

# Similar Stocks

Wei He    Yuehan Wang    Jianfeng Yu

*PBC School of Finance, Tsinghua University*

# Disclaimer

---

This presentation was produced solely by Jianfeng Yu. The opinions and statements expressed herein are those of Jianfeng Yu and are not necessarily the opinions of any other entity, including UBS AG and its affiliates. UBS AG and its affiliates accept no responsibility whatsoever for the accuracy, reliability or completeness of the information, statements or opinions contained in this presentation and will not be liable either directly or indirectly for any consequences, including any loss or damage, arising out of the use of or reliance on this presentation or any part thereof.

Reproduced with permission.

# Motivation

- ▶ **Limited Attention & Trend Chasing:** Observing a stock with high recent return, investors, especially retail investors, who do not previously own this stock, might seek to buy similar stocks, whose past returns may not be as high, to chase the trend.
- ▶ **Preference Bias:** After experiencing high return in their stock investment, investors could pay more attention to stocks with similar characteristics, because of their preference for stocks with certain characteristics and different investors have different preferences.
- ▶ **Belief Bias:** Investors could believe that firms with similar characteristics should earn similar returns on average, or simply use past realized returns of similar stocks as the approximation for future return distribution of the focal stock.
- ▶ **Experience Effect:** Investors have positive (negative) prior investment experience in some stocks, they are more (less) likely to buy new stocks with similar characteristics.
- ▶ All the above forces imply that firms with high similar-stock returns should experience high subsequent returns.

# Main Findings

- ▶ **Similarity** between two stocks is measured by the distance between their characteristics.
- ▶ After a stock's most similar stocks have experienced high (low) returns in the past month, this focal stock tends to earn an abnormally high (low) return this month.
- ▶ The long-short portfolio strategy sorted on similar-stocks' past average return earns a monthly CAPM alpha of 1.25% and Fama-French 6-factor alpha of 0.88%.
- ▶ This similarity effect is robust after controlling for style investing and a wide range of well-known firm-level characteristics.

## Related Literature

- ▶ Huang (2019) on investor trading behavior: After experiencing positive returns from their stock holdings, investors are more likely to buy other stocks in the same industry or with similar size and BM ratio.
  - ▶ we focus on the asset pricing implications of similarity investing.
- ▶ Style investing and categorical thinking
  - ▶ Barberis and Sheifer (2003), Barberis, Sheifer and Wurgler(2005): category thinking
  - ▶ Green and Hwang (2011): price
  - ▶ Woo (2004), Froot and Teo (2008): institutional investor's style investing behavior
  - ▶ Wahal and Yavuz (2013), Lewellen (2002) and Teo and Woo (2004): predictive power for individual stock returns

## Related Literature

- ▶ Economically linked firms
  - ▶ Cohen and Frazzini (2008), Menzly and Ozbas (2010): supply chain
  - ▶ Cohen and Lou (2012): same industry
  - ▶ Hoberg and Phillips (2018): same product markets
  - ▶ Lee, Sun, Wang, and Zhang (2019): technologies
  - ▶ Parsons, Sabbatucci, and Titman (2020): geographic links
- ▶ Investor attention and positive feedback trading
  - ▶ Pearson, Yang, and Zhang (2020), Ben-David, Birru, and Prokopenya (2018), Anagol, Balasubramaniam, and Ramadorai (2020), Gao, Shi, and Zhao (2019): positive investment experience and trading behavior.

# SIM

- ▶ At the end of each month  $t$ , we standardize all  $n$  characteristics (using price, size, book-to-market, operating profitability and investment as benchmark).
- ▶ In this  $n$ -dimensional space, for each stock, we compute the Euclidean distance to all other stocks.
- ▶ For example, five characteristics of stock  $i$  are  $(Prc_i, SIZE_i, BM_i, OP_i, INV_i)$ , then the distance between stock  $i$  and stock  $j$  is:

$$D_{i,j} = \sqrt{(Prc_i - Prc_j)^2 + (SIZE_i - SIZE_j)^2 + (BM_i - BM_j)^2 + (OP_i - OP_j)^2 + (INV_i - INV_j)^2} \quad (1)$$

- ▶ **SIM** for stock  $i$  is the value-weighted average excess return of its 50 nearest stocks in month  $t$ .

# Main results

## ► Equally weighted results

Panel A	Equally weighted										
	Low	2	3	4	5	6	7	8	9	High	High-Low
Excess return	0.152 (0.5)	0.447 (1.69)	0.53 (2.1)	0.7 (2.82)	0.796 (3.32)	0.838 (3.55)	0.901 (3.91)	1.009 (4.27)	1.092 (4.52)	1.303 (4.88)	1.151 (6.13)
CAPM Alpha	-0.583 (-3.49)	-0.228 (-1.7)	-0.119 (-0.95)	0.061 (0.53)	0.187 (1.61)	0.241 (2.2)	0.316 (2.94)	0.424 (3.73)	0.527 (4.04)	0.731 (4.57)	1.314 (7.19)
FF3 Alpha	-0.667 (-5.34)	-0.355 (-3.99)	-0.265 (-3.55)	-0.087 (-1.43)	0.035 (0.62)	0.083 (1.63)	0.174 (3.48)	0.27 (4.52)	0.376 (4.58)	0.585 (5.18)	1.252 (6.66)
FF3+Mom Alpha	-0.34 (-2.39)	-0.124 (-1.2)	-0.093 (-1.18)	0.041 (0.56)	0.151 (2.42)	0.181 (3.13)	0.217 (4.17)	0.315 (4.92)	0.396 (4.72)	0.608 (4.96)	0.948 (4.4)
FF5 Alpha	-0.517 (-3.19)	-0.293 (-2.45)	-0.248 (-2.7)	-0.056 (-0.68)	0.047 (0.69)	0.094 (1.68)	0.174 (3.49)	0.28 (4.55)	0.386 (4.38)	0.609 (4.94)	1.126 (4.59)
FF6 Alpha	-0.252 (-1.68)	-0.101 (-0.88)	-0.102 (-1.2)	0.051 (0.59)	0.144 (2.16)	0.177 (2.93)	0.212 (4.08)	0.318 (4.83)	0.404 (4.52)	0.628 (4.89)	0.881 (3.72)



# Main Results: Cont'd

## ► Value-weighted results

Panel B	Value-weighted										
	Low	2	3	4	5	6	7	8	9	High	High-Low
Excess return	-0.003 (-0.01)	0.378 (1.6)	0.481 (2.26)	0.554 (2.72)	0.694 (3.45)	0.763 (3.86)	0.759 (3.73)	0.852 (4.21)	0.976 (4.39)	1.088 (4.58)	1.09 (5.19)
CAPM Alpha	-0.722 (-4.8)	-0.273 (-2.21)	-0.128 (-1.43)	-0.044 (-0.54)	0.11 (1.42)	0.202 (2.78)	0.194 (2.21)	0.296 (3.41)	0.419 (3.56)	0.529 (3.92)	1.252 (6.07)
FF3 Alpha	-0.719 (-5.41)	-0.291 (-2.42)	-0.179 (-2.09)	-0.071 (-0.95)	0.074 (1.05)	0.19 (2.88)	0.219 (2.55)	0.282 (3.32)	0.413 (3.36)	0.505 (4.05)	1.224 (5.79)
FF3+Mom Alpha	-0.497 (-3.34)	-0.119 (-0.75)	-0.08 (-0.93)	0.008 (0.09)	0.093 (1.27)	0.147 (2.2)	0.144 (1.96)	0.214 (2.85)	0.329 (3.2)	0.398 (3.37)	0.895 (3.95)
FF5 Alpha	-0.537 (-3.61)	-0.206 (-1.31)	-0.208 (-2.14)	-0.041 (-0.46)	0.082 (1.16)	0.227 (3.36)	0.327 (2.92)	0.334 (3.22)	0.483 (2.79)	0.582 (3.6)	1.118 (4.05)
FF6 Alpha	-0.365 (-2.55)	-0.068 (-0.38)	-0.12 (-1.33)	0.023 (0.22)	0.097 (1.33)	0.187 (2.79)	0.251 (2.76)	0.271 (3.02)	0.406 (2.97)	0.484 (3.54)	0.849 (3.49)
HXZ Q4 Alpha	-0.325 (-1.69)	-0.004 (-0.02)	-0.129 (-1.17)	0.036 (0.29)	0.092 (1.14)	0.216 (2.73)	0.328 (2.82)	0.378 (3.1)	0.524 (2.87)	0.644 (3.43)	0.969 (3.04)

# Main Results: Cont'd

## ► Based annual SIM

Panel C	Based annual SIM										
	Low	2	3	4	5	6	7	8	9	High	High-Low
Excess return	0.282 (0.97)	0.405 (1.66)	0.519 (2.33)	0.438 (2.17)	0.624 (3.24)	0.679 (3.46)	0.71 (3.78)	0.801 (4.07)	0.908 (4.4)	1.046 (4.25)	0.764 (3.13)
CAPM Alpha	-0.432 (-2.59)	-0.246 (-1.97)	-0.101 (-0.92)	-0.138 (-1.43)	0.067 (0.83)	0.109 (1.38)	0.159 (2.01)	0.252 (2.83)	0.356 (3.32)	0.445 (3.12)	0.876 (3.64)
FF3 Alpha	-0.491 (-3.06)	-0.295 (-2.36)	-0.14 (-1.35)	-0.19 (-1.91)	0.01 (0.13)	0.069 (0.94)	0.127 (1.72)	0.214 (2.79)	0.336 (3.58)	0.472 (3.91)	0.963 (4)
FF3+Mom Alpha	0.16 (1.34)	0.172 (1.44)	0.163 (1.7)	-0.039 (-0.42)	0.08 (0.97)	0.032 (0.41)	0.021 (0.26)	0.045 (0.56)	0.069 (0.82)	0.124 (1.15)	-0.037 (-0.21)
FF5 Alpha	-0.333 (-1.78)	-0.241 (-1.55)	-0.155 (-1.31)	-0.235 (-2.17)	-0.075 (-0.87)	-0.021 (-0.27)	0.032 (0.42)	0.091 (1.09)	0.31 (2.72)	0.522 (3.92)	0.855 (2.96)
FF6 Alpha	0.209 (1.87)	0.153 (1.25)	0.106 (1.09)	-0.101 (-1.05)	-0.006 (-0.07)	-0.043 (-0.54)	-0.049 (-0.62)	-0.04 (-0.49)	0.085 (0.97)	0.219 (2.07)	0.011 (0.06)

# Single/multi-variable distance

	Low	2	3	4	5	6	7	8	9	High	High-Low
Price	-0.155 (-0.73)	-0.133 (-1.14)	-0.229 (-1.98)	-0.134 (-1.56)	0.019 (0.22)	0.162 (1.45)	0.113 (1.69)	0.075 (0.97)	0.286 (4.09)	0.427 (4.59)	0.582 (2.27)
BM	-0.105 (-1.35)	-0.106 (-1.53)	-0.073 (-1.14)	-0.049 (-0.89)	0.136 (1.9)	0.046 (0.79)	0.005 (0.07)	0.104 (1.48)	0.088 (1.26)	0.242 (3.38)	0.347 (3.03)
Size	-0.185 (-1.76)	0.012 (0.13)	-0.149 (-1.9)	0.071 (0.97)	0.032 (0.49)	0.096 (1.23)	0.186 (2.12)	0.209 (2.64)	0.224 (2.69)	0.371 (3.51)	0.556 (3.14)
Profitability	-0.185 (-1.92)	-0.029 (-0.33)	-0.001 (-0.02)	0.001 (0.01)	-0.007 (-0.13)	0.055 (0.85)	0.093 (1.48)	0.056 (0.92)	0.139 (2.19)	0.33 (3.31)	0.515 (3.38)
Investment	-0.156 (-1.94)	-0.126 (-2)	0.018 (0.29)	-0.004 (-0.06)	0.069 (0.98)	0.093 (1.47)	0.103 (1.48)	0.105 (1.48)	0.046 (0.66)	0.242 (3.12)	0.398 (3.35)
Without price	-0.256 (-1.92)	-0.161 (-1.61)	-0.193 (-2.29)	0.111 (1.48)	-0.007 (-0.1)	0.194 (2.89)	0.245 (2.82)	0.279 (3.11)	0.422 (3.63)	0.558 (4.24)	0.814 (3.61)
BM/Size	-0.4 (-3.44)	-0.319 (-3.45)	-0.107 (-1.17)	0.126 (1.6)	0.215 (3.06)	0.214 (2.84)	0.245 (3.02)	0.303 (3.17)	0.352 (3.41)	0.516 (4.71)	0.916 (4.72)
Same industry	-0.516 (-4.04)	-0.187 (-1.73)	-0.243 (-2.39)	0.065 (0.57)	0.048 (0.56)	0.087 (1.04)	0.175 (2.21)	0.25 (2.76)	0.378 (4.02)	0.378 (3.04)	0.895 (4.3)
Diff industry	-0.166 (-1)	0.038 (0.24)	0.005 (0.06)	0.08 (1.03)	0.096 (1.29)	0.226 (3.14)	0.11 (1.34)	0.268 (3.02)	0.273 (2.34)	0.457 (3.52)	0.623 (2.34)

# Orthogonal SIM

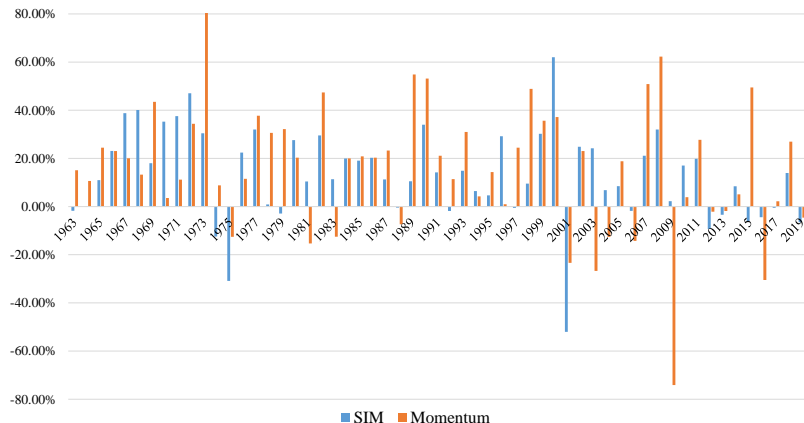
- ▶ We regress the focal stock's similar-stock average returns against the focal stock's own past-month return. Here shows the return alphas of portfolios sorted on regression residual.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Excess return	-0.103 (-0.4)	0.214 (0.93)	0.328 (1.55)	0.633 (3.2)	0.62 (3.12)	0.8 (4.28)	0.676 (3.47)	0.83 (4)	1.019 (4.8)	1.209 (5.15)	1.311 (7.01)
CAPM Alpha	-0.799 (-5.96)	-0.424 (-3.99)	-0.277 (-3.33)	0.061 (0.77)	0.046 (0.56)	0.248 (3.72)	0.134 (1.9)	0.268 (2.81)	0.466 (4.64)	0.638 (4.98)	1.437 (7.75)
FF3 Alpha	-0.794 (-6.6)	-0.432 (-4.55)	-0.305 (-3.8)	0.035 (0.48)	0.013 (0.18)	0.223 (3.38)	0.133 (2.04)	0.279 (2.97)	0.454 (4.54)	0.611 (5.31)	1.405 (7.29)
FF3+Mom Alpha	-0.58 (-4.22)	-0.357 (-3.35)	-0.228 (-2.92)	0.077 (0.96)	0.064 (0.9)	0.195 (3)	0.086 (1.29)	0.217 (2.92)	0.377 (3.73)	0.546 (4.93)	1.127 (5.44)
FF5 Alpha	-0.655 (-4.55)	-0.366 (-3.67)	-0.315 (-3.43)	0.024 (0.32)	0.012 (0.17)	0.237 (3.5)	0.178 (2.4)	0.354 (2.98)	0.508 (3.77)	0.666 (4.88)	1.321 (5.4)
FF6 Alpha	-0.487 (-3.45)	-0.309 (-2.96)	-0.248 (-2.95)	0.061 (0.75)	0.057 (0.78)	0.211 (3.15)	0.133 (1.85)	0.295 (3.08)	0.438 (3.64)	0.605 (4.96)	1.092 (4.89)
HXZ Q4 Alpha	-0.446 (-2.33)	-0.296 (-2.49)	-0.259 (-2.41)	0.081 (0.91)	0.043 (0.51)	0.271 (3.42)	0.183 (2.23)	0.4 (3.17)	0.54 (3.56)	0.758 (4.82)	1.188 (4.17)
SY4 Alpha	-0.414 (-2.61)	-0.273 (-2.33)	-0.232 (-2.56)	0.083 (0.84)	0.077 (0.87)	0.21 (2.86)	0.049 (0.66)	0.236 (2.74)	0.35 (3.16)	0.596 (5.09)	1.01 (4.53)
DHS3 Alpha	-0.033 (-0.16)	0.085 (0.57)	0.061 (0.5)	0.309 (2.61)	0.178 (1.74)	0.3 (3.09)	0.121 (1.3)	0.281 (2.3)	0.407 (2.5)	0.67 (3.82)	0.667 (2.32)

# Annual Performance

This figure plots the annual cumulative raw return spreads of long-short portfolio based on SIM and MOM from 1963 to 2019.

## Annual performance over 1963-2019



# Confounding Effects

- ▶ Momentum effects and other correlated anomalies
- ▶ Industry momentum effects
  - ▶ Grinblatt and Moskowitz (1999)
- ▶ Style investing effects
  - ▶ Barberis and Shleifer (2003), Teo and Woo,(2004), Froot and Teo, (2009), Wahal and Yavuz (2013))

# Double Sorting

	Size	BM	Profitability	Investment	R1m	R12m	Lt rev	IVOL	MAX	Skewness
High	0.405 (3.64)	0.35 (2.56)	0.333 (2.95)	0.318 (2.27)	0.539 (4.85)	-0.119 (-1.12)	0.271 (2.19)	0.193 (2.8)	0.344 (4.22)	0.292 (3.2)
4	0.201 (3.13)	0.168 (1.92)	0.218 (2.53)	0.166 (1.76)	0.272 (3.86)	-0.102 (-1.27)	0.163 (1.95)	0.111 (2.18)	0.215 (3.88)	0.187 (2.75)
3	0.029 (0.51)	0.046 (0.95)	0.118 (1.91)	0.033 (0.65)	0.166 (2.72)	-0.164 (-1.93)	0.035 (0.59)	0.078 (1.57)	0.108 (2.26)	0.053 (1.17)
2	-0.099 (-1.42)	-0.172 (-2.36)	-0.11 (-1.39)	-0.153 (-2.02)	-0.075 (-0.99)	-0.339 (-3.41)	-0.109 (-1.4)	-0.093 (-1.61)	-0.113 (-1.94)	-0.133 (-2.24)
Low	-0.384 (-3.6)	-0.363 (-3.14)	-0.464 (-4.18)	-0.254 (-2.09)	-0.31 (-2.66)	-0.515 (-3.94)	-0.259 (-2.17)	-0.308 (-3.38)	-0.201 (-1.69)	-0.242 (-2.47)
High-Low	0.79 (4.27)	0.713 (3.12)	0.798 (4.05)	0.572 (2.4)	0.849 (4.38)	0.397 (2.19)	0.53 (2.53)	0.501 (3.66)	0.546 (3.02)	0.534 (3.22)

# Fama-Macbeth Regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SIM	0.331 (5.41)	0.238 (4.56)	0.283 (5.02)	0.139 (4.42)	0.121 (4.08)	0.115 (3.80)	0.101 (3.53)
Style R1m		-0.003 (-0.10)			0.029 (1.28)		0.012 (0.52)
Style R6m		0.006 (0.11)			-0.017 (-0.42)		-0.023 (-0.59)
Style R1y		0.278 (5.00)			0.096 (2.32)		0.095 (2.41)
Industry R1m			0.236 (6.53)			0.290 (8.55)	0.288 (8.54)
Industry R6m			0.105 (1.76)			0.107 (2.04)	0.110 (2.14)
Industry R1y			0.153 (2.55)			0.096 (1.88)	0.094 (1.88)
Size				-0.209 (-4.41)	-0.243 (-4.80)	-0.193 (-4.15)	-0.227 (-4.55)
BM				0.114 (3.64)	0.049 (1.68)	0.129 (4.38)	0.063 (2.24)
profitability				0.077 (3.82)	0.078 (3.86)	0.077 (3.86)	0.077 (3.89)
Investment				-0.163 (-7.44)	-0.155 (-7.25)	-0.168 (-7.76)	-0.161 (-7.59)
R1m				-0.591 (-12.75)	-0.596 (-12.90)	-0.652 (-14.46)	-0.656 (-14.56)
R12m				0.225 (4.29)	0.223 (4.31)	0.188 (3.82)	0.187 (3.83)
IVOL				-0.258 (-3.33)	-0.242 (-3.18)	-0.270 (-3.66)	-0.255 (-3.50)
Amihud				0.046 (1.23)	0.041 (1.12)	0.047 (1.26)	0.042 (1.16)
Avg. Obs.	2088	2088	2088	2081	2081	2081	2081
R <sup>2</sup> Avg.(%)	1.113	2.217	2.531	5.846	6.149	6.801	7.073



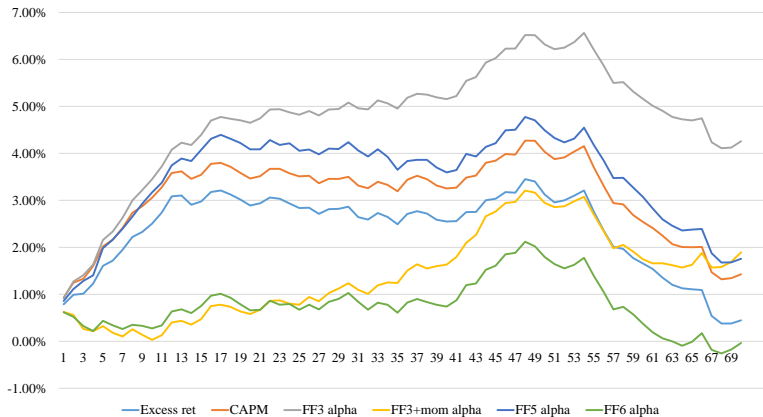
# Long-Horizon Performance

Panel A	SIM						
	1 month	3 months	6 months	1 year	3 years	5 years	6 years
Excess return	0.787 (4.29)	0.339 (2.65)	0.287 (2.98)	0.257 (3.41)	0.084 (1.71)	0.039 (1)	0.02 (0.57)
CAPM Alpha	0.927 (4.97)	0.443 (3.5)	0.358 (3.76)	0.294 (3.81)	0.094 (1.88)	0.043 (1.07)	0.023 (0.64)
FF3 Alpha	0.921 (4.63)	0.468 (3.3)	0.386 (3.81)	0.334 (4.28)	0.14 (2.91)	0.084 (2.33)	0.061 (1.84)
FF3+Mom Alpha	0.628 (2.85)	0.089 (0.66)	0.031 (0.32)	0.034 (0.5)	0.042 (0.89)	0.029 (0.82)	0.031 (0.94)
FF5 Alpha	0.86 (3.1)	0.428 (2.1)	0.358 (2.69)	0.307 (3.29)	0.105 (2.19)	0.052 (1.43)	0.027 (0.84)
FF6 Alpha	0.618 (2.44)	0.109 (0.67)	0.057 (0.57)	0.053 (0.79)	0.023 (0.5)	0.007 (0.21)	0.005 (0.15)

Panel B	Momentum						
	1 month	3 months	6 months	1 year	3 years	5 years	6 years
Excess return	1.253 (4.65)	1.132 (4.5)	0.94 (4.07)	0.578 (2.91)	0.124 (1.03)	-0.005 (-0.06)	-0.018 (-0.22)
CAPM Alpha	1.407 (5.7)	1.249 (5.32)	1.035 (4.7)	0.635 (3.22)	0.142 (1.18)	0.001 (0.01)	-0.011 (-0.13)
FF3 Alpha	1.63 (6.74)	1.486 (6.45)	1.314 (6.32)	0.951 (5.6)	0.388 (4.07)	0.2 (3.02)	0.17 (3.01)
FF3+Mom Alpha	0.339 (3.13)	0.289 (3.07)	0.3 (2.57)	0.256 (2.04)	0.149 (1.65)	0.064 (0.99)	0.078 (1.37)
FF5 Alpha	1.374 (4.4)	1.284 (4.49)	1.167 (4.69)	0.864 (4.71)	0.333 (3.59)	0.159 (2.53)	0.137 (2.56)
FF6 Alpha	0.295 (2.62)	0.28 (2.89)	0.313 (2.56)	0.276 (2.07)	0.132 (1.49)	0.046 (0.78)	0.06 (1.15)

# Long-Horizon Performance

## Long-Horizon Performance



## Returns on Earnings Announcement Days

This table reports regressions of daily firm-level stock return  $DRET_{i,d}$  for firm  $i$  on day  $d$  in month  $t$  onto firm  $i$ 's SIM in month  $t - 1$ ,  $SIM_{i,t-1}$ , earnings announcement date dummy variable (Eday) on firm  $i$  on day  $d$  in month  $t$ , and their interaction term. An earnings announcement window is defined as the one-day (Columns 1 and 2) or three-day window (Columns 3 and 4) centered on an earnings announcement date. Eday is a dummy variable which equals one if day  $d$  in month  $t$  is in this window.

	(1)	(2)	(3)	(4)
	One-day window		Three-day window	
	Dret(%)	Dret(%)	Dret(%)	Dret(%)
SIM	0.326 (18.27)	0.445 (25.10)	0.327 (18.21)	0.443 (24.84)
SIM*1(Eday)	1.312 (16.05)	1.393 (17.16)	0.433 (8.90)	0.527 (10.91)
1(Eday)	0.207 (37.95)	0.225 (41.42)	0.104 (31.76)	0.120 (37.02)
Lagged Controls	No	Yes	No	Yes
Day FE	Yes	Yes	Yes	Yes
$R^2$	0.062	0.062	0.075	0.075
Obs.	33463463	33463463	33445521	33445521

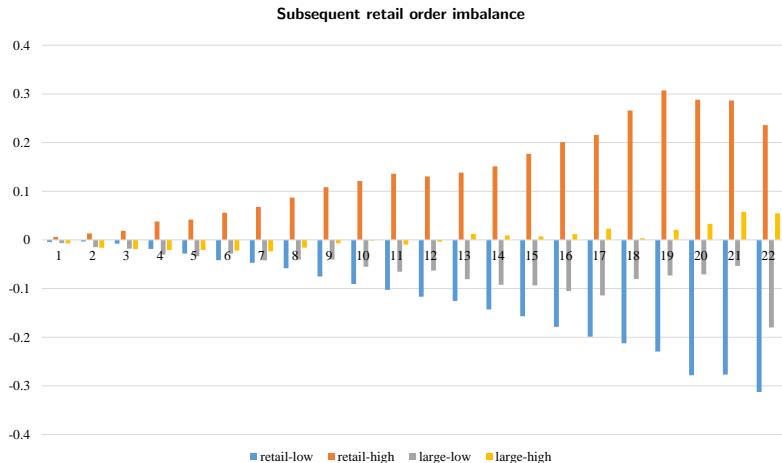
## Future Earnings Surprise

Forecasting regressions of next-quarter's standardized unexpected earnings (SUE) on SIM.

	(1) $SUE_t$	(2) $SUE_t$	(3) $SUE_t$
$SIM_{t-1}$	0.03556 (2.21)	0.03764 (1.68)	0.04368 (1.82)
$SUE_{t-1}$		0.03807 (1.74)	0.06385 (1.58)
$SUE_{t-2}$		0.00676 (0.53)	0.01910 (1.12)
$SUE_{t-3}$		-0.00817 (-0.90)	0.00500 (0.74)
$SUE_{t-4}$		-0.07389 (-1.60)	-0.00001 (-0.92)
Constant	-0.00122 (-58.49)	-0.0029 (-18.97)	-0.00285 (-23.71)
Firm FE	YES	YES	NO
Industry FE	NO	NO	YES
Year-quarter FE	YES	YES	YES
Observations	208,079	124,570	122,042
Adjusted $R^2$	0.0737	0.0570	0.0201

# Subsequent Retail Order Imbalance

At the end of each month  $t$ , we form 5 portfolios based on SIM. For each stock, we measure the change in the retail (large) order imbalance by taking the difference between daily imbalance in month  $t + 1$  and the average daily imbalance of the six-months window from month  $t - 7$  to month  $t - 2$ . Then, we calculate the cumulative retail order imbalance. To better demonstrate the results, we minus the average order imbalance from 5 groups' order imbalance in each trading day to detrend.



# Further Robustness Test

Panel A	Robusness check										
	Low	2	3	4	5	6	7	8	9	High	High-Low
Near 51-100	-0.355 (-2.6)	-0.226 (-2.18)	-0.044 (-0.46)	0.032 (0.36)	0.178 (2.51)	0.155 (2.2)	0.301 (3.71)	0.269 (2.5)	0.251 (2.55)	0.47 (3.57)	0.825 (3.52)
Near 101-150	-0.395 (-2.94)	-0.091 (-0.82)	-0.047 (-0.54)	0.112 (1.33)	0.044 (0.56)	0.113 (1.76)	0.205 (2.33)	0.272 (2.97)	0.208 (2.1)	0.526 (4.49)	0.92 (4.46)
Near 151-200	-0.242 (-1.88)	-0.062 (-0.54)	-0.021 (-0.27)	0.08 (0.95)	-0.067 (-0.91)	0.17 (2.44)	0.15 (2.21)	0.207 (2.93)	0.325 (3.65)	0.338 (2.54)	0.58 (2.62)
Near 201-250	-0.145 (-1.36)	-0.164 (-1.79)	-0.182 (-2)	0.11 (1.51)	0.169 (2.23)	0.127 (1.61)	0.104 (1.4)	0.034 (0.43)	0.178 (1.82)	0.264 (2.29)	0.408 (2.27)
Random matched	0.051 (0.83)	0.079 (1.27)	0.015 (0.21)	0.016 (0.23)	0.01 (0.15)	0.08 (1.36)	0.147 (2.08)	-0.103 (-1.79)	0.053 (0.81)	0.041 (0.7)	-0.01 (-0.11)

# Subsamples

Panel B		Two subperiods										
		Low	2	3	4	5	6	7	8	9	High	High-Low
1963-1990	CAPM Alpha	-0.691 (-3.82)	-0.358 (-2.35)	-0.106 (-0.91)	0.106 (0.99)	0.232 (2.12)	0.369 (3.74)	0.278 (2.75)	0.337 (3.09)	0.585 (4.61)	0.734 (4.34)	1.425 (5.54)
	FF3 Alpha	-0.757 (-4.43)	-0.437 (-3.18)	-0.183 (-1.56)	-0.011 (-0.1)	0.134 (1.28)	0.256 (2.8)	0.195 (2.07)	0.216 (2.26)	0.491 (4.33)	0.6 (4.54)	1.357 (5.28)
	FF6 Alpha	-0.374 (-1.94)	-0.158 (-0.98)	0.021 (0.15)	0.025 (0.2)	0.152 (1.29)	0.231 (2.36)	0.13 (1.28)	0.227 (2.15)	0.421 (2.81)	0.536 (3.35)	0.911 (2.92)
1990-2019	CAPM Alpha	-0.777 (-3.3)	-0.212 (-1.12)	-0.153 (-1.15)	-0.189 (-1.6)	0.001 (0.01)	0.045 (0.43)	0.115 (0.82)	0.253 (1.93)	0.264 (1.4)	0.343 (1.69)	1.12 (3.56)
	FF3 Alpha	-0.734 (-3.74)	-0.193 (-1.07)	-0.173 (-1.37)	-0.174 (-1.67)	-0.001 (-0.01)	0.072 (0.77)	0.175 (1.29)	0.284 (2.23)	0.3 (1.55)	0.379 (1.95)	1.113 (3.44)
	FF6 Alpha	-0.361 (-1.72)	-0.01 (-0.04)	-0.219 (-1.76)	-0.126 (-0.86)	0.01 (0.1)	0.045 (0.48)	0.254 (1.71)	0.251 (1.78)	0.327 (1.43)	0.406 (1.83)	0.767 (2.05)
Panel C		Large/small subsamples										
		Low	2	3	4	5	6	7	8	9	High	High-Low
Small Size: NYSE p50	CAPM Alpha	-0.646 (-3.62)	-0.157 (-1.07)	-0.037 (-0.27)	0.004 (0.03)	0.131 (1.05)	0.271 (2.19)	0.39 (3.06)	0.408 (3.04)	0.554 (3.73)	0.675 (4.12)	1.321 (6.84)
	FF3 Alpha	-0.736 (-5.51)	-0.289 (-3.2)	-0.18 (-2.18)	-0.159 (-2.12)	-0.018 (-0.26)	0.11 (1.61)	0.24 (3.49)	0.267 (3.16)	0.423 (4.21)	0.536 (4.45)	1.272 (6.2)
	FF6 Alpha	-0.339 (-2.09)	-0.023 (-0.24)	-0.011 (-0.11)	-0.025 (-0.23)	0.119 (1.33)	0.192 (2.33)	0.34 (4.46)	0.321 (3.67)	0.454 (4.33)	0.583 (4.5)	0.922 (3.77)
Large Size: NYSE p50	CAPM Alpha	-0.383 (-3.51)	-0.33 (-4.31)	-0.131 (-1.78)	-0.056 (-0.98)	0.092 (1.55)	0.149 (2.43)	0.091 (1.32)	0.181 (2.65)	0.242 (3.05)	0.294 (3.19)	0.677 (4.2)
	FF3 Alpha	-0.36 (-3.25)	-0.312 (-3.84)	-0.112 (-1.53)	-0.06 (-1.02)	0.085 (1.43)	0.128 (2)	0.07 (1)	0.193 (2.8)	0.263 (3.37)	0.318 (3.53)	0.678 (4.08)
	FF6 Alpha	-0.144 (-1.29)	-0.28 (-3.09)	-0.07 (-0.86)	-0.091 (-1.37)	0.064 (1.05)	0.064 (0.99)	0.044 (0.6)	0.129 (1.59)	0.178 (2.23)	0.239 (2.48)	0.383 (2.24)

# Institutional Ownership

Panel A				Panel B			
CAPM Alpha				FF6 Alpha			
SIM	Low-IO	High-IO	High-Low	SIM	Low-IO	High-IO	High-Low
High	0.602 (3.45)	0.039 (0.33)	-0.563 (-3.85)	High	0.786 (3.93)	0.07 (0.52)	-0.716 (-4.74)
4	0.474 (2.75)	0.083 (1.1)	-0.391 (-2.16)	4	0.677 (3.74)	0.021 (0.27)	-0.657 (-3.52)
3	0.394 (2.74)	-0.049 (-0.65)	-0.443 (-2.98)	3	0.555 (3.47)	-0.038 (-0.52)	-0.593 (-3.4)
2	-0.01 (-0.08)	-0.17 (-1.77)	-0.159 (-1.14)	2	0.266 (1.5)	-0.141 (-1.33)	-0.407 (-1.91)
Low	-0.549 (-2.83)	-0.327 (-2.44)	0.222 (1.32)	Low	-0.012 (-0.06)	-0.159 (-1.17)	-0.146 (-0.77)
High-Low	1.15 (4.85)	0.366 (1.75)	-0.784 (-4.31)	High-Low	0.798 (2.43)	0.229 (0.97)	-0.569 (-2.33)



# International Evidence

Country	High turnover			Country	Low turnover		
	Low	High	High-Low		Low	High	High-Low
AUS	0.018 (0.16)	0.377 (3.57)	<b>0.36</b> (4.57)	ARG	2.212 (2.82)	1.638 (1.85)	-0.645 (-0.85)
BRA	-0.014 (-0.08)	1.341 (5.51)	<b>1.355</b> (5.25)	AUT	0.974 (2.04)	0.807 (1.52)	-0.163 (-0.31)
CAN	0.104 (0.88)	0.262 (1.88)	0.158 (1.29)	BEL	0.655 (1.57)	0.778 (1.69)	0.214 (0.5)
CHE	0.142 (1.69)	0.346 (4.33)	<b>0.204</b> (2.65)	CHL	1.178 (2.75)	0.664 (1.61)	-0.514 (-1.62)
CHN	0.038 (0.29)	0.466 (3.61)	<b>0.427</b> (4.87)	DNK	0.705 (1.55)	0.924 (2.15)	0.219 (0.51)
DEU	0.088 (1.1)	0.189 (2.05)	0.101 (1.01)	ESP	0.023 (0.05)	0.734 (1.69)	<b>0.823</b> (2.11)
FIN	0.219 (2.05)	0.4 (3.89)	<b>0.181</b> (1.77)	FRA	0.893 (1.92)	0.908 (2.27)	0.016 (0.04)
GBR	0.095 (1.01)	0.055 (0.62)	-0.041 (-0.49)	IDN	1.066 (1.65)	2.265 (3.07)	<b>1.199</b> (1.91)
GRC	-0.003 (-0.02)	0.069 (0.51)	0.072 (0.62)	ISR	1.106 (2.15)	1.263 (2.79)	-0.048 (-0.18)
HKG	-0.002 (-0.01)	0.239 (2.36)	<b>0.24</b> (2.33)	MEX	0.765 (1.63)	1.096 (2.03)	0.331 (0.69)
IND	0.196 (1.36)	0.431 (3.34)	<b>0.235</b> (2.03)	MYS	0.612 (1.43)	0.802 (1.97)	0.246 (0.77)
ITA	0.256 (2.54)	0.425 (4)	<b>0.169</b> (1.75)	NLD	0.413 (0.82)	0.95 (1.66)	0.565 (1)
JPN	0.182 (2.25)	0.103 (1.29)	-0.079 (-1.45)	NOR	0.495 (0.87)	0.791 (1.53)	0.296 (0.65)
KOR	-0.153 (-0.96)	0.127 (0.89)	<b>0.28</b> (1.95)	NZL	0.978 (0.95)	1.976 (3.15)	1.036 (0.95)
SGP	0.1 (0.95)	0.378 (2.5)	0.168 (1.54)	PAK	1.319 (2.29)	1.439 (2.4)	0.12 (0.24)
SWE	0.132 (1.16)	0.302 (2.82)	<b>0.17</b> (2.06)	PHL	0.857 (1.59)	0.706 (1.03)	-0.206 (-0.31)
THA	0.242 (2.27)	0.346 (3.33)	0.105 (1.19)	POL	0.612 (1.09)	0.973 (1.72)	0.36 (0.81)
TUR	0.147 (0.84)	0.281 (1.54)	0.134 (1.13)	PRT	1.004 (1.11)	0.342 (0.34)	0.473 (0.61)
ZAF	0.289 (2.37)	0.385 (3.18)	0.096 (0.93)	RUS	2.272 (1.97)	1.775 (1.81)	-1.018 (-0.93)
Comb	0.109 (1.57)	0.348 (4.96)	<b>0.234</b> (7.91)	Comb	0.916 (2.76)	1.097 (3.41)	0.16 (1.22)

# Conclusions

- ▶ Propose a measure of similarity between two stocks based on the distance of their corresponding characteristics
- ▶ Firms with high similar-stock returns tend to earn significantly higher future returns than firms with low similar-stock returns
- ▶ Similarity effect is not completely driven by style investing and other well-known anomalies
- ▶ Our similarity effect could be consistent with both the initial underreaction or continued overreaction