Online Appendix for
Common Analysts: Method for Defining Peer Firms

Markku Kaustia[*]
Aalto University

Ville Rantala
University of Miami

---

[*] E-mail addresses: markku.kaustia@aalto.fi, vrantala@bus.miami.edu.

I. Rules of thumb for defining analyst-based peer firms

This section investigates whether it is possible to derive a meaningful "rule of thumb" for defining peer firms based on analyst coverage, but without resorting to running the full simulation. The idea is to predict the simulation-generated peer groups with only basic information on analyst coverage, available from any analyst database. We consider linear prediction rules obtained by regressing the simulation-based peer criterion, C_i, on variables measuring firm i's annual analyst coverage. We use three different regression specifications. First, we simply regress C_i on the number of analysts following firm i. The second specification adds the average number of firms followed by each analyst following firm i as an explanatory variable. The third specification combines this information into a single variable and regresses C_i on the number of analysts following firm i multiplied by the average number of firms followed by each of the analysts.

The results from these regressions are reported in Panel A of Table A.5. All three specifications explain a significant part of the variation in C. The simplest specification based on the number of analysts obtains an adjusted R^2 of 69% and the second and third specifications obtain 80% and 82%, respectively. When actual simulation-based values for the peer criterion are compared to (rounded) predicted values from the regressions, the predictions from the simplest specification are identical to the actual values 64% of the time, and the second and third specifications are identical 75% and 76% of the time.

The coefficients in the regressions imply simple rules of thumb for approximating the peer criterion. In the first specification, the intercept is 2.9 and the coefficient for the number of analyst is 0.09. So for a firm i with 10 analysts, one should require $2.9 + 0.09 \times 10 = 3.8 \approx 4$ common analysts with firm j to consider it a peer. To rephrase this in more easily memorizable format, the lowest peer criterion to use in approximation is three, and it increases by roughly one unit per every ten analysts following the firm. The group sizes of the rule-of-thumb peer groups are quite close to the actual analyst-based peer groups (Panel B, Table A.5).

In the second specification, the intercept is 2.1, the coefficient for the number of analysts is almost identical, and the coefficient for the average number of analysts following firm i is 0.06. The third

1

specification has an intercept of 3.0 and the coefficient for the number of analysts following firm i multiplied with the average number of firms followed by each of the analyst is 0.006.

To find out how these approximations perform, Table A.6 reports results from peer group homogeneity regressions using predicted instead of actual analyst-based peer groups. The adjusted R^2 values obtained with different fitted value specifications are almost identical to each other, and they indicate performance close to the actual analyst-based peer groups.

Overall, the results indicate that there is only little difference in group size or performance between the different fitted value specifications. Even the simplest rule of thumb based only on the number of analysts provides a reasonable approximation for the simulation-based peer criterion and works efficiently in most cases. It is of course possible that these linear rules work less effectively in the ends of the analyst coverage distribution.

II. Extending the peer groups to firms that are not covered by analysts

Here we discuss how to expand the reach of our method to include firms without analyst coverage. This entails placing uncovered firms into existing groups on the grounds of some other similarity metric.[1] The general idea here is to ask which firms might analysts 'assign' as peers (by their coverage choices) for an uncovered firm if it were covered.

There are many potential ways to answer this question, but as one example, here we employ Hoberg-Phillips TNIC peer groups. We take each uncovered firm and match it with its Hoberg-Phillips TNIC peers that have an analyst-based peer group. The TNIC data include similarity scores between firms, so we can find the closest such TNIC peer. We then expand the existing analyst-based peer groups so that we assign the two newly matched firms as each other's peers, and assign all the peers of the TNIC matching firm as the uncovered firm's peers. We also place the uncovered firm into all peer groups that include the TNIC match. To illustrate, suppose firm X does not have analyst-based peers. Its best TNIC match is firm Y, and

---

[1] We refer to any firm without an analyst-based peer group as uncovered. Some of these firms may be covered by individual analysts but they do not have sufficient analyst coverage to have analyst-based peers.

Y has analyst-based peers A, B, and C. After matching X with Y, X's peer group consists of Y, A, B, and C, and Y's peer group consists of X, A, B, and C. We also assign X as an additional peer firm for all other firms that have Y as their analyst-based peer. This method allows us to utilize analysts' coverage choices in peer identification even for uncovered firms.

The extended peer groups formed with this method have an average group size of 14.1 while the average group size for regular analyst-based peer groups is 11.8 (including firms that do not have a TNIC group). Table A.7 shows results from homogeneity regressions with the extended peer groups. The adjusted R^2 values indicate that the extended analyst-based peer groups outperform TNIC peer groups on 11 of the 20 test variables.

Table A.1

R²'s from Peer Group Homogeneity Regressions: Set of Common Finance Variables - Extended Sample

This table compares the explanatory power of analyst-based peer averages to the explanatory power of industry classification-based peer averages in a set of common finance variables. The sample consists of firms with CRSP share code 10 or 11 and the regressions reported in this table are identical to the regressions reported in Table 4 except for the sample composition. The results in each panel are based on the time period for which the industry groups are available. The reported figures are adjusted R^2 values from a regression where the dependent variables is a firm's variable value, and the independent variable is either the average variable value among analyst-based peers or among other firms sharing the same industry classification code. The regression includes a constant. Panel A compares the analyst-based peer groups to SIC codes and Fama-French industries, Panel B to NAICS codes, Panel C to GICS codes, and Panel D to TNIC groups. The results in each panel are based on the time period for which the industry groups are available. Analyst-based peer regressions are also run separately using analyst-weighted peer average values, with the number of common analysts between firms as weights. Industry classification regressions are run separately based on classification groups among all firms and among firms with analyst-based peers. The test variables are defined as in Table 4. All variable values except the monthly returns are annual values at the end of the year. Financial statement items for year t are based on the fiscal year ending in year t. The highest adjusted R^2 value for each variable in each panel is boldface.

	MONTHLY _RET	MONTHLY _RET_FE	BETA	MARKET_ CAP	TOTAL_ ASSETS	NET_ SALES	MARKET-TO-BOOK	DIV_ PAYMENT	BOOK_ LEVG	MARKET_ LEVG
Panel A. Analyst-Based Peers Compared to SIC Codes and Fama-French Industries										
Analyst-based peers										
Analyst-weighted average	22.4	25.4	**40.9**	**14.2**	**29.4**	**23.3**	24.8	**36.8**	**15.4**	**28.5**
Equal-weighted average	22.8	**25.5**	40.4	9.5	24.0	18.1	24.6	35.5	14.8	27.9
Industry groups based on all firms										
2-digit SIC	12.7	12.9	18.1	7.5	16.5	8.7	15.8	19.3	7.0	16.6
3-digit SIC	11.7	12.9	21.8	7.0	18.6	9.9	18.7	21.6	6.9	18.6
4-digit SIC	10.7	12.7	21.7	6.7	18.3	10.2	18.5	21.3	6.8	18.5
Fama-French Industries	13.4	13.5	20.8	6.7	13.6	7.6	18.3	21.7	7.4	18.2
Industry groups based on firms with analyst-based peers										
2-digit SIC	23.5	23.7	27.2	7.2	16.9	9.0	21.2	29.6	10.9	21.0
3-digit SIC	23.0	24.3	32.3	5.8	18.2	10.4	26.0	31.2	12.8	25.1
4-digit SIC	21.9	23.9	32.5	5.6	18.1	10.9	**26.2**	30.6	12.8	25.2
Fama-French Industries	**24.5**	24.6	29.8	5.8	14.7	7.5	25.1	33.5	12.1	23.0

4

	MONTHLY _RET	MONTHLY _RET_FE	BETA	MARKET_ CAP	TOTAL_ ASSETS	NET_ SALES	MARKET- TO-BOOK	DIV_ PAYMENT	BOOK_ LEVG	MARKET_ LEVG
Panel B. Analyst-Based Peers Compared to 3- and 5-digit NAICS Industries										
Analyst-based peers										
Analyst-weighted average	22.6	25.6	**40.8**	**14.9**	**27.2**	**23.6**	25.7	**38.1**	**15.4**	**28.5**
Equal-weighted average	22.9	**25.6**	40.4	10.1	21.7	18.3	25.4	37.0	14.8	27.9
Industry groups based on all firms										
3-digit NAICS	12.7	12.9	19.9	6.7	10.9	8.2	16.6	18.6	7.1	17.4
5-digit NAICS	10.4	12.5	21.8	7.5	12.6	9.8	18.4	19.5	6.3	18.2
Industry groups based on firms with analyst-based peers										
3-digit NAICS	**23.6**	24.0	28.8	6.6	13.4	8.9	22.0	28.7	11.5	22.2
5-digit NAICS	22.4	24.2	33.1	6.9	14.6	10.7	**26.4**	28.8	12.4	25.5
Panel C. Analyst-Based Peers Compared to 6- and 8-digit GICS Industries										
Analyst-based peers										
Analyst-weighted average	28.1	30.3	**46.1**	**14.6**	**29.5**	**20.3**	**27.2**	**42.8**	**14.6**	**26.6**
Equal-weighted average	29.3	**31.0**	45.6	9.8	24.1	15.3	26.5	41.8	13.5	25.5
Industry groups based on all firms										
6-digit GICS	28.0	28.3	39.0	6.8	19.4	7.7	23.1	36.8	9.6	19.1
8-digit GICS	27.2	28.3	39.8	6.9	21.5	10.8	22.7	35.6	11.1	21.1
Industry groups based on firms with analyst-based peers										
6-digit GICS	**30.1**	30.5	42.0	7.0	19.3	8.2	25.3	40.3	10.3	20.5
8-digit GICS	29.4	30.5	42.9	6.7	21.4	11.0	24.7	38.7	11.6	22.5
Panel D. Analyst-Based Peers Compared to TNIC 3 Groups										
Analyst-Based Peers										
Analyst-weighted average	23.3	26.1	42.4	**12.3**	**25.9**	**24.2**	24.7	**32.5**	16.3	30.1
Equal-weighted average	23.7	26.3	41.8	8.1	20.6	19.0	24.8	31.0	15.8	29.8
TNIC										
Groups based on all firms	14.7	15.5	35.8	6.8	11.4	11.9	25.0	25.2	10.4	25.9
Groups based on firms with analyst-based peers	**26.2**	**26.5**	**42.6**	5.9	13.4	13.8	**28.3**	32.2	**17.9**	**31.9**

R²'s from Peer Group Homogeneity Regressions: Accounting-Based Financial Information Variables - Extended Sample

This table compares the explanatory power of analyst-based peer averages to the explanatory power of industry classification group-based peer averages in a set of accounting-based financial information variables. The sample consists of firms with CRSP share code 10 or 11 and the regressions reported in this table are identical to the regressions reported in Table 5 except for the sample composition. The results in each panel are based on the time period for which the industry groups are available. The reported figures are adjusted R^2 values from a regression where the dependent variables is a firm's variable value, and the independent variable is either the average variable value among analyst-based peers or among other firms sharing the same industry classification code. The regression includes a constant. Panel A compares the analyst-based peer groups to SIC codes and Fama-French industries, Panel B to NAICS codes, Panel C to GICS codes, and Panel D to TNIC groups. Analyst-based peer regressions are also run separately using analyst-weighted peer average values with the number of common analysts between firms as weights. Industry classification regressions are run separately based on classification groups among all firms and among firms with analyst-based peers. The test variables are defined as in Table 5. The variables are defined as in Table 5. All variable values are annual values at the end of the year. Financial statement items for year t are based on the fiscal year ending in year t. The highest adjusted R^2 value for each variable in each panel is boldface.

	PRICE-TO-BOOK	EV-TO-SALES	PE	RNOA	ROE	AT	PROFIT_MARGIN	LEVG	SALES_GROWTH	SCALED_R&D_EXPENSE
Panel A. Analyst-Based Peers Compared to SIC Codes and Fama-French Industries										
Analyst-based peers										
Analyst-weighted average	16.2	**40.7**	6.8	**17.0**	**9.4**	82.5	21.9	**57.6**	**13.8**	**45.3**
Equal-weighted average	15.7	39.4	6.6	15.9	9.0	81.8	20.9	56.8	13.4	44.3
Industry groups based on all firms										
2-digit SIC	11.5	28.9	3.7	4.7	4.4	81.3	20.8	54.0	8.1	20.3
3-digit SIC	12.1	34.0	3.7	5.6	4.4	84.0	22.1	55.3	8.0	28.2
4-digit SIC	11.9	34.8	3.5	5.4	4.0	**84.3**	22.0	55.0	7.4	32.0
Fama-French Industries	12.7	31.1	4.9	5.9	5.3	77.5	21.4	51.4	8.2	28.1
Industry groups based on firms with analyst-based peers										
2-digit SIC	15.0	31.3	6.4	10.1	6.6	75.7	20.7	54.2	11.4	22.2
3-digit SIC	15.7	38.1	6.3	11.0	7.1	79.7	22.1	56.3	11.5	31.3
4-digit SIC	15.8	39.8	5.8	10.7	6.8	80.4	**22.2**	56.3	11.2	33.9
Fama-French Industries	**16.2**	34.9	**7.7**	11.1	7.9	72.6	21.1	52.6	11.9	32.2

	PRICE-TO-BOOK	EV-TO-SALES	PE	RNOA	ROE	AT	PROFIT_MARGIN	LEVG	SALES_GROWTH	SCALED_R&D_EXPENSE
Panel B. Analyst-Based Peers Compared to 3- and 5-digit NAICS Industries										
Analyst-based peers										
Analyst-weighted average	**17.1**	**41.3**	**6.9**	**17.1**	**9.6**	**85.1**	**23.8**	**64.1**	**13.8**	**45.7**
Equal-weighted average	16.7	40.1	6.7	16.1	9.2	84.6	22.9	63.5	13.4	44.6
Industry groups based on all firms										
3-digit NAICS	11.9	29.2	4.1	5.7	4.8	80.2	22.5	54.5	8.4	23.2
5-digit NAICS	11.6	33.0	3.4	6.1	4.2	84.4	22.9	55.9	7.6	29.3
Industry groups based on firms with analyst-based peers										
3-digit NAICS	15.3	31.8	6.6	10.9	7.4	74.8	22.5	55.9	12.1	24.0
5-digit NAICS	15.8	37.5	6.3	12.2	7.3	80.6	23.2	57.5	12.0	32.2
Panel C. Analyst-Based Peers Compared to 6- and 8-digit GICS Industries										
Analyst-based peers										
Analyst-weighted average	**18.3**	**43.8**	6.1	**20.1**	**7.7**	**84.9**	30.1	**60.3**	16.5	**54.0**
Equal-weighted average	17.3	42.0	5.9	18.6	6.9	83.8	28.7	59.0	16.2	52.4
Industry groups based on all firms										
6-digit GICS	15.7	38.5	6.2	16.0	6.7	80.3	26.9	56.1	16.2	45.7
8-digit GICS	15.2	41.5	6.3	16.5	6.8	82.9	29.1	58.4	16.1	45.3
Industry groups based on firms with analyst-based peers										
6-digit GICS	17.2	40.1	**6.5**	19.5	6.8	80.9	28.6	57.1	**17.1**	49.7
8-digit GICS	16.4	43.5	6.4	18.9	6.7	83.3	**30.8**	59.4	16.5	49.3
Panel D. Analyst-Based Peers Compared to TNIC 3 Groups										
Analyst-Based Peers										
Analyst-weighted average	14.9	**39.6**	7.0	18.0	9.6	**84.3**	22.0	58.1	15.2	**42.7**
Equal-weighted average	14.7	38.7	6.6	16.8	9.2	83.8	21.0	57.3	14.9	41.7
TNIC										
Groups based on all firms	13.9	35.3	5.7	13.9	9.7	81.3	**27.3**	**58.4**	12.2	39.8
Groups based on firms with analyst-based peers	**16.1**	38.1	**8.4**	**21.1**	**12.7**	76.9	26.4	55.5	**17.0**	40.6

Table A.3

R^2's from Peer Group Homogeneity Regressions: Common Time-Period 1996-2011 - Extended Sample

This table compares the explanatory power of analyst-based peer averages to the explanatory power of 3-digit SIC, Fama-French 49, 5-digit NAICS, 8-digit GICS, and Hoberg-Phillips TNIC classification peer averages in two sets of firm characteristic variables during the 1996 to 2011 time period. The sample consists of firms with CRSP share code 10 or 11, an analyst-based peer group, and classification data for all the industry classifications. The regressions reported in this table are identical to the regressions reported in Table 5 except for the sample composition. The reported figures are adjusted R^2 values from a regression where the dependent variables is a firm's variable value, and the independent variable is either the average variable value among analyst-based peers or among other firms in the same industry classification group. Analyst-based peer regressions are also run separately using a weighted peer average with the number of common analysts between firms as weights. The industry classification groups are formed based on firms that have analyst-based peers. The regression includes a constant. Variable List 1 consists of the common finance variables which are defined as in Table 4 and Variable List 2 consists of the financial ratios and accounting-based financial information variables defined as in Table 5. The highest adjusted R^2 value for each variable in each panel is boldface.

	Analyst-Based Peers		Industry Classification Groups Based on Firms with Analyst-Based Peers				
	Analyst-Weighted Average	Equal-Weighted Average	3-digit SIC	Fama-French	5-digit NAICS	8-digit GICS	TNIC
Variable List 1							
MONTHLY_RET	28.0	29.3	26.3	28.0	24.4	**29.7**	28.8
MONTHLY_RET_FE	30.3	**31.1**	28.3	28.0	27.4	30.8	29.9
BETA	**46.6**	46.4	37.4	34.4	37.7	43.1	44.3
MARKET_CAP	**12.5**	8.3	3.1	3.0	4.2	4.0	3.8
TOTAL_ASSETS	**26.0**	20.7	15.1	12.5	15.2	15.7	13.9
NET_SALES	**20.8**	15.6	7.4	4.6	7.8	8.8	8.5
MARKET-TO-BOOK	25.2	24.7	**25.3**	22.5	25.2	24.2	24.6
DIV_PAYMENT	**35.4**	34.3	26.1	29.5	26.6	31.3	32.5
BOOK_LEVG	**15.6**	14.3	10.5	9.5	12.3	13.8	14.5
MARKET_LEVG	**27.0**	26.2	22.7	19.2	24.2	24.2	25.8
Variable List 2							
PRICE-TO-BOOK	**16.0**	15.1	14.5	13.4	14.0	14.3	15.3
EV-TO-SALES	**43.0**	41.1	40.7	37.0	39.5	42.6	38.3
PE	6.5	6.2	6.0	6.9	6.8	7.3	**7.6**
RNOA	**19.2**	17.8	13.8	12.2	14.0	18.9	17.6
ROE	7.4	6.6	5.2	4.9	4.1	6.8	**8.1**
AT	**84.4**	83.2	84.0	76.0	82.0	82.2	73.8
PROFIT_MARGIN	28.1	26.5	28.8	26.1	28.1	**29.9**	25.4
LEVG	54.9	53.3	**56.6**	51.8	54.3	55.3	50.5
SALES_GROWTH	17.7	17.6	15.7	16.0	16.1	**18.8**	18.1
SCALED_R&D_EXPENSE	**52.2**	50.2	34.1	33.9	33.5	49.3	44.9
Average of (R^2/Highest R^2)	0.98	0.91	0.77	0.73	0.78	0.87	0.86

Table A.4

R²'s from Peer Group Homogeneity Regressions: Identifying a Single Matching Firm

This table compares the explanatory power of common analyst-matched individual comparable firms to size and market-to-book matched comparable firms based on a set of common finance variables. The sample covers years between 1983 and 2013 and includes all firms with CRSP share codes 10 or 11. Common analyst matched comparable firm for firm i is the firm with the highest number of common analysts with firm i. In case there are several firms with the same number of common analysts, one of them is selected randomly. Statistics for the common analyst matched firms are reported separately for different minimum analyst criteria ranging from 1 to 5, so that firms with less than the criterion number of common analysts with any other firm are excluded from the analysis. Common analyst matched comparable firms are compared to comparable firms matched based on the closest MARKET_CAP or MARKET-TO-BOOK value. The matches are based on annual MARKET_CAP and MARKET-TO-BOOK values defined as in Table 4. The matching is conducted separately among all firms and among firms with the same 4-digit SIC code. We also run comparisons with comparable firms selected using joint MARKET_CAP and MARKET-TO-BOOK matching, so that MARKET-TO-BOOK matching is conducted among firms whose market capitalization is between 70% and 130% of the MARKET_CAP of firm i. The reported figures are adjusted R^2 values from a regression where the dependent variable is a firm's variable value, and the independent variable is the variable value for the matched comparable firm. The regression includes a constant. All variable values except the monthly returns are annual values at the end of the year. Variable List 1 consists of the common finance variables which are defined as in Table 4 and Variable List 2 consists of the financial ratios and accounting-based financial information variables defined as in Table 5. The highest adjusted R^2 value for each variable is boldface.

Panel A. Variable List 1

	Comparables Matched Based on the Highest Number of Common Analysts					MARKET_CAP Matching		MARKET-TO-BOOK Matching		Size and M/B Matching
	Minimum Number of Common Analysts					Among All Firms	Among 4-digit SIC Group	Among All Firms	Among 4-digit SIC Group	Among All Firms
	1	2	3	4	5					
MONTHLY_RET	9.2	12.4	15.7	18.3	**20.5**	1.4	3.4	1.5	3.2	2.2
MONTHLY_RET_FE	16.8	19.8	22.8	24.9	**26.8**	10.1	11.0	10.2	10.9	10.3
BETA	23.2	28.4	32.8	36.3	**38.7**	0.5	11.6	0.7	9.9	2.6
MARKET_CAP	**17.4**	16.7	16.3	16.0	15.8	—	—	0.1	3.2	—
TOTAL_ASSETS	25.5	25.9	26.1	26.3	27.0	24.9	53.7	0.2	7.4	**66.8**
NET_SALES	23.1	23.1	23.4	23.5	23.7	33.9	45.1	0.1	4.3	**46.3**
MARKET-TO-BOOK	14.0	16.7	18.7	19.8	**20.9**	0.3	9.7	—	—	—
DIV_PAYMENT	20.8	24.4	27.5	29.3	**30.3**	2.1	14.6	0.2	7.1	4.4
BOOK_LEVG	5.7	7.6	9.1	9.4	**10.3**	0.1	2.4	0.6	3.1	1.2
MARKET_LEVG	14.5	18.0	21.0	22.1	23.2	1.1	11.1	22.3	**25.5**	22.8

9

Panel B. Variable List 2

	Comparables Matched Based on the Highest Number of Common Analysts					MARKET_CAP Matching		MARKET-TO-BOOK Matching		Size and M/B Matching
	Minimum Number of Common Analysts					Among All Firms	Among 4-digit SIC Group	Among All Firms	Among 4-digit SIC Group	Among All Firms
	1	2	3	4	5					
PRICE-TO-BOOK	8.9	10.2	11.5	12.5	13.3	1.3	7.9	42.3	38.5	**47.0**
EV-TO-SALES	27.0	30.4	33.8	36.0	**37.9**	1.2	27.0	6.2	36.3	9.2
PE	2.5	3.1	3.5	4.1	**4.2**	0.0	0.9	0.2	1.1	0.6
RNOA	7.6	9.5	11.7	13.5	**14.8**	1.4	5.2	0.4	3.1	3.2
ROE	3.3	3.9	4.8	5.4	**5.5**	0.9	2.8	0.2	1.5	1.7
AT	78.6	80.6	**80.8**	80.1	79.9	0.2	75.6	14.9	77.0	20.7
PROFIT_MARGIN	13.1	14.1	15.6	18.0	**18.6**	0.8	13.2	0.9	10.8	4.1
LEVG	49.0	52.3	52.8	**53.1**	53.0	0.1	36.7	11.7	37.7	16.5
SALES_GROWTH	5.1	6.3	7.9	8.7	**9.7**	0.5	3.2	1.2	3.9	1.5
SCALED_R&D_EXPENSE	27.9	32.9	36.8	39.7	**41.2**	0.3	15.9	0.8	14.5	1.5

Table A.5

Regressions Explaining the Simulated Peer Limit with Analyst Coverage Variables

Panel A reports results from regressions explaining the simulated peer criterion with variables related to analyst coverage. A firm's peer group in year t consists of all firms that are followed by at least the criterion number of same analysts in year t. The sample consists of firm-year observations from the time period 1983-2013. The independent variables is a firm's peer criterion and the dependent variables are NUMBER_OF_ANALYSTS (the number of analysts covering the firm), AVG_NUMBER_OF_FIRMS_COVERED (average number of firms followed by each analyst covering the firm), and the number of analysts covering the firm multiplied with the average number of firms followed by each analyst. *t*-statistics are reported below the coefficients. The percentage of firms whose (rounded) fitted value for the peer criterion in the regression is identical to the peer criterion in the simulation is reported at the bottom of the panel. Panel B reports statistics based on analyst-based peer groups that are formed using fitted values from the regressions reported in Panel A. The statistics include the percentage of firms with a peer group, average, median, and maximum peer group sizes, and standard deviation of group size. Columns 1-3 in Panel B are based on the regressions reported in the corresponding columns in Panel A. Statistics for the actual peer group formed based on the simulated peer criterion are also reported for comparison.

Panel A. Regressions Explaining the Simulated Peer Limit

	(1)	(2)	(3)
INTERCEPT	2.890	2.115	2.960
	[616.0]	[330.0]	[898.3]
NUMBER_OF_ANALYSTS	0.088	0.086	
	[313.0]	[378.5]	
AVG_NUMBER_OF_FIRMS_COVERED		0.059	
		[150.3]	
NUMBER_OF_ANALYSTS × AVG_NUMBER_OF_FIRMS_COVERED			0.006
			[444.8]
Adj. R^2	0.69	0.80	0.82
N	43,235	43,235	43,235
Same implied peer limit as in simulation (%)	64.1	74.7	75.7

Panel B. Peer Groups Based on Predicted Peer Limits Compared to Actual Analyst-Based Peer Groups

	(1)	(2)	(3)	Actual Peer Groups
% of firms with a peer group	77.49	77.59	76.63	77.21
Peer group size average	13.4	12.8	12.7	14.2
Peer group size median	11	11	12	11
Peer group size maximum	72	57	50	81
Peer group size stdev	10.38	9.35	8.81	13.56

Table A.6

The Homogeneity of Analyst-Based Peer Groups that Are Formed Based on a Predicted Peer Criterion

This table compares the explanatory power of analyst-based peer groups that are formed based on three different types of predicted peer criteria using two sets of firm characteristic variables. The sample consists of NYSE firms with CRSP Share Code 10 or 11 and a SIC code in Compustat. The reported figures are adjusted R^2 values from a regression where the dependent variables is a firm's variable value and the independent variable is the average variable value among analyst-based peers. Variable List 1 consists of the common finance variables which are defined as in Table 4 and Variable List 2 consists of the financial ratios and accounting-based financial information variables defined as in Table 5. The predicted peer criteria that are used to form the peer groups are based on fitted values from three regressions reported in Panel A of Table 8. The peer groups used in columns 1 to 3 of this table are based on fitted values from regressions reported in the corresponding columns in Table 8. The independent variable in these regressions is the firm's simulation-based peer criterion and the dependent variables are the number of analysts following the firm, the average number of firms followed by each analyst, and the number of analysts following the firm multiplied with the average number of firms followed by each analyst. All regressions include a constant. A firm's peer group in year t consists of all firms that are followed by at least the criterion number of same analysts in year t.

	Single Explanatory Variable: Number of Analysts	Two Explanatory Variables: (1) Number of Analysts, (2) Average Number of Firms per Analyst	Single Explanatory Variable: Number of Analysts Times Average Number of Firms per Analyst
	(1)	(2)	(3)
Variable List 1			
MONTHLY_RET	29.4	29.4	29.5
MONTHLY_RET_FE	31.5	31.5	31.4
BETA	45.2	44.7	44.6
MARKET_CAP	10.3	10.4	10.3
TOTAL_ASSETS	23.4	23.9	23.8
NET_SALES	13.8	13.8	13.8
MARKET-TO-BOOK	24.2	24.0	24.1
DIV_PAYMENT	38.6	37.9	38.1
BOOK_LEVG	7.1	7.0	7.2
MARKET_LEVG	19.9	19.9	19.9
Variable List 2			
PRICE-TO-BOOK	16.0	16.0	16.0
EV-TO-SALES	44.0	43.9	43.6
PE	4.6	4.6	4.8
RNOA	23.5	23.0	22.8
ROE	7.1	7.1	7.7
AT	77.6	77.9	77.8
PROFIT_MARGIN	34.5	34.3	34.8
LEVG	49.5	50.0	49.9
SALES_GROWTH	16.7	16.5	16.4
SCALED_R&D_EXPENSE	52.3	51.2	52.2

Table A.7

The Homogeneity of Extended Analyst-Based Peer Groups

This table reports the explanatory power of extended analyst-based peer groups in homogeneity regressions using two sets of firm characteristic variables. These peer groups are analyst-based peer groups that have been extended to include firms that are not covered by analysts. The sample consists of NYSE firms with CRSP Share Code 10 or 11 and a TNIC classification. The reported figures are adjusted R^2 values from a regression where the dependent variable is a firm's variable value and the independent variable is the average variable value among peers. Variable List 1 consists of the common finance variables which are defined as in Table 4 and Variable List 2 consists of the financial ratios and accounting-based financial information variables defined as in Table 5. The extended analyst-based peer groups are formed so that each firm without analyst-based peers is first match with the closest comparable firm among the firms that have an analyst-based peer group. We conduct the matching based on the similarity scores of the Hoberg-Phillips TNIC classification. We then expand the existing peer groups so that we assign the two firms as each other's peers and assign the matched firm's peers as the uncovered firm's peers. We also place the uncovered firm into all peer groups that include the closest match.

Variable List 1		Variable List 2	
MONTHLY_RET	25.1	PRICE-TO-BOOK	10.6
MONTHLY_RET_FE	27.6	EV-TO-SALES	39.0
BETA	35.6	PE	3.1
MARKET_CAP	3.2	RNOA	14.8
TOTAL_ASSETS	12.0	ROE	5.7
NET_SALES	5.4	AT	67.7
MARKET-TO-BOOK	19.3	PROFIT_MARGIN	28.4
DIV_PAYMENT	25.2	LEVG	34.3
BOOK_LEVG	5.8	SALES_GROWTH	16.1
MARKET_LEVG	17.0	SCALED_R&D_EXPENSE	45.7