Have we solved the idiosyncratic volatility puzzle?

Roger Loh [1] Kewei Hou[2]

[1]Singapore Management University

[2]Ohio State University

Presented by Roger Loh
Proseminar
SMU Finance Ph.D class

The idiosyncratic volatility puzzle

- Ang, Hodrick, Xing, & Zhang (2006) find that idiosyncratic volatility (IVOL) and next-month cross-sectional returns are negatively related.
  - Puzzling because according to standard asset-pricing models (e.g. CAPM), non-systematic risk should not be priced (Fama and MacBeth, 1973)
  - Or if priced, the relation should be positive (Merton, 1987; Hirshleifer, 1988). Investors with undiversified portfolios demand positive premium for holding stocks with high idiosyncratic risk
- Many papers try to explain the puzzle. But not clear which explanation is best or whether the puzzle is fully explained.

Our paper

Provides a method to objectively quantify the marginal contribution of each existing story that claims to explain the puzzle.

Our contribution

1. Objective and agnostic approach
    - Most papers aim to remove the IVOL puzzle with their favorite explanation. We treat each potential candidate explanation seriously, without favorites.
    - Most papers just aim to make the IVOL coefficient insignificant. We can quantify the fraction of the puzzle that a candidate explains.
2. We pit existing explanations against one another
    - A common framework, standard sample, and fair horse race between explanations.
    - Existing papers usually do not consider competing explanations.
3. Our method can be used to evaluate any anomaly in asset-pricing (e.g. Chen, Strebulaev, Zhang, and Xing (2014), Bao, Chen, Hou, and Lu (2015))

Candidate explanations

1) Lottery Preference

1. Skewness (Barberis & Huang, 2008)
2. Co-skewness (Chabi-Yo & Yang, 2009)
3. Expected idiosyncratic skewness (Boyer, Mitton, & Vorkink, 2010)
4. Maximum daily return (Bali, Cakici, Whitelaw, 2011)
5. Retail-trading proportion (Han & Kumar, 2013)

2) Market Frictions

6. Lag Return (Fu, 2009; Huang, Liu, Rhee, & Zhang, 2009)
7. Amihud illiquidity (Han & Lesmond, 2009)
8. Zero-return measure (Han & Lesmond, 2009)
9. Bid-ask spread (Han & Lesmond, 2009)

3) Others

10. Dispersion (Ang et al., 2009)
11. Average variance beta (Chen & Petkova, 2012)
12. SUE (Wong, 2009; Jiang, Xu, & Yao, 2009)

Conditioning variables

We also examine the success of the best candidates in subsamples associated with a stronger IVOL puzzle:

1. Non-penny stocks (e.g. $>$ \$5, Bali & Cakici, 2008)
2. Low analyst coverage (George and Hwang, 2011)
3. Poor credit ratings (Avramov, Chordia, Jotova, & Philipov, 2013)
4. High short-sale constraints (George & Hwang, 2011)
5. High leverage (Johnson, 2004; Ang et al. 2009)
6. Low institutional ownership (Nagel, 2009)
7. High growth firms (Barinov, 2014)
8. Non-Nasdaq stocks (Bali & Cakici, 2008)
9. Non-January months (Doran, Jiang, & Peterson, 2012)

Decomposition methodology

- Start from Fama-MacBeth cross-sectional regressions each month t for all stocks i.

$$R_{it} = \alpha_t + \gamma_t IVOL_{it-1} + \epsilon_{it} \quad (1)$$

- Suppose we have a candidate explanation. $Candidate_{it-1}$ must be correlated with $IVOL_{it-1}$ to explain the IVOL puzzle. So we regress:

$$IVOL_{it-1} = a_{t-1} + \delta_{t-1} Candidate_{it-1} + \mu_{it-1} \quad (2)$$

- From above, we can decompose $IVOL_{it-1}$ into 2 components, $(\delta_{t-1} Candidate_{it-1})$ and $(a_{t-1} + \mu_{it-1})$.
  - First is the component of IVOL related to the candidate.
  - Second is a residual component unrelated to the candidate.

Decomposition methodology

- Using the linearity property in covariances, we decompose the estimated γ_t coefficient in equation (1): $R_{it} = \alpha_t + \gamma_t IVOL_{it-1} + \epsilon_{it}$.

$$\begin{aligned}
\gamma_t &= \frac{Cov[R_{it}, IVOL_{it-1}]}{Var[IVOL_{it-1}]} \\
&= \frac{Cov[R_{it}, (\delta_{t-1} Candidate_{it-1}) + (a_{t-1} + \mu_{it-1})]}{Var[IVOL_{it-1}]} \\
&= \frac{Cov[R_{it}, (\delta_{t-1} Candidate_{it-1})]}{Var[IVOL_{it-1}]} + \frac{Cov[R_{it}, (a_{t-1} + \mu_{it-1})]}{Var[IVOL_{it-1}]} \\
&= \gamma_t^C + \gamma_t^R \qquad (3)
\end{aligned}$$

- γ_t^C / γ_t is the fraction explained by the *Candidate*.
- We can obtain the mean explained fraction using Fama-MacBeth time-series averages: $\overline{\gamma_t^C} / \overline{\gamma_t}$, and the variance of this ratio using the multivariate delta method.

Relating to the conventional approach

- Conventional approach:

$$R_{it} = \tilde{\alpha}_t + \tilde{\gamma}_t^R IVOL_{it-1} + \tilde{\gamma}_t^C C_{it-1} + \tilde{\epsilon}_{it}. \quad (4)$$

- Which can be re-written as:

$$\begin{aligned} R_{it} &= \tilde{\alpha}_t + \tilde{\gamma}_t^R(a_{t-1} + \mu_{it-1} + \delta_{t-1}C_{it-1}) + \tilde{\gamma}^C C_{it-1} + \tilde{\epsilon}_{it} \\ R_{it} &= \tilde{\alpha}_t + \tilde{\gamma}_t^R(a_{t-1} + \mu_{it-1}) + \bar{\gamma}^C C_{it-1} + \tilde{\epsilon}_{it} \quad (5) \end{aligned}$$

where $\bar{\gamma}_t^C = \tilde{\gamma}_t^C + \delta_{t-1}\tilde{\gamma}_t^R$, is the coefficient when R_{it} is regressed on C_{it-1}.

- We can then rewrite our Equation 3 as follows:

$$\begin{aligned} \gamma_t^C &= \frac{Cov[R_{it}, \delta_{t-1}C_{it-1}]}{Var[IVOL_{it-1}]} \\ &= \frac{Cov[R_{it}, \delta_{t-1}C_{it-1}]}{Var[\delta_{t-1}C_{it-1}]} \times \frac{Var[\delta_{t-1}C_{it-1}]}{Var[IVOL_{it-1}]} \\ &= \frac{\bar{\gamma}_t^C}{\delta_{t-1}} \times \frac{Var[\delta_{t-1}C_{it-1}]}{Var[IVOL_{it-1}]} \\ &= (\frac{\tilde{\gamma}_t^C}{\delta_{t-1}} + \tilde{\gamma}_t^R) \times \frac{Var[\delta_{t-1}C_{it-1}]}{Var[IVOL_{it-1}]} \quad (6) \end{aligned}$$

Example with Skewness as candidate, Table 3A

Stage	Description	Variable	Skewness	
1	Regress returns on IVOL	Intercept	0.353***	(6.47)
		IVOL	-17.401***	(-8.47)
2	Add candidate variable	Intercept	0.355***	(6.47)
		IVOL	-16.145***	(-7.67)
		Candidate	-0.099***	(-5.53)
3	IVOL on candidate variable	Intercept	2.398***	(90.46)
		Candidate	0.367***	(34.31)
		Adj R-Sq	4.3%	
4	Decompose Stage 1 IVOL coefficient	Candidate	-1.785	
			10.3%***	(6.73)
		Residual	-15.615	
			89.7%***	(58.88)
		Total	-17.401***	(-8.47)
			100%	
	sample		1963 to 2012	
	avgnfirms		3563.7	

- IVOL-return relation $\overline{\gamma_t} = -17.401$ percent. Skewness can explain ($\overline{\gamma_t^C} = -1.785$) 10.3% of this relation.

Explained fraction of each univariate candidate

Story	No.	Candidate Variable	Fraction explained
Lottery preference	1	Skewness	10.3%***
	2	CoSkewness	1.9%
	3	E(idioskew)	14.7%***
	4	Maxret	112.0%***
	5	RTP	22.3%***
Market friction	8	Lag Return	33.7%***
	9	Amihud Illiquidity	-2.4%
	10	Zero Return Proportion	0.9%
	11	Bid-Ask Spread	30.4%***
Others	12	Analyst forecast Dispersion	5.3%*
	13	Average Variance Beta	1.0%*
	14	SUE	10.9%***

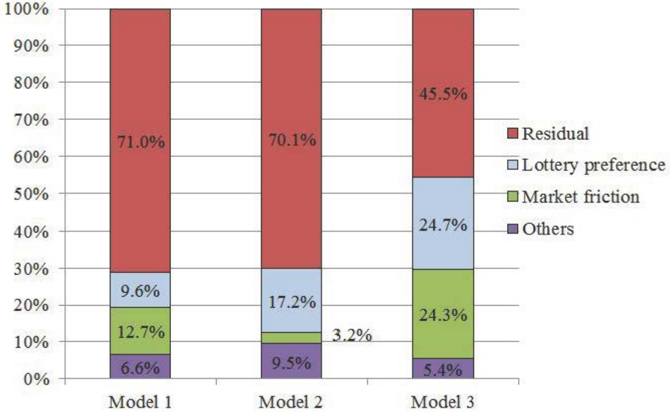
- Many variables explain less than 10% of the puzzle (from Table 3).

All candidates in multivariate setting

Variable	Model 1			Model 2			Model 3		
	Coeff.	Fraction	t-stat	Coeff.	Fraction	t-stat	Coeff.	Fraction	t-stat
Skew	-0.450	2.4%	(1.51)	-0.432	3.0%	(1.56)	-1.246	6.5%***	(6.35)
Coskew	-0.520	2.8%	(0.99)	-0.505	3.5%	(0.73)	-0.593	3.1%***	(2.95)
E(IdioSkew)	-0.772	4.2%**	(2.13)	-1.516	10.7%**	(1.98)	-2.874	15.1%***	(6.24)
RTP	-0.043	0.2%	(0.08)						
Lagret	-1.050	5.7%	(1.03)	-0.072	0.5%	(0.07)	-4.085	21.5%***	(5.74)
Amihud	0.351	-1.9%	(-0.69)	-0.531	3.7%	(0.69)	-0.726	3.8%	(1.60)
Zeroret	-0.248	1.3%	(0.28)	0.136	-1.0%	(-0.47)	0.186	-1.0%	(-1.02)
Spread	-1.412	7.6%	(0.52)						
Dispersion	-0.640	3.4%***	(2.66)	-0.793	5.6%***	(3.22)			
AvgVar β	-0.150	0.8%	(0.81)	0.032	-0.2%	(-0.12)	-0.060	0.3%	(0.67)
SUE	-0.448	2.4%***	(2.76)	-0.579	4.1%***	(3.12)	-0.973	5.1%***	(7.58)
Residual	-13.178	71.0%***	(5.86)	-9.972	70.1%***	(6.56)	-8.657	45.5%***	(10.06)
Total	-18.560***	100%	(-3.17)	-14.231***	100%	(-3.49)	-19.028***	100%	(-8.89)
Sample	1984 to 2001			1982 to 2012			1971 to 2012		
Avg # firms/mth	1524.4			1806.0			2752.4		

- Lottery and friction variables dominate other explanations (from Table 5).

Fig 1A: Summary of explained fraction



- All existing explanations explain 30-55%. Lottery-preference and market friction-based stories are the most successful.
- We can plot such pie charts because the contributions add up to 100%. Can't be done with conventional approach.

Flexibility of our decomposition

1. Portfolios
   - Can be applied to cross-sectional regressions on portfolios sorted by IVOL (portfolios help reduce measurement error which causes downward bias in fraction explained).
2. Non-linear specifications.
   - Replace continuous IVOL with a dummy variable indicating high IVOL, and/or replace candidate with dummy variable.
   - We show non-linear specifications produce similar set of best candidates.
3. Decompose other anomalies.
   - We can flip the analysis to see how much of other anomalies (e.g. Maxret, SUE) are explained by IVOL.
   - Our method can be easily applied to other anomalies.

Conclusion

- We survey explanations for the IVOL puzzle and propose a simple methodology to quantify the success of each explanation.
- We find that most explanations explain $<10\%$ of the puzzle.
- The most promising explanations are lottery preference and market friction explanations.
- Across various specifications, the residual part of the IVOL puzzle that remains unexplained by the best candidates is statistically significant.
- Our simple methodology can be used to compare competing explanations for other anomalies.