Online Appendix for "Search-Based Peer Firms: Aggregating Investor Perceptions through Internet Co-Searches"

Charles M.C. Lee, Paul Ma, and Charles C.Y. Wang

Table IA1.
Variable Description

This table reports the construction of variables used in our regressions. We use CRSP monthly stock returns and Compustat quarterly data for the sample period 2009–2011. CRSP variable names are in parentheses and Compustat variable names are in square brackets. After collecting the raw Compustat data, in accordance with Bhojraj, Lee, and Oler (2003), we drop all firm–quarter observations missing data on total assets [atq], total long term debt [dlttq], net income before extraordinary items [ibq], debt in current liabilities [lctq], or operating income after depreciation [oiadpq]. Further, we require the raw share price on the last day of each fiscal quarter to be greater than $3, both total common equity [ceqq] and total shareholder equity [seqq] to be positive, and net sales [saleq] to be more than $100 million.

Variable	Description	Calculation
returns	Monthly cum-dividend stock returns	(ret)
Valuation Multiples		
pb	Price-to-book ratio	Market cap / total common equity [ceqq]
evs	Enterprise value- to- sales ratio	(Market cap + long-term debt [dlttq]) / net sales [saleq]
pe	Price-to-earnings ratio	Market cap net income before extraordinary items [ibq]
Financial Statement Ratios		
rnoa	Return on net operating assets	Net operating income after depreciation [oiadpq] / (property, plant, and equipment [ppentq] + current assets [actq] - current liabilities [lctq])
roe	Return on equity	Net income before extraordinary items [ibq] / total common equity [ceqq]
at	(Inverse of) Asset turnover	Total assets [atq] / net sales [saleq]
pm	Profit margin	Net operating income after depreciation [oiadp] / net sales [saleq]
lev	Leverage	Long-term debt [dlttq] / total stockholder's equity [seqq]
Other Financial Information		
salesgrowth	One-year-ahead realized sales growth	(Net sale one year ahead in the future - current year net sales) / current year net sales [saleq]
rdpersales	R&D expense- to- sales ratio	R&D expense [xrdq] / net sales [saleq]
ltgrowth	Median analyst long-term growth forecast	
ltgrowth spread	Standard deviation in analyst long-term growth forecast	
eps spread	Standard deviation in analyst one-year-ahead EPS forecast	
coverage	Number of analysts covering firm	
size	Market capitalization	Price (prc) × shares outstanding (shrout)

Table IA2.
Comparison of R^2 Values: Monthly Returns

This table compares the average R^2 values from monthly cross-sectional regressions of the form

$$R_{i,t} = \alpha_t + \beta_t R_{p_i,t} + \epsilon_{i,t}$$

using CRSP returns data from January 2009 to December 2011. Columns (1)∼(3) report average R^2s from monthly cross-sectional regressions, regressing base firm (i) returns in a given month (t) on the concurrent returns of a portfolio (p_i) of 10 peers. Column (1) considers an equal-weighted portfolio of 10 randomly selected firms from the base firm's GICS6 industry; Column (2) considers an equal-weighted portfolio (SBP EW) of the top 10 SBP firms, ranked by the prior calendar year's *Annual Search Fraction* f_{ij}, defined as the percentage of daily unique users searching for firm j after searching for firm i on the same day; Column (3) considers a portfolio (SBP TW) consisting of the top 10 SBP firms, with each peer firm weighted by the prior calendar year's *Annual Search Fraction* (relative to the top 10 peer firms). SBPs and portfolio weights are generated based on prior calendar year's EDGAR search traffic (e.g., the regressions in 2009 are generated with weights from calendar year 2008). Columns (4) and (5) test for the significance of the differences in average R^2s between the two SBP portfolio formulations and the GICS6 peer portfolios. Column (6) tests for the significance of the differences in average R^2s between the equal-weighted and traffic-weighted SBP portfolios.

The results are reported under 4 major groupings depending on whether one restricts the SBP firms to be from the S&P1500 or the CRSP universe, and whether one restricts the SBP firms to also be within the same GICS2 sector as the base firm. Under each major grouping we report results for the S&P500 and S&P1500 base firms respectively. To facilitate comparisons, all the regressions are conducted using the same underlying set of firms. For example, when a base firm does not have at least 10 GICS peers but does have 10 peers based on the search traffic (and vice versa), we exclude the firm from the regression. When imposing the GICS2 restriction, we drop all firms not from the same GICS2 sector prior to forming our portfolio of 10 SBP firms. The variable N in parentheses represents the average cross-sectional sample size for each monthly regression and standard errors are reported in square brackets. Significance levels are indicated by *, **, *** for 10%, 5%, and 1%, respectively.

Table IA2.
(Continued)

	GICS6 (1)	SBP EW (2)	SBP TW (3)	(2)-(1) (4)	(3)-(1) (5)	(3)-(2) (6)
Random 10 SP1500 Peers						
SP500 Base Firms (N=407)	0.116*** [0.011]	0.189*** [0.013]	0.214*** [0.013]	0.073*** [0.007]	0.098*** [0.007]	0.024*** [0.005]
SP1500 Base Firms (N=1240)	0.069*** [0.007]	0.114*** [0.008]	0.125*** [0.010]	0.044*** [0.003]	0.055*** [0.005]	0.011*** [0.003]
Random 10 CRSP Peers						
SP500 Base Firms (N=455)	0.054*** [0.006]	0.205*** [0.012]	0.228*** [0.011]	0.151*** [0.008]	0.174*** [0.008]	0.023*** [0.006]
SP1500 Base Firms (N=1373)	0.037*** [0.004]	0.125*** [0.008]	0.131*** [0.009]	0.089*** [0.006]	0.094*** [0.006]	0.005 [0.003]
Random 10 SP1500 GICS2 Peers						
SP500 Base Firms (N=406)	0.116*** [0.011]	0.199*** [0.012]	0.228*** [0.014]	0.083*** [0.006]	0.111*** [0.007]	0.029*** [0.005]
SP1500 Base Firms (N=1239)	0.069*** [0.007]	0.123*** [0.008]	0.135*** [0.010]	0.053*** [0.003]	0.066*** [0.006]	0.012*** [0.003]
Random 10 CRSP GICS2 Peers						
SP500 Base Firms (N=455)	0.054*** [0.006]	0.215*** [0.012]	0.239*** [0.012]	0.161*** [0.008]	0.185*** [0.009]	0.024*** [0.006]
SP1500 Base Firms (N=1372)	0.037*** [0.004]	0.135*** [0.009]	0.141*** [0.009]	0.098*** [0.006]	0.104*** [0.007]	0.006** [0.003]
Number of Months	36	36	36	36	36	36

Table IA3.
Comparison of Slope Coefficient Values: Monthly Returns

This table compares the average slope coefficient (β) values from monthly cross-sectional regressions of the form

$$R_{i,t} = \alpha_t + \beta_t R_{p_i,t} + \epsilon_{i,t}$$

using CRSP returns data from January 2009 to December 2011. Columns (1)$\sim$(3) report average βs from monthly cross-sectional regressions, regressing base firm (i) returns in a given month (t) on the concurrent returns of a portfolio (p_i) of 10 peers. Column (1) considers an equal-weighted portfolio of 10 randomly selected firms from the base firm's GICS6 industry; Column (2) considers an equal-weighted portfolio (SBP EW) of the top 10 SBP firms, ranked by the prior calendar year's *Annual Search Fraction* f_{ij}, defined as the percentage of daily unique users searching for firm j after searching for firm i on the same day; Column (3) considers a portfolio (SBP TW) consisting of the top 10 SBP firms, with each peer firm weighted by the prior calendar year's *Annual Search Fraction* (relative to the top 10 peer firms). SBPs and portfolio weights are generated based on prior calendar year's EDGAR search traffic (e.g., the regressions in 2009 are generated with weights from calendar year 2008). Columns (4) and (5) test for the significance of the differences in average βs between the two SBP portfolio formulations and the GICS6 peer portfolios. Column (6) tests for the significance of the differences in average βs between the equal-weighted and traffic-weighted SBP portfolios.

The results are reported under 2 major groupings depending on whether one restricts the SBP firms to also be within the same GICS2 sector as the base firm. Under each major grouping we report results for the S&P500 and S&P1500 base firms respectively. To facilitate comparisons, all the regressions are conducted using the same underlying set of firms. For example, when a base firm does not have at least 10 GICS peers but does have 10 peers based on the search traffic (and vice versa), we exclude the firm from the regression. When imposing the GICS2 restriction, we drop all firms not from the same GICS2 sector prior to forming our portfolio of 10 SBP firms. The variable N in parentheses represents the average cross-sectional sample size for each monthly regression and standard errors are reported in square brackets. Significance levels are indicated by *, **, *** for 10%, 5%, and 1%, respectively.

Table IA3.
(Continued)

	GICS6 (1)	SBP EW (2)	SBP TW (3)	(2)-(1) (4)	(3)-(1) (5)	(3)-(2) (6)
Random 10 SP1500 Peers						
SP500 Unrestricted	0.627***	0.937***	0.816***	0.310***	0.189***	-0.121***
(N=407)	[0.024]	[0.021]	[0.021]	[0.024]	[0.024]	[0.024]
SP1500 Unrestricted	0.597***	0.861***	0.713***	0.263***	0.115***	-0.148***
(N=1240)	[0.019]	[0.019]	[0.023]	[0.015]	[0.019]	[0.017]
Random 10 SP1500 GICS2 Peers						
SP500 Restricted	0.628***	0.865***	0.771***	0.237***	0.143***	-0.094***
(N=406)	[0.024]	[0.016]	[0.019]	[0.021]	[0.023]	[0.022]
SP1500 Restricted	0.597***	0.823***	0.685***	0.225***	0.087***	-0.138***
(N=1239)	[0.020]	[0.017]	[0.021]	[0.016]	[0.019]	[0.016]
Number of Months	36	36	36	36	36	36

Table IA4.
Comparison of R^2 Values: Valuation Multiples, Financial Ratios, and Other Financial Information

This table compares the average R^2 from several monthly cross-sectional regressions of the form

$$Var_{i,t} = \alpha_t + \beta_t Var_{p_i,t} + \epsilon_{i,t}$$

using most recently observable quarterly financial statement data from Compustat and market capitalization data from CRSP on March, June, September, and December of each year from 2009 to 2011. Columns (1)∼(3) report average R^2s from quarterly cross-sectional regressions, regressing base firm (i) Var in a given month (t) on the concurrent Var of a portfolio (p_i) of 10 peers. Each row considers a different Var, as defined in Table 1. Column (1) considers an equal-weighted portfolio of 10 randomly selected firms from the base firm's GICS6 industry; Column (2) considers an equal-weighted portfolio of the top 10 SBP firms, ranked by the prior calendar year's *Annual Search Fraction* f_{ij}, defined as the percentage of unique users searching for firm j after searching for firm i on the same day; Column (3) considers a portfolio consisting of the top 10 SBP firms, with each peer firm weighted by the prior calendar year's *Annual Search Fraction* (relative to the top 10 SBP firms). SBP firms and portfolio weights are generated based on prior calendar year's EDGAR search traffic (e.g., the regressions in 2009 are generated with weights from calendar year 2008). Columns (4) and (5) test for the significance of the differences in average R^2s between the two SBP portfolio formulations and the GICS6 peer portfolios. Column (6) tests for the significance of the differences in average R^2s between the equal-weighted and traffic-weighted Search Traffic Peer portfolios.

Panels A and B (C and D) report regressions performed for the S&P500 (S&P1500) base firm sample. Whereas Panel A and C match each base firm to benchmark firms from the S&P1500 universe, Panels B and D restricts SBPs to also come from the same GICS2 sector. To facilitate comparisons, all the regressions are conducted using the same underlying set of base firms. For example, when a base firm does not have at least 10 GICS peers but does have 10 peers based on the search traffic (and vice versa), we exclude the firm from the regression. In addition to the data filters mentioned, for regressions involving *pe* we also drop observations with negative net income before extraordinary items and for regressions involving *rnoa* we drop observations when values are missing for current assets, current liabilities, or property, plant, and equipment. The variable N in parentheses represents the average cross-sectional sample size for each quarterly regression and standard errors are reported in square brackets. Significance levels are indicated by *, **, *** for 10%, 5%, and 1%, respectively.

Table IA4.
(Continued)

Panel A: S&P500 Base Firms

	GICS6 (1)	SBP EW (2)	SBP TW (3)	(2)-(1) (4)	(3)-(1) (5)	(3)-(2) (6)
Valuation Multiples						
pb (N=297)	0.048***	0.127***	0.121***	0.079***	0.072***	-0.007
	[0.004]	[0.010]	[0.008]	[0.007]	[0.006]	[0.005]
evs (N=298)	0.252***	0.428***	0.475***	0.176***	0.223***	0.047***
	[0.012]	[0.012]	[0.009]	[0.013]	[0.012]	[0.004]
pe (N=260)	0.019***	0.034***	0.039***	0.015	0.020*	0.005
	[0.004]	[0.010]	[0.010]	[0.010]	[0.009]	[0.006]
Financial Statement Ratios						
rnoa (N=292)	0.249***	0.274***	0.334***	0.025**	0.085***	0.060***
	[0.009]	[0.009]	[0.013]	[0.008]	[0.011]	[0.008]
roe (N=296)	0.037***	0.083***	0.094***	0.047***	0.057***	0.010**
	[0.009]	[0.010]	[0.009]	[0.007]	[0.007]	[0.003]
at (N=296)	0.407***	0.624***	0.662***	0.217***	0.254***	0.037***
	[0.011]	[0.011]	[0.012]	[0.009]	[0.008]	[0.004]
pm (N=295)	0.130***	0.224***	0.278***	0.094***	0.148***	0.054***
	[0.006]	[0.015]	[0.018]	[0.014]	[0.016]	[0.005]
lev (N=299)	0.041***	0.114***	0.132***	0.073***	0.090***	0.017***
	[0.006]	[0.016]	[0.018]	[0.011]	[0.014]	[0.005]
Other Financial Information						
ltgrowth (N=268)	0.150***	0.216***	0.223***	0.066***	0.073***	0.007
	[0.016]	[0.022]	[0.030]	[0.012]	[0.020]	[0.012]
salesgrowth (N=291)	0.175***	0.222***	0.258***	0.047***	0.084***	0.036***
	[0.018]	[0.020]	[0.024]	[0.013]	[0.014]	[0.007]
rdpersales (N=296)	0.698***	0.698***	0.744***	0.001	0.046***	0.046***
	[0.007]	[0.007]	[0.007]	[0.008]	[0.008]	[0.002]
Number of Quarters	12	12	12	12	12	12

Table IA4.
(Continued)

Panel B: S&P500 Base Firms, Restricted Model

	GICS6 (1)	SBP EW (2)	SBP TW (3)	(2)-(1) (4)	(3)-(1) (5)	(3)-(2) (6)
Valuation Multiples						
pb (N=297)	0.048***	0.130***	0.123***	0.082***	0.074***	-0.008
	[0.004]	[0.011]	[0.009]	[0.009]	[0.007]	[0.005]
evs (N=298)	0.252***	0.452***	0.491***	0.200***	0.239***	0.039***
	[0.012]	[0.012]	[0.012]	[0.010]	[0.011]	[0.003]
pe (N=260)	0.019***	0.029***	0.036***	0.010	0.017**	0.007
	[0.004]	[0.007]	[0.009]	[0.007]	[0.007]	[0.005]
Financial Statement Ratios						
rnoa (N=292)	0.249***	0.274***	0.335***	0.025**	0.086***	0.062***
	[0.009]	[0.008]	[0.013]	[0.008]	[0.010]	[0.008]
roe (N=296)	0.037***	0.074***	0.088***	0.037***	0.051***	0.015***
	[0.009]	[0.010]	[0.009]	[0.006]	[0.006]	[0.003]
at (N=296)	0.407***	0.613***	0.670***	0.205***	0.263***	0.057***
	[0.011]	[0.013]	[0.013]	[0.010]	[0.009]	[0.005]
pm (N=295)	0.130***	0.229***	0.297***	0.099***	0.167***	0.068***
	[0.006]	[0.014]	[0.018]	[0.012]	[0.018]	[0.007]
lev (N=299)	0.041***	0.119***	0.147***	0.078***	0.106***	0.028***
	[0.006]	[0.016]	[0.020]	[0.011]	[0.015]	[0.006]
Other Financial Information						
ltgrowth (N=268)	0.150***	0.222***	0.238***	0.072***	0.088***	0.016
	[0.016]	[0.026]	[0.032]	[0.012]	[0.021]	[0.013]
salesgrowth (N=291)	0.175***	0.213***	0.256***	0.038***	0.081***	0.043***
	[0.018]	[0.020]	[0.023]	[0.009]	[0.011]	[0.006]
rdpersales (N=296)	0.698***	0.707***	0.763***	0.009	0.065***	0.056***
	[0.007]	[0.007]	[0.008]	[0.006]	[0.005]	[0.002]
Number of Quarters	12	12	12	12	12	12

Table IA4.
(Continued)

Panel C: S&P1500 Base Firms

	GICS6 (1)	SBP EW (2)	SBP TW (3)	(2)-(1) (4)	(3)-(1) (5)	(3)-(2) (6)
Valuation Multiples						
pb (N=793)	0.028***	0.121***	0.124***	0.093***	0.096***	0.003
	[0.003]	[0.008]	[0.011]	[0.009]	[0.012]	[0.005]
evs (N=795)	0.234***	0.403***	0.439***	0.169***	0.205***	0.036***
	[0.006]	[0.009]	[0.008]	[0.008]	[0.006]	[0.003]
pe (N=676)	0.022***	0.025***	0.035***	0.003	0.013*	0.009***
	[0.005]	[0.005]	[0.007]	[0.005]	[0.006]	[0.003]
Financial Statement Ratios						
rnoa (N=783)	0.192***	0.238***	0.278***	0.046***	0.086***	0.041***
	[0.012]	[0.012]	[0.012]	[0.009]	[0.009]	[0.005]
roe (N=790)	0.021***	0.056***	0.067***	0.036***	0.046***	0.010***
	[0.006]	[0.007]	[0.008]	[0.005]	[0.005]	[0.003]
at (N=792)	0.434***	0.596***	0.617***	0.161***	0.183***	0.021***
	[0.009]	[0.008]	[0.008]	[0.005]	[0.006]	[0.001]
pm (N=790)	0.150***	0.261***	0.298***	0.111***	0.147***	0.036***
	[0.009]	[0.015]	[0.016]	[0.013]	[0.013]	[0.003]
lev (N=797)	0.029***	0.065***	0.066***	0.037***	0.037***	0.000
	[0.003]	[0.008]	[0.005]	[0.007]	[0.005]	[0.004]
Other Financial Information						
ltgrowth (N=658)	0.125***	0.171***	0.180***	0.046***	0.055***	0.009
	[0.007]	[0.010]	[0.015]	[0.008]	[0.012]	[0.007]
salesgrowth (N=761)	0.128***	0.159***	0.175***	0.030**	0.047***	0.016***
	[0.012]	[0.014]	[0.017]	[0.010]	[0.013]	[0.004]
rdpersales (N=792)	0.629***	0.655***	0.686***	0.025***	0.056***	0.031***
	[0.008]	[0.007]	[0.008]	[0.004]	[0.004]	[0.003]
Number of Quarters	12	12	12	12	12	12

Table IA4.
(Continued)

Panel D: S&P1500 Base Firms, Restricted Model

	GICS6 (1)	SBP EW (2)	SBP TW (3)	(2)-(1) (4)	(3)-(1) (5)	(3)-(2) (6)
Valuation Multiples						
pb (N=793)	0.028***	0.100***	0.089***	0.071***	0.061***	-0.011***
	[0.003]	[0.007]	[0.007]	[0.006]	[0.007]	[0.002]
evs (N=794)	0.234***	0.398***	0.444***	0.164***	0.210***	0.046***
	[0.006]	[0.008]	[0.009]	[0.006]	[0.005]	[0.003]
pe (N=676)	0.022***	0.026***	0.034***	0.004	0.012*	0.008**
	[0.005]	[0.005]	[0.006]	[0.005]	[0.006]	[0.003]
Financial Statement Ratios						
rnoa (N=782)	0.192***	0.240***	0.291***	0.048***	0.099***	0.051***
	[0.012]	[0.012]	[0.013]	[0.007]	[0.007]	[0.005]
roe (N=789)	0.021***	0.050***	0.056***	0.029***	0.035***	0.006*
	[0.006]	[0.007]	[0.008]	[0.004]	[0.005]	[0.003]
at (N=791)	0.434***	0.596***	0.632***	0.161***	0.197***	0.036***
	[0.009]	[0.010]	[0.009]	[0.006]	[0.006]	[0.003]
pm (N=790)	0.150***	0.256***	0.306***	0.106***	0.156***	0.050***
	[0.009]	[0.013]	[0.015]	[0.010]	[0.013]	[0.004]
lev (N=797)	0.029***	0.078***	0.074***	0.049***	0.045***	-0.003
	[0.003]	[0.008]	[0.006]	[0.006]	[0.004]	[0.004]
Other Financial Information						
ltgrowth (N=658)	0.125***	0.180***	0.190***	0.055***	0.065***	0.010
	[0.007]	[0.009]	[0.015]	[0.006]	[0.011]	[0.008]
salesgrowth (N=761)	0.128***	0.152***	0.170***	0.023**	0.042***	0.019***
	[0.012]	[0.011]	[0.015]	[0.008]	[0.010]	[0.004]
rdpersales (N=791)	0.630***	0.672***	0.712***	0.042***	0.082***	0.040***
	[0.008]	[0.007]	[0.009]	[0.004]	[0.004]	[0.002]
Number of Quarters	12	12	12	12	12	12