>2.0</td>
<td>100.0</td>
<td>1.01</td>
<td>100.0</td>
<td>1.01</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>1.0</td>
<td>3.0</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>0.99</td>
<td>100.0</td>
<td>0.99</td>
</tr>
<tr>
<td>5.0</td>
<td>1.0</td>
<td>99.9</td>
<td>1.00</td>
<td>99.9</td>
<td>1.01</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>5.0</td>
<td>2.0</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>5.0</td>
<td>3.0</td>
<td>100.0</td>
<td>1.00</td>
<td>100.0</td>
<td>0.99</td>
<td>100.0</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Suggests non-linearities unlikely the main driver
Mispricing: A Trading Strategy

Premise of Strategy

- Trading against mispricing should provide arbitrage profits
- ... while trading against a misspecified model should not

We design trading strategy based on affine-$Q$ specification

- Posits that affine model correctly represents “true value” of claims
- Exploits “mispricing” on long end arising due to overreaction
- Trades long maturity claims hedged with short maturities (a convergence trade via replication)

Details

- Strictly real-time information, estimate model recursively (250 days)
- Trade only if observed mispricing is “large enough”
- Limit leverage, volatility of strategy (capital/margin constraints)
A Trading Strategy

- Premise of Strategy: The model’s no-arbitrage conditions hold on average

\[ p_{t+n,N} = (b_N)' P_{t+n,1:K} \]  

\(^(*)\)

- At time \( t \), identifying mispricings using perfect replication

<table>
<thead>
<tr>
<th>Date</th>
<th>Strategy ( A )</th>
<th>Ongoing Value</th>
<th>Cash Flows</th>
<th>Strategy ( B )</th>
<th>Ongoing Value</th>
<th>Cash Flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>( p_{t,N+n} )</td>
<td></td>
<td></td>
<td>( b'<em>N P</em>{t,1+n:K+n} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t+1 )</td>
<td>( p_{t+1,N+n-1} )</td>
<td>( x_{t+1} )</td>
<td></td>
<td>( + (1 - b'<em>N \mathbf{1}) P</em>{t,n} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t+n )</td>
<td>( p_{t+n,N} )</td>
<td>( x_{t+n} )</td>
<td></td>
<td>( b'<em>N P</em>{t,1+n-1:K+n-1} )</td>
<td>( x_{t+1} )</td>
<td>( + (1 - b'<em>N \mathbf{1}) P</em>{t,n-1} )</td>
</tr>
</tbody>
</table>

\( \vdots \)

- \( A \) and \( B \) replicate \(^(*)\) at \( t \), and replicates all intermediate cash flows

- At time \( t \), trade if \( p_{t,N+n} \neq b'_N P_{t,1+n:K+n} + (1 - b'_N \mathbf{1}) P_{t,n} \)
# Implementation with Variance Swaps

## Annualized Sharpe Ratios: 1-month Holding Period

<table>
<thead>
<tr>
<th>Mispricing Threshold</th>
<th>Longest Maturity Traded</th>
<th>Variance Swaps</th>
<th>Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Missing Factor</td>
</tr>
<tr>
<td>50</td>
<td>15</td>
<td>0.73</td>
<td>-0.01</td>
</tr>
<tr>
<td>50</td>
<td>18</td>
<td>1.17</td>
<td>-0.01</td>
</tr>
<tr>
<td>50</td>
<td>21</td>
<td>0.94</td>
<td>-0.01</td>
</tr>
<tr>
<td>50</td>
<td>24</td>
<td>0.56</td>
<td>-0.01</td>
</tr>
<tr>
<td>75</td>
<td>15</td>
<td>1.43</td>
<td>0.00</td>
</tr>
<tr>
<td>75</td>
<td>18</td>
<td>1.68</td>
<td>0.00</td>
</tr>
<tr>
<td>75</td>
<td>21</td>
<td>1.37</td>
<td>0.00</td>
</tr>
<tr>
<td>75</td>
<td>24</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>15</td>
<td>1.56</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>18</td>
<td>1.96</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>21</td>
<td>1.91</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>24</td>
<td>1.61</td>
<td>0.00</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>1.28</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Implementation with No-Arbitrage Counterfactuals

1. Truth is 2-factors, estimated model has 1-factor
2. Truth borderline non-stationary ARFIMA ($d = 0.49$)
3. Truth is non-linear STAR model

<table>
<thead>
<tr>
<th>Mispricing Threshold</th>
<th>Longest Maturity Traded</th>
<th>Variance Swaps</th>
<th>Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Missing Factor</td>
</tr>
<tr>
<td>50</td>
<td>15</td>
<td>0.73</td>
<td>-0.01</td>
</tr>
<tr>
<td>50</td>
<td>18</td>
<td>1.17</td>
<td>-0.01</td>
</tr>
<tr>
<td>50</td>
<td>21</td>
<td>0.94</td>
<td>-0.01</td>
</tr>
<tr>
<td>50</td>
<td>24</td>
<td>0.56</td>
<td>-0.01</td>
</tr>
<tr>
<td>75</td>
<td>15</td>
<td>1.43</td>
<td>0.00</td>
</tr>
<tr>
<td>75</td>
<td>18</td>
<td>1.68</td>
<td>0.00</td>
</tr>
<tr>
<td>75</td>
<td>21</td>
<td>1.37</td>
<td>0.00</td>
</tr>
<tr>
<td>75</td>
<td>24</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>15</td>
<td>1.56</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>18</td>
<td>1.96</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>21</td>
<td>1.91</td>
<td>0.00</td>
</tr>
<tr>
<td>90</td>
<td>24</td>
<td>1.61</td>
<td>0.00</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>1.28</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Trading Strategy: Riskiness

- Possible that high S.R.s in variance swap data due to risk exposure

- Essentially no correlation with obvious risk factors
  - $E[r_{strat}] = 24\%$ p.a.
  - $\alpha_{FF} = 23\%$
  - $\alpha_{TS} = 24\%$

- Positive skewness
## Limits to Arbitrage

### Annualized Sharpe Ratios: 1-month Holding Period

<table>
<thead>
<tr>
<th>Mispricing Percentile</th>
<th>Longest Maturity Traded</th>
<th>0% TC</th>
<th>1% TC</th>
<th>2% TC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1M</td>
<td>3M</td>
<td>6M</td>
</tr>
<tr>
<td>50</td>
<td>18</td>
<td>1.17</td>
<td>1.26</td>
<td>0.98</td>
</tr>
<tr>
<td>50</td>
<td>21</td>
<td>0.94</td>
<td>1.12</td>
<td>1.10</td>
</tr>
<tr>
<td>75</td>
<td>18</td>
<td>1.68</td>
<td>1.34</td>
<td>1.52</td>
</tr>
<tr>
<td>75</td>
<td>21</td>
<td>1.37</td>
<td>1.46</td>
<td>1.43</td>
</tr>
<tr>
<td>90</td>
<td>18</td>
<td>1.96</td>
<td>2.26</td>
<td>1.70</td>
</tr>
<tr>
<td>90</td>
<td>21</td>
<td>1.91</td>
<td>2.45</td>
<td>1.54</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>1.28</td>
<td>1.28</td>
<td>1.12</td>
</tr>
</tbody>
</table>

### Panel A: Sharpe Ratio

<table>
<thead>
<tr>
<th>Mispricing Percentile</th>
<th>Longest Maturity Traded</th>
<th>0% TC</th>
<th>1% TC</th>
<th>2% TC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1M</td>
<td>3M</td>
<td>6M</td>
</tr>
<tr>
<td>50</td>
<td>18</td>
<td>0.55</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td>50</td>
<td>21</td>
<td>0.54</td>
<td>0.51</td>
<td>0.49</td>
</tr>
<tr>
<td>75</td>
<td>18</td>
<td>0.33</td>
<td>0.28</td>
<td>0.30</td>
</tr>
<tr>
<td>75</td>
<td>21</td>
<td>0.33</td>
<td>0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>90</td>
<td>18</td>
<td>0.16</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>90</td>
<td>21</td>
<td>0.15</td>
<td>0.13</td>
<td>0.14</td>
</tr>
</tbody>
</table>

### Panel B: Trading Frequency
A Model of Extrapolation

- What type of investor behavior might lead to mispricing?
- “Extrapolation: The Most Important Idea in Behavioral Economics”
  (Barberis and Shleifer 2003, Gennaioli and Shleifer 2014, Barberis, Greenwood, Jin, Shleifer (2015a,b), and others)
- Little exploration of how expectation formation can vary with horizon of expectation
- A stylized model of extrapolation lines up well with term structure facts

This assumption is usually motivated by Kahneman & Tversky’s (1974) representativeness heuristic. According to this heuristic, people expect even small samples of data to reflect the properties of the parent population. As a result, they draw overly strong inferences from these small samples, and this can lead to over-extrapolation. –Barberis (2013)
A Model of Extrapolation

- Process pinning down “true value”

\[ x_{t+1} = (1 - \rho)\mu + \rho x_t + \epsilon_{t+1} \]

- Extrapolative expectations summarized by replacing long run mean, \( \mu \), with distorted mean

\[ \mu_t^\theta = \mu + \theta(x_t - \mu) \]

- Distortion represents investor’s tendency to over-emphasize recent data when contemplating cash flow distribution

- Forecast/price for horizon \( n \) has exact 1-factor structure:

\[ f_{t,n} = E_t^\theta[x_{t+n}] = (1 - \rho^n)\mu_t^\theta + \rho^n x_t = \left[ \theta + (1 - \theta)\rho^n \right] x_t + \text{constant} \]

  - Non-geometric decay

- By Proposition 1, model admits arbitrage
Conclusion

- Internal inconsistency in the behavior of term structures
- Excess volatility for long maturity claims
- Pervasive phenomenon, many asset classes
- Not explicable with discount rate variation
- Evidence points to mispricing driven by overreaction

What We Require

- Affine (or exponential-affine) autoregression under $\mathbb{Q}$
- Unrestrictive ... $K$ is arbitrary, chosen by data

What We Do Not Require

- Do not require model/estimates of $\mathbb{P}$ dynamics
- No data other than asset prices (no need for cash flows $x_t$)
- Nor do we require a model of risk premia/discount rates