Firm geographic dispersion and financial analysts' forecasts

Published Version:
Platikanova, P., Mattei, M.M. (2016). Firm geographic dispersion and financial analysts' forecasts. JOURNAL OF BANKING & FINANCE, 64, 71-89 [10.1016/j.jbankfin.2015.11.012].

Availability:
This version is available at: https://hdl.handle.net/11585/555740 since: 2016-07-18

Published:
DOI: http://doi.org/10.1016/j.jbankfin.2015.11.012

Firm geographic dispersion and financial analysts' forecasts

Petya Platikanova [a,*], Marco Maria Mattei [b]

[a] ESADE-University Ramon Llull, Av Pedralbes 60, 08034 Barcelona, Spain
[b] Department of Management, University of Bologna, via Capo di Lucca, 34, 40126 Bologna, Italy

Using a text-based measure of geographic dispersion that captures the economic ties between a firm and its geographically distributed economic interests, this study provides evidence that financial analysts issue less accurate, more dispersed and more biased earnings forecasts for geographically dispersed firms. We observe the degree to which a firm has an overlapping distribution of economic centers in comparison to industry competitors and suggest that geographically similar firms have lower information gathering costs and thereby more precise earnings forecasts. Empirical evidence supports this prediction. We further find that the geographic dispersion across the U.S. is less likely to affect forecast precision when a firm has economic activities in states with highly correlated local shocks. Our findings suggest that the effect of geographic dispersion is more pronounced for soft-information environments where information is more difficult to make impersonal by using technological advances. Consistent with the information asymmetry argument, we find that geographically dispersed firms have less comparable and more discretionary managed earnings, have less extensive than industry competitors segment information, are more likely to restate sale segment information, and issue annual and quarterly filings with a delay.

1. Introduction

A growing body of literature in finance suggests that relevant information regarding the future cash flows and earnings of publicly traded U.S. firms is geographically dispersed across states in the U.S. (Garcia and Norli, 2012; Giroud, 2013; Addoum et al., 2014; Bernile et al., 2015). Recent studies demonstrate that the geographic dispersion of corporate activities across multiple U.S. states presents a significant problem in making efficient investment and financial decisions. Giroud (2013) finds that managers are more likely to locate plants close to headquarters to maintain better control over production. Landier et al. (2009) show the distance between divisions and headquarters affects the decision regarding who to layoff off and which division to divest. Coval and Moskowitz (1999) find that mutual fund managers are more likely to invest in local stocks and earn abnormal returns from these investments, suggesting either improved monitoring capabilities or access to private information. Addoum et al. (2014) show that firms exhibit stronger post-earnings-announcement drift and stronger momentum in returns when geographic information is more dispersed and difficult to aggregate.

In this paper, we investigate the extent to which the degree of geographic dispersion across U.S. states affects the ability of financial analysts to produce more precise earnings forecasts. The geographic dispersion may simplify the forecasting task of financial analysts by reducing performance volatility and facilitating high-quality forecasting model inputs, thus yielding more precise earnings forecasts. Alternatively, the spatial distribution of firms' activities could result in more complex forecasting of future earnings. First, relevant information regarding past performance and future trends may not be available for geographically dispersed firms because managers may not be efficient in aggregating and reporting value-relevant information regarding centers of business activities (Addoum et al., 2014). Second, the presence of business activities in multiple U.S. states is likely to increase management discretion and operating flexibility (e.g., shifting profits to Delaware; Dyreng et al., 2013). When such actions increase information asymmetry between analysts and management of dispersed firms, the precision of analysts' forecasts is likely to decrease. Moreover, the variation across states (e.g., tax codes) may increase complexity in forecasting when following firms with interstate operations (e.g., Florian and Ljungqvist, 2015).

Using a text-based measure of geographic dispersion, which captures the economic ties between a firm's headquarter and its geographically distributed economic interests (Garcia and Norli, 2012; Addoum et al., 2014), we provide empirical evidence that financial analysts issue less accurate, more dispersed and more biased forecasts for geographically dispersed firms. Empirical findings are consistent with the notion that a less transparent and more discretional aggregation of geographically dispersed information into financial reports increases the information asymmetry between management and financial analysts. We expect information gathering costs to be lower for geographically similar firms where comparable disclosure is provided by industry competitors. Consistent with this expectation, we find that geographically similar firms have less dispersed and biased forecasts compared to geographically different firms. We also predict that analysts experience forecasting difficulties for geographically dispersed firms, when economic shocks are imperfectly correlated across the U.S. Using the correlation between the headquarter state and relevant economic centers, we show that firms with highly correlated economic centers have more precise analysts' forecasts compared to other firms.

Innovation in information technology may reduce any potential difficulties in following firms with operations in multiple states. Petersen and Rajan (2002) suggest that such innovation explains the increasing distance between primarily lenders and borrowing firms. Following Landier et al. (2009), we use information regarding changes in the distance between borrowing firms and primary lenders from the National Survey of Small Business Finance and characterize firms as operating in a soft- (hard-) information environment. The empirical findings suggest that the effect of geographic dispersion on analysts' precision is more pronounced in soft-information environments. Therefore, the interstate dispersion is likely to reduce analysts' precision when information asymmetries are already high.

We next examine possible reasons for higher precision of financial analysts' forecasts for less dispersed firms' stocks. Because accounting comparability and discretionary accruals can significantly increase the forecasting error (Francis et al., 2004; De Franco et al., 2011; Veenman, 2012), we investigate the relation between accounting properties and geographic dispersion. Our results show that financial information in earnings of geographically dispersed firms is less comparable and more discretionarily managed than that of local firms. The presence of information asymmetry between management and outside may increase the demand for voluntary disclosure, providing management incentives to increase disclosure (Grossman and Hart, 1980). Using the information quality of corporate disclosure, we find that geographically dispersed firms do not provide more decomposed segment disclosure, have more often restated geographic sales, and delay the release of annual and quarterly filings. Lower information quality is consistent with the argument that accurate information may be more difficult to collect for economically dispersed relevant centers.

The contribution of this study is threefold. First, recent studies demonstrate that distance is relevant to improving operating efficiency (Giroud, 2013) and achieving superior trading performance (Coval and Moskowitz, 1999; Hau, 2001). Malloy (2005) argues that the distance between financial analysts and management is also important in determining the precision of forecast analysts. This study suggests that the distance between a firm's headquarters and relevant economic centers affects analyst performance, and provides novel evidence that firms' geographic dispersion across the U.S. determines analysts' forecast precision. Second, a large body of accounting literature investigates the determinants of a manager's reporting and disclosure choices (Verrecchia, 2001; Dechow et al., 2010). We show that the spatial distribution of firms' activities affects the quality of corporate disclosure and relevant properties of accounting information. Third, we complement existing literature on international and industrial diversification (e.g., Duru and Reeb, 2002) by presenting empirical evidence that the within-country variation in the distribution of economic activities also influences the forecasting task and analyst performance.

This paper is organized as follows. Section 2 outlines and develops the hypotheses related to the influence of geographic dispersion on analysts' forecast precision, accounting properties, and disclosure choices. Section 3 details the sample formation process and defines the variables. The empirical link between geographic dispersion across U.S. states and analysts' precision, performance volatility, and financial information quality are presented in Section 4. Section 5 presents additional empirical analysis, and Section 6 concludes the paper.

2. Hypothesis development

The degree of information asymmetry between managers and outsiders may differ for dispersed versus local firms. Aggregated cash flows and other diversification-related information problems may make it more difficult for analysts to forecast cash flows of diversified firms in comparison to focused firms. Consistent with such an interpretation, Thomas (2002) argues that diversification-related information problems explain larger forecast errors and greater dispersion among analysts' forecasts. In the same vein, Litov et al. (2012) argue that a firm's diversification across different industries requires either multiple analysts to collaboratively evaluate the firm or analysts to develop expertise across multiple industries. Duru and Reeb (2002) additionally identify the operating flexibility of international diversification as a possible source of information asymmetry between management and outsiders and an additional layer of difficulty to analysts' forecasting of future performance.

Consistent with the literature on industrial and international diversification, we conjecture that firms with largely dispersed business activities across the U.S. states may have higher information asymmetry problems and thus less precise analyst forecasts. One possible source of information asymmetry is the aggregation of financial information (Thomas, 2002; Frankel et al., 2006; Addoum et al., 2014). Whereas managers of geographically diversified firms can observe cash flows in each U.S. state, outsiders can observe only noisy estimates of these cash flows. Thus, the mapping of geographically dispersed cash flows into consolidated earnings may be less transparent to outsiders, and as a result, reported earnings may convey less value-relevant information. The precision of analyst estimates for geographically dispersed firms therefore may depend on the extent to which analysts understand the underlying earnings generation process in the presence of multiple centers of business activities (Abarbanell and Lehavy, 2003; Burgstahler and Eames, 2006; Coles et al., 2006). Because the level of discretion in accessing and aggregating geographically dispersed information may be high for spatially dispersed firms, financial analysts may encounter difficulties incorporating accrual reversals and estimating discretionary accruals for such firms compared to purely local firms. Another possible source of information asymmetry is the uncertainty regarding performance effects stemming from greater operating flexibility (Duru and Reeb, 2002). Managers enjoy discretion in shifting income across U.S. states and organizing activities across the U.S. (Dyreng et al., 2013), thus introducing difficulty in forecasting corporate actions and the associated effect on firm performance. Taken together, we expect problems arising from asymmetric information as reflected in the precision of analyst forecasts to be more severe for more dispersed firms.

Alternatively, geographic dispersion may reduce the information asymmetry between management and financial analysts. In the context of industrially diversified firms, Thomas (2002) suggests that asymmetric information regarding each segment's performance can be, in part, diversified away across segments (i.e. the information diversification hypothesis). We extend such arguments to the spatial distribution of a firm's economic activity in the U.S. and conjecture that the errors that outsiders make in forecasting cash flows generated in dispersed economic centers may be imperfectly correlated across U.S. economic centers. In this setting, even if the errors that outsiders make in forecasting cash flows of dispersed firms' economic centers are larger than the errors that they make in forecasting local firms' cash flows, the consolidated forecast may be more accurate. Moreover, the dispersion of economic activities may reduce the volatility of performance indicators when cash flows are imperfectly correlated across states in the U.S. (i.e. the portfolio hypothesis, Shapiro, 1978), further reducing analysts' forecasting errors.

Our first hypothesis addresses the relationship between the spatial dispersion of economic activities and the precision of financial analysts' forecasts. Consistent with the transparency argument, we conjecture that, *ceteris paribus*, earnings forecasts are less precise for geographically dispersed firms than for local firms. Hypothesis 1 (in alternative form) is presented as follows:

H1: Analysts make more (less) precise earnings forecasts for firms with less (more) geographically dispersed economic activities.

In the second set of hypotheses, we examine the relationship between geographic dispersion, financial information quality and performance indicator volatility. The comparability of financial statements enhances the precision of analysts' forecasts (De Franco et al., 2011), explaining the greater accuracy and lower dispersion of earnings forecasts for more comparable firms. Because the extent of opportunistic earnings management is increasing with the level of information asymmetry (Dye, 1986; Trueman and Titman, 1988; Richardson, 2000), we further expect earnings management to be more pronounced for geographically dispersed firms than for local firms. Therefore, we argue that, *ceteris paribus*, both comparability of financial statements and earnings quality is greater for local firms than for geographically dispersed firms. Hypothesis 2a (in alternative form) proposes the following:

H2a [Information quality effect]: Financial information quality, as proxied by accounting comparability and earnings quality, is higher (lower) for firms with less (more) geographically dispersed economic activities.

Alternatively, the precision of earnings forecasts may be determined by the riskiness of firms' activities rather than by higher levels of information asymmetry. Geographically dispersed firms may have lower forecast errors because the spatial distribution of economic activities decreases the volatility of performance indicators when the cash flows of multiple states are imperfectly correlated (i.e., the portfolio effect; Shapiro, 1978). Despite such benefits stemming from geographic dispersion, the exposure to state-specific macroeconomic factors and regulatory constraints may also increase performance volatility. To investigate the effect of geographic dispersion on the analysts' forecasting difficulties stemming from the volatility of performance metrics, we formulate Hypothesis 2b (in alternative form):

H2b [Portfolio effect]: Performance indicators are more (less) volatile for firms with less (more) geographically dispersed economic activities.

The presence of information asymmetry between management and outsides is likely to create a demand for disclosure, providing management incentives to increase disclosure because the value of additional information is greater in such settings (e.g., Grossman and Hart, 1980). In the context of international diversification, Burgstahler and Eames (2006) and Webb et al. (2008) argue that higher information asymmetry stemming from international activities strengthens the incentives to increase voluntary disclosure. Extending this argument to geographically dispersed firms, we may expect to find stronger incentives for managers of dispersed firms to reduce information asymmetry by voluntarily increasing disclosure. Alternatively, geographically dispersed firms may find it difficult to increase voluntary disclosure when obtaining precise information at the economic-center level is harder to collect and release on the market. The notion in the accounting literature is that the precision of a manager's private information determines the probability of disclosure increases (Verrecchia, 1990). Because the precision of private information may be lower for managers of dispersed firms, the incentives to increase voluntary disclosure may not be significantly different for dispersed firms. Furthermore, to the extent that aggregation problems may reduce the earnings quality of dispersed firms, the level of voluntary disclosure is likely to decrease for more dispersed firms (e.g., on the relation between earnings quality and voluntary disclosure, see Francis et al., 2008). To examine the relationship between geographic dispersion and voluntary disclosure, we formulate Hypothesis 3 (in alternative form):

H3 [Disclosure choice effect]: Voluntary disclosure of financial information is higher (lower) for firms with more (less) geographically dispersed economic activities.

3. Sample selection and methodology

3.1. Sample selection and primarily data sources

The sample selection process begins by including all firms at the intersection of the I/B/E/S Summary History file, the Compustat database, CRSP files and the GN data set (Garcia and Norli, 2012) over the 11-year period from 1997 to 2008. The sample period is determined by the availability of the geographic dispersion measure. The empirical tests exclude the following: (*i*) non-U.S. firms; (*ii*) firm-year observations with fewer than three earnings forecasts; (*iii*) financial and utility firms (firm-level SIC codes 6000-6999 and 4900-4999); and (iv) firms with negative total assets, negative stock prices and missing accounting information. Screening firms as described above results in a sample of 17,316 firm-year observations for 3,591 firms over the 1997–2008 period.

3.2. Variable definitions and supplementary data sources

3.2.1. Geographic dispersion, geographic similarity and geographic correlation

The degree of geographic dispersion of a firm's business operations is measured by Garcia and Norli (2012) using data from 10-K filings. Form 10-K is an annual report required by the SEC that provides a comprehensive summary of a public company's performance and operations. In addition to financial data, the annual report typically includes information on the evolution of the firm's operations during that year and details on its organizational structure, including information on the firm's properties. For example, firms may include sales at stores in different states and list the manufacturing facilities under their operation along with the city and state where those facilities are located.

Using a computerized parsing of 10-Ks, Garcia and Norli (2012) count the number of times each 10-K mentions a U.S. state's name in four main sections of the 10-K filings: "Item 1: Business", "Item 2: Properties", "Item 6: Consolidated Financial Data", and "Item 7: Management's Discussion and Analysis". Firms that do not mention any state in 10-Ks are excluded from the analysis. Addoum et al. (2014) also parse these four sections based on the argument that these sections summarize the locality of a firm's main business operations, including the firm's plants and equipment, major physical assets, store locations, office locations, and acquisition activities; therefore, the citation count measure captures the economic ties between a firm's headquarter and its geographically distributed economic interests. The states most frequently mentioned in the 10-Ks of the sampled firms are (% of firms) California (15.8%), Texas (7.8%), New York (6.9%), Delaware (6.2%), Massachusetts (4.3%) and Florida (3.8%). The least frequently mentioned states are Rhode Island, North Dakota and South Dakota (each state: 0.2%).

We construct the following concentration measure of geographic dispersion using the GN dataset with the state citations in 10-K filings (Garcia and Norli, 2012). To measure the degree of geographic dispersion across firms, we compute a normalized Herfindahl–Hirschman Index (HHI) of state activities as follows. First, we calculate the sum of the squared relative state citations:

$$SS_{i,t} = \left(\frac{\#Alabama_{i,t}}{\#Total\_US\ states_{i,t}}\right)^2 + \cdots + \left(\frac{\#New\ York_{i,t}}{\#Total\_US\ states_{i,t}}\right)^2 + \cdots + \left(\frac{\#Wyoming_{i,t}}{\#Total\_US\ states_{i,t}}\right)^2 \quad (1)$$

where $SS_{i,t}$ is the sum of the squared relative state counts for firm i in year t. Next, we obtain the normalized concentration as follows:

$$CONCENTRATION_{i,t} = \frac{SS_{i,t} - 1/50}{1 - 1/50} \quad (2)$$

If a firm's activities are exclusively concentrated in one state, *CONCENTRATION* is equal to one. In contrast, if a firm's activities are equally dispersed across the 50 U.S. states, the concentration index is equal to zero. Lower *CONCENTRATION* values indicate a firm's tendency to spread its business activities across a larger number of states.

To test the robustness of the empirical findings, we also construct simple measures of geographic dispersion. We employ the simple count of the number of U.S. states mentioned in the 10-K filings (*NSTATES*). Following Garcia and Norli (2012), firms are additionally classified as local firms if one or two states are mentioned in the annual reports, and zero otherwise (*LOCAL*).

We also construct a novel measure of geographic dispersion similarity, *GEOSIMILAR*. We calculate the pairwise similarity of any two firms in a particular industry group using the relative state citations and obtain a cosine similarity measure for each firm (See Appendix A for details on the variable construction).[1] This measure captures the extent to which a firm has an overlapping portfolio of economic centers with industry competitors defined at the three-digit SIC code level. Our measure is unique in that it represents the distances between firms in the 10-K based geographic dispersion space and indicates the degree to which a firm is different from industry rivals.

To further explore the cross-section variation in the geographic dispersion across U.S. states, we calculate a weighted-average correlation between local shocks across a firm's relevant economic states (*GEOCORR*; see Appendix B for details on the variable construction). We obtain archival economic time series at the state level for the period 1959–2013 from the Federal Reserve Bank of St. Louis, Archival Federal Reserve Economic Data (ALFRED). We detrend the economic series for each state and define the state-specific (local) shocks as the variance in state-level personal income per capita, which is not explained by the time trends and country-level fluctuations in personal income per capita. The correlation between local shocks is calculated over a 10-year window and weighted by the relative importance of each state for each firm. Higher values of *GEOCORR* indicate that either a firm concentrates its economic activities in one state or operates in states with highly correlated local shocks.

3.2.2. Financial analysts' forecast precision

We construct the measure of analyst earnings forecast accuracy as the negative of the absolute value of the analyst forecast (i.e., the difference between the firm's consensus forecast annual earnings per share (EPS) and the firm's actual EPS), deflated by the stock price at the beginning of the reporting period. We use forecasts from very near the end of the forecasting period (i.e., the IBES statistical period, statpers, ending in the last month of a firm's fiscal year, following Elton et al. (1984) and Easterwood and Nutt (1999)). We transform the variable such that higher values indicate higher accuracy:

$$ACCURACY_{i,t} = -1 \times \frac{\left|EPS_{i,t}^{Forecast} - EPS_{i,t}^{Actual}\right|}{P_{i,t-1}} \quad (3)$$

where $EPS_t^{Forecast}$ is the mean forecast of earnings per share and P_{t-1} is the price in period $t-1$.

DISPERSION is the cross-sectional standard deviation of individual analysts' annual forecasts for a given firm scaled by price at the beginning of the reporting period. *BIAS* is the signed forecast error:

$$BIAS_{i,t} = \frac{EPS_{i,t}^{Forecast} - EPS_{i,t}^{Actual}}{P_{i,t-1}} \quad (4)$$

where $EPS_t^{Forecast}$ is the median forecast of earnings per share and P_{t-1} is the price in period $t-1$.

3.2.3. Other variables

Several firm-specific factors are likely to determine analysts' incentives to acquire information or the complexity of the forecasting task. Empirical studies suggest that the demand for investment advice is significantly greater for larger firms, explaining greater coverage and improved information environment for such firms (Atiase, 1985; Freeman, 1987; King et al., 1990). We control for the quality of the information environment using size (*SIZE*), analyst coverage (*COVERAGE*), profitability (*LOSS*), abnormal earnings (*SUPRISE*), and growth opportunities (*MB*) (Hwang et al., 1996; Barth et al., 2001; Barron et al., 2002; Matsumoto, 2002). We include volatility of performance indicators, which is likely to increase the complexity of forecasting, namely earnings volatility (*EVOL*), cash flow volatility (*CFVOL*), and return volatility (*RETVOL*). Spatially dispersed firms may operate in several different industrial or geographic segments, which may contribute to the forecast precision (Duru and Reeb, 2002; Thomas, 2002). Using Compustat Segment Data, we control for firm diversification using the number of industry segments (*INDSEG*) (Thomas, 2002; Frankel et al., 2006; Franco et al., 2015) and the number of geographic segments (*INTSEG*) (Duru and Reeb, 2002; Doukas and Pantzalis, 2003). In robustness tests, we use the degree of geographic dispersion proxied by the relative importance of foreign income and the concentration of geographic sales at the firm level.

[1] Basic cosine similarity is a widely used method for evaluating textual similarity. In the finance literature, cosine similarity has recently been introduced by Hoberg and Phillips (2010), whereby product similarity is calculated by measuring overlapping portfolios of products listed in business descriptions of 10-K filings.

Table 1
Descriptive statistics and correlation matrix.

Variable	N	Mean	P50	St Dev.	P25	P75
Panel A. Descriptive statistics						
1. Geographic variables						
CONCENTRATION	17,136	0.34	0.29	0.23	0.17	0.45
NSTATES	17,136	10.01	7.00	9.69	4.00	12.00
LOCAL	17,136	0.12	0.00	0.33	0.00	0.00
GEOSIMILAR	13,843	0.25	0.22	0.16	0.12	0.36
GEOCORR	16,534	0.77	0.82	0.19	0.69	0.90
HEADCONC	17,034	0.39	0.35	0.27	0.17	0.58
NSUBS	11,922	5.42	3.00	6.24	2.00	6.00
2. Analysts' forecast precision						
ACCURACY	17,136	−1.06	−0.16	3.77	−0.51	−0.05
DISPERSION	17,136	0.55	0.10	1.95	0.04	0.28
BIAS	17,136	0.28	−0.03	2.63	−0.17	0.11
3. Firm characteristics						
SIZE	17,136	6.91	6.76	1.63	5.77	7.91
SURPRISE	17,136	1.70	0.50	4.37	0.20	1.21
COVERAGE	17,136	9.26	7.00	6.49	4.00	12.00
LOSS	17,136	0.25	0.00	0.43	0.00	0.00
MB	17,136	3.47	2.47	4.32	1.57	4.04
LEV	17,136	0.21	0.19	0.20	0.02	0.34
ROA	17,136	0.00	0.05	0.24	0.00	0.09
R&D	17,136	0.05	0.01	0.09	0.00	0.07
CFVOL	17,136	3.23	2.54	2.33	1.61	4.11
EVOL	17,136	2.29	1.26	2.76	0.64	2.71
RETVOL	17,136	14.04	11.95	8.72	8.37	17.27
INTSEG	15,193	3.04	2.00	2.26	2.00	4.00
INDSEG	16,075	1.65	1.00	1.06	1.00	2.00
4. Disclosure environment						
E_QUALITY	12,553	−5.57	−4.00	5.06	−6.76	−2.46
COMPARABILITY	11,253	−0.42	−0.19	0.84	−0.40	−0.10
IND_SCORE	15,193	−0.05	−0.06	0.21	−0.18	0.04
INT_SCORE	16,076	0.07	0.05	0.22	−0.10	0.27
SEG_RESTATE	15,193	0.16	0.00	0.36	0.00	0.00
LATE10K	17,136	0.03	0.00	0.18	0.00	0.00
LATE10K/Q	17,136	0.06	0.00	0.23	0.00	0.00

Panel B: Correlation matrix	*CONCENTRATION*	*NSTATES*	*LOCAL*	*ACCURACY*	*DISPERSION*	*BIAS*
CONCENTRATION	1.000					
NSTATES	−0.551*	1.000				
LOCAL	0.688*	−0.325*	1.000			
ACCURACY	−0.019*	0.037*	−0.005	1.000		
DISPERSION	0.033*	−0.053*	0.006	−0.779*	1.000	
BIAS	−0.009	0.001	−0.001	−0.573*	0.366*	1.000

Notes: This table provides summary statistics (Panel A) and correlation matrix (Panel B). The correlation is statistically significant at 5 (*) level of significance.

To estimate the effect of geographic dispersion on the properties of financial information, we construct and employ in further empirical analysis the following variables. First, we observe the degree of comparability of financial statements across firms in a given industry (*COMPARABILITY*). De Franco et al. (2011) calculate the degree of comparability by comparing the information content of financial statements for each firm *i*-firm *j* combination in a given industry defined at the two-digit SIC code. We obtain the comparability measure from the data library of Professor Verdi.[2] Higher *COMPARABILITY* values suggest that the financial statements of a particular firm are highly comparable to industry competitors. Second, we estimate the quality of reported earnings (*E_QUALITY*) following the estimation procedure used by Francis et al. (2004) and Veenman (2012). We take the negative of the earnings quality measure, where higher *E_QUALITY* values represent high-quality accruals.

To examine a manager's incentives to reduce information asymmetries, we construct several proxies of voluntary disclosure quality. SFAS 131 requires firms to present relevant business and geographic segments, which firms strategically disaggregate (Nagarajan and Sridhar, 1996; Botosan and Stanford, 2005; Bens et al., 2011; Hope et al., 2013). Following Franco et al. (2015), we construct a segment-disclosure index as the industry-adjusted percentage of 16 segment-level items the firm discloses in Compustat Segment Data. These 16 items include segment-level information regarding sales, operating income before depreciation, depreciation and amortization, operating income after depreciation, capital expenditures, total assets, equity in earnings, investments at equity, number of employees, research and development expenses, order backlog, export sales, pretax income, income before extraordinary items, net income, and operating profit. If a firm discloses more than one segment, we use the percentage of segment items disclosed averaged at the firm-year level. Similar to Franco et al. (2015), we assign sample firms to the high (low) index group if the firm's segment-disclosure index is above (below) the sample median. We also compare the disclosure of sale levels across geographic segments over time and identify restated geographic information. Analogously to earnings restatements (See more in Dechow et al., 2010), we argue that the restatement of geographic segment information indicates low-quality information.

Geographic dispersion of a firm is potentially important to internal information flow. Especially when information cannot be transferred through technological means and when information

[2] Source: http://www.mit.edu/rverdi/

is not easily verifiable, distance is likely to inhibit the flow of information. We expect the effect of dispersion on analysts' precision to be more pronounced for soft-information industries in which greater dispersion of economic activities is likely to affect the within-firm information asymmetries. Following Petersen and Rajan (2002) and Landier et al. (2009), we characterize firms as operating in soft or hard information environments using the industry-level distance between firms and the branch of the main lending institution. Petersen and Rajan (2002) find that the distance between banks and their borrowers has been increasing and further argue that the means of information collection are getting more impersonal with time. Using the National Survey of Small Business Finance (1993 and 2003), we observe the change in the distance between firms and banks in different industries.[3] Following Landier et al. (2009), for each survey year and two-digit SIC code, we compute the mean distance of firms to their primary lending institution and observe the change in distance over time. In hard-information industries, we would expect lenders to take advantage of technological developments and deregulation, leading to a greater distance between the bank lending office and the borrowing firm. We characterize firms as operating in a soft (hard) information environment if the change in the industry-level distance between the firm and the primary lending institution is below (above) the sample median.

3.3. Sample description

Panel A of Table 1 presents the descriptive statistics, and Panel B presents the correlation matrix. To reduce the effect of outliers, all variables are winsorized at the 1 and 99 percentiles. All variables are defined in Appendix C.

The mean *CONCENTRATION*, the citation-based HHI, is 0.34. A typical (median) U.S. firm is geographically present in seven U.S. states (*NSTATES*), which is comparable to the results reported by Garcia and Norli (2012) and Addoum et al. (2014). Approximately 12% of the firms in our sample have economic centers in less than three states. An average firm has a cosine similarity of geographically dispersed activities relative to its industry competitors (*GEOSIMILAR*) of 0.25, and the correlation between local shocks in states with relevant economic activities (*GEOCORR*) is 0.77. The concentration of economic activities in the headquarter state (*HEADCONC*) is 0.39. A typical (median) U.S. firm in the sample has subsidiaries in five different states (*NSUBS*).

The mean (median) forecast *ACCURACY* is 1.06% (0.16%) of the share price, which is consistent with existing studies. The mean (median) *DISPERSION* of analysts' forecasts is 0.55% (0.10%) of the share price, and the mean (median) unsigned forecast error, *BIAS*, is 0.28% (−0.03%).

The descriptive statistics regarding firm characteristics are comparable to those of other studies. Approximately nine analysts follow firms that have an average of approximately 21% debt out of total assets, −0.01% return on assets, a market value of equity three times larger than their book value, 1.70% earnings surprise, 2.3% earnings volatility, 3.23% cash flow volatility, and 14.0% return volatility. Approximately one-fourth of the observations contain negative earnings before the inclusion of extraordinary items during the sample period.

The disclosure environment is characterized by an average five-year standard deviation of the accrual model residuals, *E_QUALITY*, of 5.57%, consistent with Veenman (2012). The average accounting comparability, *COMPARABILITY*, is −0.42, indicating that an average error in quarterly earnings between a pair of comparable firms is 0.42% of market value. Following Franco et al. (2015), we construct a segment-disclosure index as the industry-adjusted percentage of 16 segment-level items the firm discloses in Compustat Segment Data. An average firm reports less (more) segment items in business (geographic) segment disclosure when compared to industry peers (*IND_SCORE* and *INT_SCORE*, respectively). Approximately two out of ten firms restate sale segment data reported in geographic segment files (*SEG_RESTATE*), and 3% (6%) of sampled firms file late 10-K (and 10-Q) filings (*LATE10K* and *LATE10K/10Q*, respectively).

4. Empirical results

4.1. The effect of geographically dispersed activities on analysts' forecast precision

In this section, we study the relationship between geographic dispersion and the precision of financial analysts' earnings forecasts. Hypothesis 1 predicts that the information environment of financial analysts is adversely affected by geographic dispersion. We estimate the effect of geographic dispersion on the properties of analysts' forecasts with the following baseline specification:

$$\begin{aligned} Precision_{i,t} = {} & \beta_0 + \beta_1 GEODISP_{i,t} + \beta_2 SIZE_{i,t} + \beta_3 SURPRISE_{i,t} \\ & + \beta_4 COVERAGE_{i,t} + \beta_5 LOSS_{i,t} + \beta_6 MB_{i,t} + \beta_7 R\&D_{i,t} \\ & + \beta_8 EVOL_{i,t} + \beta_9 CFVOL_{i,t} + \beta_{10} RETVOL_{i,t} \\ & + \beta_{11} INTSEG_{i,t} + \beta_{12} INDSEG_{i,t} + \varepsilon_{i,t} \end{aligned} \quad (5)$$

where *Precision* (analysts' forecast precision) is equivalent to *ACCURACY*, *DISPERSION* and *BIAS*. *GEODISP* is constructed from the state counts in the GN dataset and takes the following values: *CONCENTRATION* (HHI based on the economic activities across U.S. states), *NSTATES* (the number of states with economic activities), and *LOCAL* (an indicator variable for truly local firms). The control variables are defined in Appendix C. The baseline estimates are obtained with industry and year fixed effects and two-way clustered standard errors (i.e., by firm and by year).[4] The baseline model includes industry fixed effects using the 48 Fama-French industry classification.

Table 2 presents the baseline regression estimates. We tabulate three specifications, which include different sets of control variables. Panel A tabulates the estimated effect of geographic dispersion on forecast accuracy; Panel B, on forecast dispersion; and Panel C, on forecast bias. Higher values of *ACCURACY* (*BIAS*) signify that earnings forecasts are more accurate (positively biased). Higher values of *DISPERSION* suggest a high level of disagreement (variation) among analysts regarding future earnings. The existing literature indicates high dispersion with the use of private information before issuing earnings forecasts and the divergent interpretation of public disclosure (Kim and Verrecchia, 1994; Kandel and Pearson, 1995; De Franco et al., 2011).

All panels in Table 2 show similar results. The ability of financial analysts to produce more accurate, less dispersed and less biased forecasts is adversely affected by geographic dispersion. This result confirms Hypothesis 1, which predicted that the precision of

[3] The surveys are made available by the Federal Reserve Board (http://www.federalreserve.gov/pubs/oss/oss3/nssbftoc.htm). We use the sample weights provided with the data and the variables idist1 and h7_1 for 1998 and 2003, respectively. These variables measure the distance in miles from the main office of the firm to the office or branch of the bank's main lending institution.

[4] Following Petersen (2009), Gow et al. (2010), and Thompson (2011), we cluster standard errors on two dimensions to handle persistent common shocks in the panel. Persistent common shocks, like business cycles, can induce correlation between different analysts' forecasts in different years (e.g. exogenous changes in analysts' coverage following broker mergers or changes in information quality of corporate disclosure across business cycles). Because the assumption of a constant firm or time effect (i.e., time-invariant fixed effects) may not fully remove the dependence between observations, standard errors are clustered on two dimensions without assuming constant effects, thus reporting results that are not subject to the related estimation bias.

Table 2
Geographic dispersion and analysts' forecast precision.

	(1)	(2)	(3)	(4)	(5)
Panel A. Dependent variable: ACCURACY					
CONCENTRATION	0.461**	0.464***	0.368**		
	(2.39)	(2.47)	(2.26)		
NSTATES				−0.010***	
				(−2.86)	
LOCAL					−0.005
					(−0.05)
SIZE	0.316***	0.225***	0.267***	0.268***	0.261***
	(6.89)	(3.97)	(5.23)	(5.26)	(5.06)
SURPRISE	−0.028***	−0.008	−0.007	−0.007	−0.007
	(−3.17)	(−1.17)	(−0.95)	(−0.99)	(−0.96)
COVERAGE	0.001	0.005	−0.001	−0.001	−0.002
	(0.10)	(0.51)	(−0.16)	(−0.11)	(−0.18)
LOSS	−1.866***	−1.333***	−1.265***	−1.267***	−1.271***
	(−8.58)	(−6.19)	(−6.79)	(−6.78)	(−6.78)
MB	−0.010	0.006	0.004	0.004	0.005
	(−0.88)	(0.53)	(0.43)	(0.44)	(0.56)
R&D	−4.819***	−2.596**	−2.023	−1.969	−1.915
	(−3.42)	(−1.98)	(−1.55)	(−1.51)	(−1.45)
EVOL		−0.212***	−0.189***	−0.190***	−0.190***
		(−5.05)	(−4.99)	(−5.00)	(−5.02)
CFVOL		0.021	0.008	0.009	0.010
		(0.65)	(0.25)	(0.29)	(0.31)
RETVOL		−0.034***	−0.034***	−0.034***	−0.034***
		(−2.64)	(−2.51)	(−2.51)	(−2.49)
INTSEG			−0.025	−0.026	−0.023
			(−1.29)	(−1.32)	(−1.18)
INDSEG			−0.118***	−0.116***	−0.118***
			(−4.07)	(−4.02)	(−4.07)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes
N	17,136	17,136	14,391	14,391	14,391
R^2	0.125	0.146	0.138	0.138	0.137

Panel B. Dependent variable: DISPERSION	(1)	(2)	(3)	(4)	(5)
CONCENTRATION	−0.213**	−0.229***	−0.166**		
	(−2.36)	(−2.51)	(−2.31)		
NSTATES				0.003**	
				(2.28)	
LOCAL					−0.049
					(−0.87)
SIZE	−0.151***	−0.100***	−0.116***	−0.115***	−0.113***
	(−5.22)	(−3.14)	(−3.32)	(−3.31)	(−3.25)
SURPRISE	0.011**	0.003	0.004	0.004	0.004
	(1.94)	(0.50)	(0.65)	(0.67)	(0.65)
COVERAGE	−0.002	−0.004	−0.001	−0.001	−0.001
	(−0.38)	(−0.76)	(−0.11)	(−0.14)	(−0.12)
LOSS	0.813***	0.575***	0.556***	0.557***	0.558***
	(9.82)	(6.70)	(6.71)	(6.68)	(6.67)
MB	0.008	−0.000	−0.000	−0.001	−0.001
	(1.43)	(−0.03)	(−0.10)	(−0.13)	(−0.19)
R&D	3.818***	2.838***	2.094***	2.065***	2.059***
	(5.16)	(4.12)	(2.88)	(2.84)	(2.81)
EVOL		0.079***	0.065***	0.066***	0.066***
		(4.76)	(4.82)	(4.84)	(4.85)
CFVOL		0.001	0.001	0.001	0.001
		(0.09)	(0.11)	(0.05)	(0.06)
RETVOL		0.022***	0.023***	0.023***	0.022***
		(3.04)	(3.09)	(3.08)	(3.08)
INTSEG			0.001	0.002	0.001
			(0.15)	(0.16)	(0.06)
INDSEG			0.047***	0.047***	0.048***
			(3.39)	(3.35)	(3.38)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes
N	17,136	17,136	14,391	14,391	14,391
R^2	0.140	0.156	0.140	0.140	0.140

Table 2 (*continued*)

Panel C. Dependent variable: BIAS	(1)	(2)	(3)	(4)	(5)
CONCENTRATION	−0.289**	−0.271**	−0.221*		
	(−2.29)	(−2.06)	(−1.82)		
NSTATES				0.007***	
				(2.50)	
LOCAL					0.020
					(0.31)
SIZE	−0.128***	−0.127***	−0.167***	−0.168***	−0.164***
	(−5.32)	(−4.78)	(−7.40)	(−7.16)	(−6.96)
SURPRISE	0.013***	0.009**	0.006	0.006	0.006
	(3.21)	(2.25)	(1.25)	(1.28)	(1.28)
COVERAGE	0.005	0.005	0.009**	0.008*	0.009**
	(0.98)	(0.92)	(1.94)	(1.87)	(1.95)
LOSS	1.174***	1.097***	1.072***	1.072***	1.075***
	(7.16)	(7.96)	(9.36)	(9.34)	(9.30)
MB	0.008	0.006	0.006	0.006	0.005
	(1.30)	(1.07)	(0.97)	(1.02)	(0.87)
R&D	0.177	−0.173	−0.285	−0.314	−0.355
	(0.17)	(−0.16)	(−0.27)	(−0.29)	(−0.33)
EVOL		0.051**	0.031	0.031	0.032
		(2.25)	(1.33)	(1.36)	(1.37)
CFVOL		−0.014	−0.003	−0.004	−0.005
		(−0.80)	(−0.21)	(−0.26)	(−0.29)
RETVOL		−0.005	0.001	0.001	0.001
		(−0.62)	(0.19)	(0.17)	(0.15)
INTSEG			0.010	0.010	0.008
			(0.89)	(0.96)	(0.76)
INDSEG			0.078***	0.077***	0.078***
			(5.93)	(5.82)	(5.98)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes
N	17,136	17,136	14,391	14,391	14,391
R^2	0.055	0.056	0.060	0.060	0.059

Notes: This table presents the effect of a firm's geographic dispersion on analysts' accuracy (Panel A), dispersion (Panel B), and bias (Panel C). In Columns (1) to (3), we explain the variance in analysts' precision with a firm's geographic concentration. In Columns (4) and (5), we use the number of states with relevant economic activities and an indicator variable for local firms as a proxy for geographic dispersion. Three model specifications are tabulated. Models (1) and (3) explain the variance in analysts' precision with geographic dispersion after controlling for size, earnings surprises, analyst coverage, losses, market-to-book ratio, and research and development expenses. In model (2), we additionally control for cash flow volatility, earnings volatility and market return volatility. Model (3) further controls for corporate diversification. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***, **, * denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

financial analysts' forecasts is lower for more geographically dispersed firms. Empirical findings are robust to the inclusion of performance volatility indicators (Column (2)). Furthermore, results in Columns (3)–(5) suggest that the effect of geographic dispersion on analysts' precision differs from that of industrial and international diversification. Both geographic dispersion and across-country (industry) diversification may enable firms to exploit economies of scope (e.g., synergies between business centers or units in manufacturing, distribution or product promotion activities; Panzar and Willig, 1981). In contrast to within-country dispersion, international diversification may increase firms' exposure to systematic risk, including foreign exchange risk and political country risk (Reeb et al., 1998; Duru and Reeb, 2002). Moreover, home country institutional settings may represent a significant challenge to implementing diversification strategies (Wan and Hoskisson, 2003) and achieving a homogeneous level of information quality Leuz et al. (2003), further increasing analysts' difficulties in forecasting performance. By focusing on the within-country dispersion of economic activities, we examine the effect of geographic dispersion on analysts' precision in relatively homogeneous institutional (legal and political) settings. The results in Table 2 suggest that the spatial distribution of economic activities presents an additional difficulty in forecasting future performance, which is not attributable to a firm's exposure to largely different institutional settings or organisational complexity stemming from industrial diversification.

The estimated effect is both statistically and economically significant. Using the coefficient estimates in Column (3), we predict that a one-standard-deviation change in *CONCENTRATION* is associated with a 0.085 increase in accuracy, which represents a 8% improvement in a firm's accuracy.[5] A lower number of economic centers across U.S. states, *NSTATES*, is also associated with more accurate earnings forecasts. A one-standard-deviation increase in *NSTATES* is predicted to decrease average accuracy by 9.1%. The baseline results in Table 2 suggest that 'local' firms do not benefit from more precise analyst forecasts.[6]

Table 3 presents the empirical analysis using the degree of similarity in the geographic distribution patterns within a given industry. Higher values of *GEOSIMILAR* indicate that a firm's distribution of relevant economic activities across the U.S. is similar to that of industry competitors defined at the three-digit SIC level. By allowing analysts to use information from comparable firms as an additional input in their earnings forecasts, we expect greater similarity in the spatial distribution of relevant economic centers within an industry to facilitate the forecasting process (i.e., the benchmarking effect).[7] Moreover, as the number of competitors with economic activities in a given state increases, it is likely that

[5] The marginal effect is obtained as follows: 0.079 = (0.368 × 0.23)/1.06, where the denominator is the average accuracy.

[6] A supplementary analysis, however, shows that the results are not robust to the definition of 'localness'. Forecast accuracy and dispersion (accuracy and bias) are determined by the quartile local proxy (*LOCAL_PR*) (the top–bottom local measure). See more in the Online Supplementary Material. In further analysis, we focus on the HHI-based proxy, which contains a richer set of information regarding a firm's interstate dispersion in comparison to a simple (count or indicator) variable. By incorporating information about a state's relative importance to a given firm, the HHI-based measure more precisely captures interstate dispersion.

[7] The predicted effect of geographic similarity on analysts' forecasts is similar to the effect of accounting comparability in De Franco et al. (2011).

Table 3
Geographic similarity, geographic correlation and analysts' forecast precision.

Dependent variable	*ACCURACY* (1)	*DISPERSION* (2)	*BIAS* (3)	*ACCURACY* (4)	*DISPERSION* (5)	*BIAS* (6)
GEOSIMILAR	0.394	−0.206*	−0.304**			
	(1.60)	(−1.69)	(−2.02)			
GEOCORR				0.464***	−0.291***	−0.146
				(2.51)	(−2.80)	(−1.15)
SIZE	0.250***	−0.114***	−0.161***	0.268***	−0.119***	−0.166***
	(4.80)	(−3.20)	(−7.55)	(5.14)	(−3.37)	(−6.91)
SURPRISE	−0.007	0.004	0.005	−0.003	0.001	0.003
	(−0.93)	(0.72)	(0.87)	(−0.38)	(0.15)	(0.71)
COVERAGE	−0.001	−0.000	0.010**	−0.003	0.000	0.009**
	(−0.10)	(−0.05)	(2.10)	(−0.37)	(0.01)	(1.97)
LOSS	−1.276***	0.569***	1.068***	−1.245***	0.539***	1.041***
	(−6.97)	(6.66)	(10.30)	(−6.21)	(6.26)	(8.33)
MB	0.006	−0.001	0.005	0.004	−0.000	0.006
	(0.67)	(−0.19)	(0.73)	(0.42)	(−0.08)	(0.93)
R&D	−2.191*	2.128***	−0.269	−2.219*	2.133***	−0.187
	(−1.68)	(2.93)	(−0.25)	(−1.67)	(2.91)	(−0.17)
EVOL	−0.180***	0.061***	0.036	−0.187***	0.068***	0.029
	(−5.07)	(4.90)	(1.38)	(−4.91)	(4.95)	(1.28)
CFVOL	0.003	0.003	−0.006	0.007	0.001	−0.002
	(0.09)	(0.22)	(−0.32)	(0.22)	(0.08)	(−0.14)
RETVOL	−0.035***	0.023***	0.001	−0.035**	0.023***	0.002
	(−2.46)	(2.98)	(0.16)	(−2.45)	(2.99)	(0.25)
INTSEG	−0.025	0.003	0.010	−0.023	−0.001	0.011
	(−1.22)	(0.27)	(0.89)	(−1.14)	(−0.13)	(0.94)
INDSEG	−0.116***	0.045***	0.075***	−0.109***	0.043***	0.070***
	(−3.59)	(3.02)	(4.91)	(−3.57)	(2.96)	(5.16)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	13,843	13,843	13,843	13,878	13,878	13,878
R^2	0.137	0.141	0.060	0.138	0.141	0.057

Notes: This table presents the effect of geographic similarity and geographic correlation on analysts' precision (Columns (1) to (3) and Columns (4) to (6), respectively). We estimate the marginal effects after controlling for relevant firm characteristics, including performance indicator volatility and corporate diversification. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***,**,* denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

analysts are more likely to find in any public disclosure or private communications relevant information, such as benefits and costs of operating in such locations, thereby improving forecast precision (i.e., the information effect). Taken together, we expect such effects of geographic similarity to reduce the cost of information acquisition and processing, increasing analysts' precision.

Columns (1) to (3) present the regression coefficients of geographic similarity obtained after controlling for known determinants of forecast precision, including industrial and international diversification. We find that the dispersion and bias of analysts' forecasts is significantly lower for firms with a more similar distribution of activities across U.S. states. Using the coefficients in Column (2), we predict that a one-standard-deviation increase in *GEOSIMILAR* is associated with a decrease of approximately 6% in the average dispersion. Empirical results are consistent with the notion that a firm's unique distribution of relevant economic activities generates different analysts' beliefs about future prospects in spread economic centers, leading to greater forecast dispersion. The results in Column (3) suggest that geographic similarity is also associated with less biased forecasts, with an estimated effect of approximately 17% improvement in the average bias. Ke and Yu (2006) argues that analysts use biased earnings forecasts to curry favor with firm management to obtain better access to management's private information and find the currying-favor effect to be stronger for firms whose earnings are harder to forecast. The results are consistent with such argument in that there may be a lower demand for private information for more geographically comparable firms, explaining less pronounced favor-currying behavior using biased forecasts for such firms.

In Columns (4) to (6) of Table 3, we report the estimated effect on analysts' precision of the correlation of economic trends in states with relevant economic activities. Higher values of *GEOCORR* indicate that relevant centers of economic activities are located in states whose local shocks are more positively correlated. Because the demand to obtain additional information regarding economic trends in multiple locations is likely to be lower for highly correlated states, we expect analysts' precision to be higher for firms with dispersed activities in states with correlated local economic shocks. The results suggest that both accuracy and dispersion are significantly determined by the degree of co-movements of economic shocks in multiple states. The estimated effect is both statistically and economically significant. We predict that a one-standard-deviation change in geographic correlation is associated with an improvement in forecast accuracy of approximately 8.3% in the average accuracy and a reduction of approximately 10.1% in the average dispersion.

A possible concern is that geographic dispersion and volatility in performance indicators, a significant determinant of forecast precision, are inherently correlated. Because controlling for volatility in multivariate analysis may not account for the possibility that dispersed and local firms have different distributions of volatility, we carry out further empirical analysis using control groups. In Panel A of Table 4, we report the results obtained using exactly matched pairs. To test the robustness of the results, we also use propensity score matching to identify control firms, which are identical in important aspects but differ in geographic dispersion (Panel B of Table 4). The exact matching is performed using size and cash flow volatility at the three-digit SIC code level, and the treatment (control) group is drawn from the top (bottom) quartile of the concentration distribution. Because the exact matching is based on three characteristics, we have a significant number of firm-years for which a control was not possible to identify. To overcome such a limitation, we additional conduct a propensity score matching. The propensity score matching is the probability of

Table 4
Geographic dispersion and analysts' precision: Matched-pair analysis and Heckman correction for sample selection.

Panel A. Exact matched-pair sample using size and volatility			
Dependent variable	*ACCURACY* (1)	*DISPERSION* (2)	*BIAS* (3)
CONCENTRATION	0.727***	−0.399***	−0.285***
	(3.96)	(−3.87)	(−2.89)
Other controls	Yes	Yes	Yes
Year & Industry FE	Yes	Yes	Yes
N	2,422	2,422	2,422
R^2	0.162	0.146	0.084
Panel B. Matched sample using propensity score mathing			
	(1)	(2)	(3)
CONCENTRATION	0.610***	−0.273***	−0.492***
	(3.69)	(−3.19)	(−4.05)
Other controls	Yes	Yes	Yes
Year & Industry FE	Yes	Yes	Yes
N	7,115	7,115	7,115
R^2	0.159	0.171	0.081
Panel C1. Heckman correction. First-stage regression estimates			
Variable	Coefficient (1)	*t*-statistics (2)	Marginal effect (3)
SIZE	−0.081***	(−5.12)	−0.024
ROA	−0.022	(−0.37)	−0.006
MB	0.021***	(5.78)	0.006
CAPEX	0.156*	(1.74)	0.047
CFVOL	0.034***	(3.40)	0.010
EVOL	−0.018**	(−2.45)	−0.005
CONC_IND	5.326***	(13.52)	1.600
Year & Industry FE	Yes		
N	14,376		
Pseudo *R2*	0.097		
Panel C2. Heckman correction. Second-stage regression estimates			
Dependent variable	*ACCURACY* (1)	*DISPERSION* (2)	*BIAS* (3)
CONCENTRATION	0.399***	−0.173**	−0.264**
	(2.49)	(−2.30)	(−2.29)
INV.MILLS	0.146	0.026	−0.244**
	(0.75)	(0.30)	(−2.00)
Other controls	Yes	Yes	Yes
Year & Industry FE	Yes	Yes	Yes
N	14,376	14,376	14,376
R^2	0.136	0.136	0.060

Notes: This table presents the effect of geographic concentration on analysts' precision using matched sample analysis (Panels A and B), and after correcting for a possible sample selection bias (Panel C). Panel A tabulates the results obtained using an exact matched pair sample. We match firms on size, cash flow volatility and industry. Panel B tabulates the estimates obtained following a propensity score matching procedure. The probability of receiving treatment (i.e., control sample) is calculated using size, book-to-market, cash flow volatility, earnings volatility and industry. Panel C tabulate both first and second stage results. At the first stage, we explain the probability of being geographically dispersed with size, book-to-market, return on assets, capital intensity, cash flow volatility, earnings volatility, industry-level geographic concentration levels, and industry fixed effects. At the second stage, we correct for the self-selection bias by including the inverse mills ratio. All specifications include the following controls: size, earnings surprises, analyst coverage, losses, market-to-book ratio, research and development expenses, cash flow volatility, earnings volatility, market return volatility, and corporate (industrial and geographic) diversification. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***, **, * denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

receiving treatment conditional on a set of observed covariates, namely, size, book-to-market, cash flow volatility, earnings volatility and industry assignation. In this empirical part, all control variables in the baseline model, including industrial and international diversification, are part of the regression model. The matched-sample analysis further supports the prediction that geographic dispersion has a significant effect on analysts' precision, which is not explained by the possible differences in volatility.

Because the spread in economic activities across multiple states is a choice variable, we test the implications of geographic dispersion on analysts' precision correcting for a possible sample selection bias (i.e., Heckman correction). In Panel C1, we report the results obtained from the first stage regression model. Firm-specific characteristics influence the decision of firms to disperse activities across the U.S. We expect that (1) firms with low profitability in their current operations may diversify into other states in search of more lucrative opportunities, (2) firms with a high level of investment in current operations may be less likely to disperse across U.S. states, and (3) firms with high volatility in cash flows and earnings may prefer to enter other states to benefits from non-correlated economic shocks across the U.S. Following selection models in corporate finance (e.g. Campa and Kedia, 2002), we explain firms' localness with size, book-to-market, return on assets, capital intensity, and cash flow and earnings volatility. We additionally include the average geographic dispersion in a given industry (defined at the 2-digit SIC level) to capture time-variant industry-specific trends towards more (less) concentrated economic activities in the U.S. We run a logit model, where the dependent variable is an indicator for a local firm (i.e., top quartile of concentration), and zero otherwise. The marginal effects suggest that industry trends and capital intensity significantly increase the probability of being local. In Panel C2, we tabulate the second stage results. The insignificant coefficient of the Inverse Mills ratio for forecast accuracy and dispersion suggests that the sample selection bias is not driving the effect of geographic dispersion.

Empirical findings so far are consistent with the notion that the less transparent and more discretional aggregation of geographically dispersed information into consolidated reports increases the information asymmetry between management and financial analysts, reducing precision of earnings forecasts. Addoum et al. (2014) suggest that managers may not be efficient in aggregating and reporting value-relevant information regarding centers of business activities. Such aggregation difficulties may be surprising when surveyance and other data gathering techniques, reducing the cost of information collection and monitoring of division-level performance, are becoming more commonly employed in trend analysis.[8] Petersen and Rajan (2002) suggest that innovations in information technology explain the decreasing distance between primarily lenders and borrowing firms, and Landier et al. (2009) find that the decrease is more pronounced in some industries than in others. The notion in the literature is that in certain environments, information cannot be made impersonal cheaply (i.e., hardened), explaining the predominant trend to lend to proximate firms in such settings. Following Landier et al. (2009), we classify industries based on the changes in the distance between borrowing firms and primary lenders and expect the effect of geographic dispersion on analysts' accuracy to be more pronounced for firms in soft-information environments.

We tabulate the results for soft- and hard-information environments in Table 5. Consistent with the expectation that soft-information environments are characterized by more severe information asymmetry problems, we find that the sensitivity of analysts' precision to geographic dispersion is significantly higher in such environments. The spread of economic centers across multiple locations, therefore, is likely to add an additional layer of forecasting difficulty in environments where information asymmetries

[8] We would like to thank the anonymous referee for noting the importance of technological advances in monitoring. See examples about business applications of technology in Rothfeld and Patterson (2013) and Cameron (2014) (Source: Wall Street Journal; http://www.wsj.com/articles/SB10001424052702303497804579240182187225264; http://www.wsj.com/articles/digitalglobe-cleared-to-sell-sharper-images-1402519967).

Table 5
Geographic dispersion and analysts' precision in soft and hard information environments.

Dependent variable	*ACCURACY*		*DISPERSION*		*BIAS*	
Information environment:	*SOFT* (1)	*HARD* (2)	*SOFT* (3)	*HARD* (4)	*SOFT* (5)	*HARD* (6)
CONCENTRATION	0.564***	0.008	−0.290***	0.027	−0.271*	−0.070
	(2.73)	(0.04)	(−2.97)	(0.27)	(−1.88)	(−0.52)
SIZE	0.297***	0.247***	−0.101**	−0.140***	−0.220***	−0.122***
	(4.36)	(3.72)	(−2.21)	(−3.38)	(−6.46)	(−3.06)
SURPRISE	0.004	−0.006	−0.009	0.008	0.000	0.006
	(0.38)	(−0.37)	(−1.36)	(0.82)	(0.06)	(0.90)
COVERAGE	0.000	−0.014	−0.001	0.006	0.009	0.015**
	(0.04)	(−1.27)	(−0.20)	(1.13)	(1.36)	(2.10)
LOSS	−1.607***	−0.869***	0.675***	0.386***	1.389***	0.728***
	(−7.29)	(−4.83)	(6.98)	(3.70)	(8.73)	(6.39)
MB	−0.007	0.017	0.000	−0.002	0.024*	−0.011
	(−0.52)	(1.04)	(0.03)	(−0.23)	(1.70)	(−0.95)
EVOL	−0.206***	−0.175***	0.069***	0.059***	0.028	0.044
	(−4.33)	(−4.04)	(3.43)	(3.01)	(0.97)	(1.28)
CFVOL	0.049	−0.042	−0.008	0.017	−0.020	0.009
	(1.40)	(−0.97)	(−0.55)	(0.98)	(−1.08)	(0.35)
RETVOL	−0.023*	−0.051***	0.021**	0.026***	−0.003	0.008
	(−1.63)	(−3.44)	(2.07)	(4.82)	(−0.50)	(0.74)
R&D	−1.759	−2.412	1.793	2.392***	−2.210	0.818
	(−0.53)	(−1.45)	(1.10)	(2.47)	(−1.45)	(0.69)
INTSEG	−0.028	−0.022	0.005	0.002	−0.012	0.029
	(−1.02)	(−0.77)	(0.45)	(0.14)	(−1.15)	(1.51)
INDSEG	−0.156***	−0.087***	0.056**	0.044***	0.122***	0.029
	(−3.18)	(−2.49)	(2.11)	(2.54)	(4.60)	(1.05)
t-stats 'soft' vs. 'hard'	1.87**		−2.12**		−1.08	
(*p*-value)	(0.06)		(0.03)		(0.28)	
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	6,982	6,666	6,982	6,666	6,982	6,666
R^2	0.150	0.140	0.141	0.148	0.081	0.050

Notes: This table presents the effect of geographic concentration on analysts' precision across sub-samples (*SOFT* and *HARD*). We characterize firms as operating in a soft (hard) information environment if the change in the industry-level distance between the firm and the primary lending institution is below (above) the sample median. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. *t*-stats and *p*-value refer to the test for significance in the effects of geographic dispersion on analysts' precision in soft versus hard information environments. ***, **, * denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

– between the headquarters and economic centers or between managers and outsiders – are already high.

4.2. Geographic dispersion, financial information quality and performance volatility

In this section, we empirically examine the effect of geographic dispersion on information content in earnings and properties of performance indicators. First, we directly test for the possibility that geographically dispersed firms have less comparable and more discretionary manipulated earnings. Next, we investigate the association between the spatial distribution of economic activities and the volatility of earnings, cash flows and market returns.

4.2.1. Geographic dispersion and financial information quality

We conjecture that greater information asymmetry between management and financial analysts contributes to less precise earnings forecasts for geographically dispersed firms. In Table 6, we test the effect of geographic dispersion on financial information quality, measured by accounting comparability and discretionary accruals manipulations. We tabulate three model specifications, including variables for industrial and international diversification. Following Francis et al. (2004) and De Franco et al. (2011), we control for cash flow and market return volatility.

The empirical results in Table 6 support the prediction that financial information in the earnings of geographically dispersed firms is less comparable and more discretionarily managed than that of truly local firms, thus likely contributing to the lower precision of earnings forecasts for geographically dispersed firms. Although lower comparability and earnings quality may be associated with a firm's risk profile, we find that the effect is significant after controlling for cash flow and market return volatility. Moreover, regression estimates suggest that the effect of geographic dispersion on earnings comparability and quality is not stemming from industrial and international diversification. The effect of geographic concentration is robust to the inclusion of control for the degree of corporate diversification. Using the coefficients in Columns (1) and (4), we predict that a one-standard-deviation change in concentration is associated with an improvement of approximately 2.5% in the average comparability and of 3.4% in the average earnings quality. Local firms are estimated to have on average more comparable and higher quality earnings (i.e., using the coefficients in Columns (3) and (6), the estimated effect is 2.9% and 3.1%, respectively).

4.2.2. Geographic dispersion and volatility

We also conjecture that geographic dispersion of relevant economic activities affects the volatility of performance indicators, thus determining the precision of analysts' forecasts. On one side, managers of geographically dispersed firms may efficiently reduce volatility by extending economic activities across multiple states (i.e., the portfolio effect in Shapiro, 1978). On the other side, environmental factors may introduce additional risk stemming from more dispersed economic activities (Duru and Reeb, 2002).

In Table 7, we present the empirical link between performance volatility and geographic dispersion. Consistent with the portfolio effect, we find empirical evidence, albeit not robust, that cash flow and market return volatility are higher for more concentrated firms

Table 6
Geographic dispersion, comparability and earnings quality.

Dependent variable	*COMPARABILITY* (1)	*COMPARABILITY* (2)	*COMPARABILITY* (3)	*E_QUALITY* (4)	*E_QUALITY* (5)	*E_QUALITY* (6)
CONCENTRATION	0.046*			0.819**		
	(1.82)			(2.34)		
NSTATES		0.001			0.003	
		(0.67)			(0.52)	
LOCAL			0.037***			0.519***
			(2.63)			(2.54)
INTSEG	−0.001	−0.001	−0.001	−0.055*	−0.050	−0.052*
	(−0.37)	(−0.24)	(−0.33)	(−1.75)	(−1.58)	(−1.67)
INDSEG	−0.017	−0.017*	−0.018*	0.013	0.013	0.011
	(−1.60)	(−1.63)	(−1.63)	(0.22)	(0.22)	(0.18)
SIZE	0.005	0.004	0.005	0.192***	0.176***	0.182***
	(0.91)	(0.70)	(0.84)	(3.57)	(3.27)	(3.36)
CASHFLOW	−0.255***	−0.256***	−0.254***	−3.563***	−3.581***	−3.551***
	(−2.58)	(−2.62)	(−2.58)	(−3.79)	(−3.79)	(−3.80)
CFVOL	−0.002	−0.002	−0.003	−0.489***	−0.486***	−0.488***
	(−0.45)	(−0.38)	(−0.45)	(−9.09)	(−9.02)	(−8.98)
LOSS	−0.144***	−0.144***	−0.144***	−1.009***	−1.018***	−1.013***
	(−5.20)	(−5.23)	(−5.23)	(−5.11)	(−5.23)	(−5.14)
MB	0.003*	0.004**	0.003*	−0.053***	−0.050***	−0.051***
	(1.80)	(1.94)	(1.85)	(−2.93)	(−2.78)	(−2.87)
R&D	0.406***	0.419***	0.410***	−9.702***	−9.508***	−9.622***
	(3.13)	(3.19)	(3.18)	(−5.21)	(−5.11)	(−5.21)
LEVERAGE	−0.148***	−0.156***	−0.150***	−0.656	−0.751*	−0.702*
	(−3.23)	(−3.35)	(−3.24)	(−1.56)	(−1.80)	(−1.69)
ROA	−0.073	−0.071	−0.073	3.682***	3.707***	3.688***
	(−0.81)	(−0.78)	(−0.80)	(3.19)	(3.19)	(3.20)
RETVOL	−0.002*	−0.002	−0.002	−0.078***	−0.077***	−0.077***
	(−1.64)	(−1.59)	(−1.59)	(−3.41)	(−3.35)	(−3.40)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	9,414	9,414	9,414	10,571	10,571	10,571
R^2	0.683	0.683	0.683	0.307	0.306	0.307

Notes: This table presents the empirical link between geographic concentration, financial statement comparability (Columns (1)–(3)), and earnings quality (Columns (4) and (5)). The controls include size, cash flow, cash flow volatility, losses, market-to-book, leverage, research and development expenses, return on assets, and return volatility. All other specifications additionally include proxies for corporate diversification. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***, **, * denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

(Columns (2) and (3)). Because geographic dispersion and volatility are significantly associated, jointly determining analysts' precision, the results further strengthen the motivation to include risk-related controls in the baseline model (Tables 2, 3, and 5) and to conduct matched sample analysis (Table 4).

4.3. Geographic dispersion and disclosure choices

The predicted adverse effect of geographic dispersion on analysts' forecasts implicitly assumes that firms with multiple centers of economic activities do not attempt to reduce information asymmetry by providing voluntary information to the market, or if they do, such efforts do not effectively translate into lower asymmetry problems. In Table 8, we empirically test the prediction that the demand for voluntary disclosure is higher for geographically dispersed firms, explaining the higher quantity and quality of voluntary disclosure by such firms. Consistent with Botosan and Stanford (2005), we expect more extensive segment disclosure to facilitate analysts' forecasting activities and expect that dispersed firms, in search of transparency, have more detailed segment disclosure than that of industry competitors. Following Franco et al. (2015), we compare the level of segment disclosure to industry competitors and predict that in such settings, geographic concentration is negatively associated with disclosure scores. In Columns (1) and (3), we assign disclosure scores in deciles and estimate the effect of concentration using a Tobit model. Higher values of IND_SCORE_D (INT_SCORE_D) indicate that a firm tabulates a greater number of business (geographic) segment items in comparison to industry competitors. In Columns (2) and (4), we assign firms to two groups based on the sample median segment disclosure index, IND_SCORE_M and INT_SCORE_M (i.e., the dependent variable is equal to one for firms with disclosure scores above the sample median, and zero otherwise).

Contrary to the prediction in Hypothesis 3, we find that firms with less dispersed economic activities are more likely to provide a higher quantity of segment information. Because we obtain coefficient estimates after controlling for industrial and international diversification, we attribute the effect to the spatial distribution and not to corporate diversification. Verrecchia (1990) advances the argument that only precise and credibly disclosed voluntary information reduces market asymmetries. Although the demand for information may be higher for firms with geographically dispersed activities, managers may not have credible signals to voluntarily communicate on the market. For instance, firms may not be capable of obtaining precise information at the economic-center level, thus explaining the possibility for estimation errors in segment disclosure (e.g., accruals and tax provisions in multiple, dispersed locations). This interpretation is consistent with Jennings and Tanlu (2014), who find that managers indeed find it difficult to provide high-quality guidance when business, international and cost structures are complex. Moreover, extensive segment disclosure may not effectively reduce information asymmetry for firms, when earnings information is low quality (Francis et al., 2008). The results in Table 6 suggest that earnings quality is adversely affected by geographic dispersion, likely reducing the credibility of voluntarily disclosed information.

In Table 8, we also report the empirical link between geographic dispersion and other disclosure variables, including the release of high-quality segment information and the delay in filing 10-K/10-Q reports. The results in Column (5) suggest that the likelihood of restating sale segment information is higher for firms with dispersed activities. Columns (6) and (7) present empirical

Table 7
Geographic dispersion and performance indicator volatility.

	EVOL (1)	*CFVOL* (2)	*RETVOL* (3)	*EVOL* (4)	*CFVOL* (5)	*RETVOL* (6)	*EVOL* (7)	*CFVOL* (8)	*RETVOL* (9)
CONCENTRATION	−0.132	0.254**	1.087**						
	(−1.21)	(2.29)	(2.37)						
NSTATES				−0.001	−0.001	−0.012			
				(−0.50)	(−0.32)	(−0.94)			
LOCAL							−0.140**	0.101	0.290
							(−2.09)	(1.41)	(0.92)
INTSEG	0.031**	−0.031***	0.113***	0.030**	−0.030**	0.115***	0.031**	−0.030**	0.118***
	(2.02)	(−2.53)	(2.82)	(1.94)	(−2.44)	(2.82)	(2.00)	(−2.44)	(2.92)
INDSEG	−0.026	−0.013	−0.496***	−0.025	−0.013	−0.493***	−0.025	−0.014	−0.497***
	(−1.07)	(−0.43)	(−4.90)	(−1.07)	(−0.42)	(−4.86)	(−1.03)	(−0.45)	(−4.93)
SIZE	−0.192***	−0.356***	−1.056***	−0.188***	−0.360***	−1.067***	−0.191***	−0.360***	−1.075***
	(−9.98)	(−16.68)	(−10.34)	(−9.56)	(−16.75)	(−10.20)	(−9.93)	(−16.61)	(−10.86)
CASHFLOW	0.542	−1.073***	−3.685***	0.544	−1.075***	−3.693***	0.537	−1.071***	−3.685***
	(0.81)	(−3.25)	(−2.73)	(0.82)	(−3.25)	(−2.71)	(0.81)	(−3.23)	(−2.71)
LOSS	1.427***	−0.022	3.431***	1.428***	−0.024	3.425***	1.427***	−0.023	3.425***
	(12.00)	(−0.38)	(5.89)	(11.98)	(−0.42)	(5.90)	(12.03)	(−0.40)	(5.90)
MB	0.070***	0.080***	0.167***	0.070***	0.081***	0.169***	0.070***	0.081***	0.170***
	(8.70)	(11.66)	(3.98)	(8.71)	(11.76)	(4.06)	(8.72)	(11.72)	(4.02)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	14,391	14,391	14,391	14,391	14,391	14,391	14,391	14,391	14,391
R^2	0.410	0.314	0.410	0.410	0.313	0.409	0.410	0.314	0.409

Notes: This table presents the empirical link between geographic dispersion and performance indicator volatility. In Columns (1) to (3), we explain the variance in performance indicator volatility with geographic concentration. In Columns (4) to (6), we present the results obtained using the number of states. In Columns (7) to (9), an indicator variable for truly local firms is used as a proxy for geographic dispersion. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. *** and ** denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

Table 8
Geographic dispersion, segment disclosure quality and late filings of financial reports.

Dependent variable	*IND_SCORE$_D$* (1)	*IND_SCORE$_M$* (2)	*INT_SCORE$_D$* (3)	*INT_SCORE$_M$* (4)	*SEG_RESTATE* (5)	*LATE10K* (6)	*LATE10K/Q* (7)
CONCENTRATION	0.337*	−0.132	0.764***	0.572***	−0.300*	−0.492**	−0.471**
	(1.81)	(−0.99)	(4.05)	(3.49)	(−1.88)	(−2.04)	(−2.36)
SIZE	−0.177***	0.051*	−0.228***	−0.189***	0.019	−0.296***	−0.261***
	(−4.08)	(1.74)	(−5.58)	(−5.44)	(0.54)	(−4.77)	(−5.28)
SURPRISE	−0.005	0.006	−0.018***	−0.010**	−0.001	0.013*	0.013**
	(−0.86)	(1.61)	(−3.42)	(−2.39)	(−0.09)	(1.65)	(1.98)
COVERAGE	−0.005	0.007	0.033***	0.025***	0.003	−0.018	−0.012
	(−0.65)	(1.11)	(3.70)	(3.50)	(0.40)	(−1.20)	(−1.01)
MB	0.019**	−0.011*	0.039***	0.027***	0.002	−0.020	−0.024*
	(2.02)	(−1.84)	(4.71)	(4.06)	(0.24)	(−1.24)	(−1.92)
CFVOL	0.006	0.019	0.005	0.002	−0.005	−0.039	−0.049**
	(0.26)	(1.35)	(0.25)	(0.10)	(−0.26)	(−1.22)	(−2.01)
EVOL	−0.025	0.022**	0.011	0.020	−0.010	0.042**	0.042***
	(−1.45)	(2.03)	(0.68)	(1.53)	(−0.71)	(2.16)	(2.65)
RETVOL	0.001	0.000	0.001	0.004	0.004	0.022***	0.024***
	(0.34)	(0.00)	(0.19)	(1.03)	(0.97)	(2.76)	(3.20)
INTSEG	−0.478***	0.333***	−0.092***	−0.127***	0.247***	0.052**	0.053***
	(−14.83)	(12.06)	(−3.44)	(−5.95)	(14.67)	(2.37)	(2.86)
INDSEG	−0.260***	0.222***	−1.298***	−2.282***	0.142***	−0.099	−0.070
	(−5.07)	(6.39)	(−21.89)	(−21.18)	(3.68)	(−1.40)	(−1.32)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model	Tobit	Logit	Tobit	Logit	Logit	Logit	Logit
N	14,391	14,391	14,391	14,391	14,358	14,156	14,279
Pseudo R^2	0.033	0.099	0.056	0.327	0.091	0.146	0.129

Notes: This table presents the empirical link between geographic concentration and management disclosure choices. In Columns (1) to (4), we explain the choice to increase segment disclosure with geographic concentration. In Column (5), we estimate the likelihood of restating segment disclosure as a function of geographic concentration. In Columns (6) and (7), we relate geographic concentration to the likelihood of issuing annual and quarterly filings with a delay. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***, **, * denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

evidence that dispersed firms are more likely to file late annual and quarterly reports.[9] Taken together, the empirical findings in Table 8 suggest that the quality and quantity of segment disclosure is higher for less dispersed firms, likely contributing to the higher precision of analysts' forecasts.

5. Additional empirical analysis and robustness tests

In Table 9, we employ alternative definitions for geographic dispersion, such as the concentration of economic activities in the

[9] Although the filing of 10-K and 10-Q reports is requested by the SEC for publicly listed firms, it is a management choice to strictly follow the deadlines and present relevant financial information in a timely manner. For this reason, we tabulate the results obtained for late filings along with empirical tests regarding the quantity and quality of segment information.

Table 9
Additional analysis: Headquarters, subsidiaries and analysts.

Dependent variable	*ACCURACY* (1)	*DISPERSION* (2)	*BIAS* (3)	*ACCURACY* (4)	*DISPERSION* (5)	*BIAS* (6)
HEADCONC	0.490***	−0.244***	−0.160			
	(3.89)	(−4.14)	(−1.36)			
NSUBS				−0.021***	0.008***	0.015***
				(−3.63)	(2.83)	(3.71)
SIZE	0.277***	−0.118***	−0.171***	0.245***	−0.080***	−0.182***
	(5.48)	(−3.43)	(−7.75)	(4.70)	(−2.56)	(−6.35)
SURPRISE	−0.004	0.002	0.004	−0.018*	0.008	0.008
	(−0.49)	(0.39)	(0.83)	(−1.67)	(0.99)	(1.23)
COVERAGE	−0.004	0.000	0.010**	0.009	−0.009***	0.006
	(−0.47)	(0.00)	(2.19)	(1.16)	(−2.50)	(1.15)
LOSS	−1.232***	0.529***	1.046***	−1.279***	0.531***	1.126***
	(−6.37)	(6.29)	(8.66)	(−4.82)	(5.09)	(7.67)
MB	0.003	0.000	0.006	−0.003	0.002	0.008
	(0.30)	(0.02)	(0.98)	(−0.24)	(0.25)	(1.44)
EVOL	−0.188***	0.066***	0.032	−0.204***	0.068***	0.041
	(−5.09)	(4.91)	(1.45)	(−4.34)	(3.58)	(1.07)
CFVOL	0.004	0.004	−0.002	−0.007	0.007	−0.001
	(0.13)	(0.31)	(−0.11)	(−0.27)	(0.70)	(−0.05)
RETVOL	−0.035***	0.022***	0.001	−0.020*	0.018**	−0.004
	(−2.54)	(3.10)	(0.17)	(−1.80)	(2.31)	(−0.48)
R&D	−2.173*	2.186***	−0.251	−1.473	1.343*	−0.825
	(−1.66)	(3.05)	(−0.24)	(−0.90)	(1.86)	(−0.58)
INTSEG	−0.025	−0.000	0.012	−0.025	0.005	0.006
	(−1.23)	(−0.03)	(1.02)	(−1.22)	(0.54)	(0.49)
INDSEG	−0.112***	0.045***	0.074***	−0.046*	0.009	0.057***
	(−3.77)	(3.21)	(5.36)	(−1.67)	(0.65)	(3.59)
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	14,302	14,302	14,302	10,020	10,020	10,020
R^2	0.139	0.141	0.059	0.147	0.143	0.069

Notes: This table tabulates additional empirical tests. We explain the precision of analysts' forecasts with the concentration of relevant activities in the headquarter state (Columns (1) to (3)) and with the presence of subsidiaries in different states (Columns (4) to (6)). The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***,**,* denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively. Variables are defined in Appendix C.

Table 10
Robustness tests.

	Variable	*CONCENTRATION*	*t*-stats	*N*	R^2
a. GDP-weighted concentration	*ACCURACY*	0.332**	2.415	14,391	0.138
	DISPERSION	−0.178**	−2.553	14,391	0.140
	BIAS	−0.166*	−1.894	14,391	0.060
b. Controlling for earnings skewness	*ACCURACY*	0.509***	2.958	10,040	0.158
	DISPERSION	−0.222***	−2.980	10,040	0.156
	BIAS	−0.280**	−2.122	10,040	0.065
c. Excluding retail industry (SIC 52-59)	*ACCURACY*	0.380**	2.136	12,960	0.137
	DISPERSION	−0.174**	−2.223	12,960	0.140
	BIAS	−0.229*	−1.704	12,960	0.056
d. Industry adj. *CONCENTRATION*	*ACCURACY*	0.359**	2.132	14,391	0.138
	DISPERSION	−0.139*	−1.879	14,391	0.140
	BIAS	−0.236**	−2.017	14,391	0.060
e. Using 2-digit codes	*ACCURACY*	0.378**	2.343	14,391	0.138
	DISPERSION	−0.172**	−2.275	14,391	0.140
	BIAS	−0.219*	−1.802	14,391	0.060
f. Using state fixed effects	*ACCURACY*	0.305*	1.813	14,356	0.150
	DISPERSION	−0.133*	−1.802	14,356	0.152
	BIAS	−0.164	−1.494	14,356	0.068
g. Using foreign income as a proxy for geographic diversification	*ACCURACY*	0.373**	2.267	14,391	0.138
	DISPERSION	−0.169**	−2.395	14,391	0.140
	BIAS	−0.230*	−1.917	14,391	0.060
h. Using HHI sale concentration as a proxy for geographic diversification	*ACCURACY*	0.376**	2.269	14,386	0.138
	DISPERSION	−0.172**	−2.376	14,386	0.140
	BIAS	−0.215*	−1.711	14,386	0.060

This table presents the robustness tests. All specifications include the following controls: size, earnings surprises, analyst coverage, losses, market-to-book ratio, research and development expenses, cash flow volatility, earnings volatility, market return volatility, and corporate industrial and geographic diversification. The estimates are obtained with industry and year fixed effects. We tabulate two-way clustered standard errors. ***, **, * denote statistically significant coefficients at the 1%, 5% and 10% levels, respectively.

headquarter state and the number of subsidiaries across multiple states. Management and outsiders may not be efficient in collecting and assimilating information originating from dispersed centers of economic activities (Giroud, 2013; Addoum et al., 2014; Bernile et al., 2015). In such settings, we expect the concentration of activities in the headquarter state to increase analysts' forecast precision. We count the number of times a firm's headquarter state is mentioned in the 10-K filing and divide it to total state count (*HEADCONC*). All control variables of the baseline specification are included. The results in Columns (1) to (3) confirm the reported positive effect of geographic concentration on analysts' precision.

We next consider the dispersion of a firm's subsidiaries across the U.S. Our prediction is that locating subsidiaries in differen states increases the difficulties in aggregating information, thus making forecasting more difficult. We observe the location of firms' subsidiaries and expect interstate dispersion of subsidiary activities to determine analysts' information environment. The empirical results suggest that forecast precision is adversely affected by multi-state location of subsidiary activities (Columns (4) to (6)).

In Table 10, we tabulate additional robustness tests. We present the following variations in the estimation model:

(a) *GDP-weighted concentration*. We weight the relative importance of economic centers by the state-level gross domestic product. It is likely that business activities in some states - for instance, growing states – have a differential effect on firm performance in aggregate. While the data do not indicate the fraction of a firm's operations by state, we assess robustness of results to weighting the state citations by state-level GDP. The GDP-weighted concentration measure gives more weight to larger and economically growing states, which are likely to have a larger contribution to a firm's overall financial performance. The GDP-weighted concentration measure confirms the economic significance of geographic dispersion in forming analysts' expectations.

(b) *Controlling for earnings skewness*. Negatively (positively) skewed earnings increase forecasting difficulties, leading to more negative (positive) forecast bias (Duru and Reeb, 2002). We estimate earnings skewness over a period of 10 consecutive years and expect highly (positively or negatively) skewed earnings to be negatively associated with analysts' precision. The significant effect of geographic concentration on forecast precision is robust to the inclusion of earnings skewness in the estimation model.

(c) *Excluding retail industry (SIC 52-59)*. We test the robustness of empirical findings by excluding firms from the retail sector. It is possible that analysts do not encounter forecasting more difficult for firms whose business model requires more dispersed organizational activities, such as retail firms. Our findings are robust to the exclusion of retail firms whose activities are expected to be dispersed over the U.S. states.

(d) *Industry-adjusted concentration*. We assess the robustness of geographic dispersion to adjusting interstate dispersion for industry standards. We adjust the concentration measure by subtracting the average concentration at the industry level (two-digit SIC codes). We expect that financial analysts follow several firms in a particular industry and face forecasting difficulties for firms that have more dispersed activities in comparison to industry peers. Empirical evidence supports the prediction that (industry-adjusted) geographic concentration significantly affects forecast precision.

(e) and (f) *Using 2-digit or state fixed effects*. We include 2-digit SIC code groups to test the robustness of our results obtained using Fama-French industry classification. Additionally, we include state fixed effects to capture state-specific shocks to analysts' forecasts. Regression coefficients are not affected by the selection of fixed effects.

(g) and (h) *Using alternative controls for international diversification*. We consider two alternative proxies for international diversification. In part g., we control for the relative importance of foreign income (i.e. foreign income divided by total assets). In part h., we calculate the Herfindahl–Hirschman Index of sales in different geographic segments using Compustat segment data. Empirical findings are robust to such alternative model specifications and confirm the prediction that analysts' precision is lower for geographically dispersed firms.

6. Conclusions and implications

Using a text-based measure of geographic dispersion, which captures the economic ties between a firm and its geographically distributed economic interests (Garcia and Norli, 2012; Addoum et al., 2014), we provide empirical evidence that financial analysts issue less accurate, more dispersed and more biased forecasts for geographically dispersed firms. Empirical findings are consistent with the notion that a less transparent and more discretional aggregation of geographically dispersed information into financial reports increases the information asymmetry between management and financial analysts. The empirical evidence further suggests that geographic dispersion is also less likely to reduce forecast precision when a firm has economic activities in states with highly correlated local shocks. Consistent with the information asymmetry argument, we find that geographically dispersed firms have less comparable and more discretionary managed earnings, have less extended segment information than that of the industry competitors, are more likely to restate sale segment information, and issue annual and quarterly filings with a delay. Additionally, we show that the effect of geographic dispersion is more pronounced for soft-information environments where information is more difficult to make impersonal by using technological advances.

This study contributes to the current research on corporate geography by suggesting that management and financial analysts may not be prepared to efficiently aggregate information for more dispersed centers of economic activities. Using information regarding within-country or within-industry variation in corporate activities, we show that geographic dispersions represent a significant challenge in analysts' forecasting and attribute such results to the greater information asymmetries stemming from the dispersion of relevant information across multiple U.S. states.

Acknowledgments

We would like to thank Diego Garcia for sharing the dataset with the state counts in 10-K filings. The paper has significantly benefited from the comments and recommendations of Calin Arcalean, Emanuele Bajo, Antonio Matacena, Stefano Mengoli, Fabrizio Palmucci, and anonymous referees. Petya Platikanova gratefully acknowledges financial support from Government of Catalonia (Grant 2014-SGR-1079), Banc Sabadell and la Caixa.

Appendix A. Cosine similarity based on the spatial distribution of economic activities.

In this section, we describe the calculation and present an example of how the cosine similarity is computed. Consider an industry with three firms and suppose that the three firms have the distribution of relevant economic activities across the U.S. states as follows:

U.S. state a	Arizona	California	Delaware
Firm i	0.4	0.6	0.0
Firm j	0.0	0.8	0.2
Firm k	0.0	0.5	0.5

To compute a Firm i's similarity to its peers, we first calculate the angular separation similarity measure between Firm i and Firm j as the cosine of the angle between two vectors:

$$Similarity_{ij} = \sum_{a=1}^{3} \frac{x_{ia}x_{ja}}{\left(\sum_{a=1}^{3} x_{ia}^2 \sum_{a=1}^{3} x_{ja}^2\right)^{1/2}} \tag{6}$$

where a takes the values of the relative frequency of a firm's economic activities in Arizona, California and Delaware.

$$Similarity_{ij} = \frac{0.48}{(0.52 \times 0.68)^{1/2}} = 0.807 \tag{7}$$

The cosine similarity takes values in the range (0, 1). High similarity values suggest that the two firms have similar spatial distributions of relevant economic activities. In this example, both Firm i and Firm j have a significant concentration of economic activities in California, explaining the relatively high similarity between the two firms.

We next calculate the cosine similarity between Firm i and Firm k (0.589) and take the average value in the industry group defined by the three-digit SIC codes, excluding the firm itself (by construct, the self-similarity equals one). Because our sample is homogeneous (i.e., firms with analyst followings), we do not use weights and calculate a simple average of the pair-wise similarity measures. In empirical analysis, we use the aggregate measure of fluidity which in this example for Firm i is equal to 0.698.

For comparison, consider the spatial distribution of economic activities of Firm j. Both Firm i and Firm j have a significant concentration of economic activities in California, translating into a high similarity between the two firms (calculated above, 0.807). Moreover, Firm j and Firm k have both relevant activities in the same states (i.e., a cosine similarity of 0.857). The overall similarity of Firm j is 0.832. We conclude that Firm j is more comparable to industry competitors in comparison to Firm i.

Appendix B. Detrended economic series and geographic correlations.

In this section, we present the estimation procedure followed to obtain the weighted average correlation between the headquarter state of a firm i and U.S. states with relevant economic activities.

We first estimate the trend in the economic series data of a state j as follows (Wooldridge, 2009):

$$\begin{aligned} Personal\ Income_{j,t} = \alpha + \beta_1 Personal\ Income_{j,t-1} \\ + \beta_2 Personal\ Income_{U\ S,t} + \beta_3 t + \epsilon_t \end{aligned} \tag{8}$$

where $Personal\ Income_j$ is the personal income per capita in state j, $Personal\ Income_j$ is the aggregate personal income per capita at the country level, and t is a time variable (t = 1, 2, ..., 55). Archival economic time series at the state level are available from the Federal Reserve Bank of St. Louis, Archival Federal Reserve Economic Data (ALFRED). We estimate the trend in personal income per capita using data for the period 1959–2013. We do not use gross domestic product data because such economic series are available at the state level only for the period 1997–2013. We obtain the residual (i.e., detrended economic series) for each state j (j = 1, 2, ..., 50). The residual captures the deviations from the predicted trend ('local shocks'), which are not explained by the last year's level of personal income in the state, the trends over time, or the macroeconomic trends at the country level.

We next calculate the correlation between state j and state k using a rolling window of 10 years. We weight the calculated correlation using the relative count of US states in the annual filings and merge the pair-wise correlations between the headquarter state and the state j with the main dataset.

A firm with economic activities only in its headquarter states obtains a weighted correlation equal to 1. A firm with significant activities in a state whose local shocks are not perfectly correlated with the headquarter state has a weighted correlation less than 1. Higher weighted correlations suggest that the firm has a high concentration of activities in the headquarter state and/or in other states whose local shocks are highly correlated with those in the headquarter state.

Appendix C. Variable definitions

Variable	*Definition*	*Data source*
1. Geographic variables		
CONCENTRATION	A normalized Herfindahl–Hirschman Index (HHI) of state activities. Using computerized parsing of 10-Ks, the relative importance of a state to a firm is defined as the number of times a U.S. state is mentioned in 10-K filings, divided by the total number of times all U.S. states are mentioned in the filing	Garcia and Norli (2012)
NSTATE	A count variable equal to the number of states mentioned in the 10-K filings	Garcia and Norli (2012)
LOCAL	An indicator variable equal to one for truly local firms (i.e., firms with one or two states mentioned in the annual 10-K filings), and zero otherwise	Garcia and Norli (2012)
GEOSIMILAR	A cosine similarity of a firm's geographic dispersion to its industry peers measured at three-digit SIC level	See Appendix A for details

Appendix C (*continued*)

Variable	*Definition*	*Data source*
GEOCORR	The average correlation between the economic shocks of a firm's headquarter state and other states with relevant economic activities, weighted by the state citation. An economic shock is the variance in a state's personal income (per capita) level, which is not explained by time trends and country-level fluctuations	See Appendix B for details
HEADCONC	A count variable equal to the citations of a firm's headquarter state in the 10-K filing, divided by the total number of state citations in the filing	Garcia and Norli (2012)
NSUBS	A count variable equal to the number of different states where a firm's subsidiaries are located	Dyreng et al. (2013)
2. Analysts' forecast precision		
ACCURACY	The absolute difference between the mean earnings per share forecast (meanest) and the actual earnings per share (actual), divided by the stock price at the beginning of the reporting period (prc). The variable is multiplied by minus one such that higher values stand for more precise forecasts	I/B/E/S Summary Statistics and CRSP
DISPERSION	The standard deviation of the earnings per share forecast (stdev), divided by the stock price at the beginning of the reporting period (prc)	I/B/E/S Summary Statistics and CRSP
BIAS	The signed forecast error of analysts' forecasts, which is the difference between the median earnings per share forecast (medianest) and the actual earnings per share (actual), divided by the stock price at the beginning of the reporting period (prc)	I/B/E/S Summary Statistics
3. Firm characteristics		
INTSEG	A count variable equal to the number of a firm's geographic segments	COMPUSTAT
INDSEG	A count variable equal to the number of a firm's unique four-digit business segments	COMPUSTAT
SIZE	A logarithmic transformation of a firm's market value of equity ($prcc\_f \times csho$) at the end of the reporting period	COMPUSTAT
SURPRISE	The absolute difference between annual earnings (ib) at time t and annual earnings at time $t-1$, divided by the annual earnings at time $t-1$	COMPUSTAT
COVERAGE	A count variable equal to the number of analysts issuing earnings per share forecast (numest)	I/B/E/S Summary Statistics
LOSS	An indicator variable equal to one for negative actual earnings per share before extraordinary items (epspx), and zero otherwise	COMPUSTAT
MB	The market value of equity($prcc\_f \times csho$), divided by the book value of equity (ceq)	COMPUSTAT
LEV	Total debt (dltt plus dlc), divided by total assets (at)	COMPUSTAT
ROA	The income before extraordinary items (ib), divided by the total assets at the end of the reporting period (at)	COMPUSTAT
R&D	Research and development expense (xrd), divided by total assets (at)	COMPUSTAT
CFVOL	The standard deviation of cash flows from operating activities over the past 12 quarters. Cash flows from operating activities (using date-to-date oancfy) is divided by total assets in a given quarter (atq)	COMPUSTAT
EVOL	The standard deviation of returns on assets over the past 12 quarters. Returns on assets is defined as the income before extraordinary income (ibq) to total assets (atq) in a given quarter	COMPUSTAT
RETVOL	The standard deviation of stock returns over the past 12 months (ret)	CRSP

(*continued on next page*)

Appendix C (*continued*)

Variable	Definition	Data source
4. Disclosure environment		
E_QUALITY	The standard deviation over the past three to five years of the firm-specific residuals of the modified Dechow and Dichev (2002) model and estimated following Veenman (2012). The variable is multiplied by minus one such that higher values indicate for high quality earnings	COMPUSTAT
COMPARABILITY	A comparability measure of a firm's financial statements to four closest peers, as defined by De Franco et al. (2011)	Verdi's Data Library
IND_SCORE	A disclosure score capturing the level of disclosure in business segment data; The industry-adjusted percentage disclosure of 16 segment items in the business segment files	COMPUSTAT
INT_SCORE	A disclosure score capturing the level of disclosure in geographic segment data; The industry-adjusted percentage disclosure of 16 segment items in the geographic segment files	COMPUSTAT
SEG_RESTATE	An indicator variable equal to one for a restatement in geographic sale segment information, and zero otherwise	COMPUSTAT
LATE10K	An indicator variable equal to one for a delayed filing of 10-K forms, and zero otherwise	EDGAR
LATE10K/Q	An indicator variable equal to one for a delayed filing of 10-K or 10-Q firms, and zero otherwise	EDGAR

Appendix D. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbankfin.2015.11.012.

References

Abarbanell, J., Lehavy, R., 2003. Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts' earnings forecasts. Journal of Accounting and Economics 36 (1–3), 105–146.
Addoum, J., Kumar, A., Law, K., 2014. Geographic diffusion of information and stock returns. Working papers, University of Miami.
Atiase, R., 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. Journal of Accounting Research 23 (1), 21–36.
Barron, O., Byard, D., Kile, C., Riedl, E., 2002. High-technology intangibles and analysts forecasts. Journal of Accounting Research 40 (2), 289–312.
Barth, M., Kasznik, R., McNichols, M., 2001. Analyst coverage and intangible assets. Journal of Accounting Research 39 (1), 1–34.
Bens, D., Berger, P., Monahan, S., 2011. Discretionary disclosure in financial reporting: an examination comparing internal firm data to externally reported segment data. The Accounting Review 86 (2), 417–449.
Bernile, G., Kumar, A. and Sulaeman, J., 2015. Home away from home: geography of information and local investors. Review of Financial Studies.
Botosan, C., Stanford, M., 2005. Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on analysts' information environment. The Accounting Review 80 (3), 751–772.
Burgstahler, D., Eames, M., 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. Journal of Business Finance & Accounting 33 (5–6), 633–652.
Cameron, D., 2014. DigitalGlobe Cleared to Sell Sharper Images. Wall Street Journal.
Campa, J., Kedia, S., 2002. Explaining the diversification discount. Journal of Finance 57 (4), 1731–1762.
Coles, J., Hertzel, M., Kalpathy, S., 2006. Earnings management around employee stock option reissues. Journal of Accounting and Economics 41 (12), 173–200.
Coval, J., Moskowitz, T., 1999. Home bias at home: local equity preference in domestic portfolios. The Journal of Finance 54 (6), 2045–2073.
Dechow, P., Dichev, D., 2002. The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review 77, 35–59.
Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: a review of the proxies, their determinants and their consequences. Journal of Accounting and Economics 50 (23), 344–401.
De Franco, G., Kothari, S., Verdi, R., 2011. The benefits of financial statement comparability. Journal of Accounting Research 49 (4), 895–931.
Doukas, J., Pantzalis, C., 2003. Geographic diversification and agency costs of debt of multinational firms. Journal of Corporate Finance 9 (1), 59–92.
Duru, A., Reeb, D., 2002. International diversification and analysts' forecast accuracy and bias. The Accounting Review 77 (2), 415–433.
Dye, R., 1986. Proprietary and nonproprietary disclosures. The Journal of Business 59 (2), 331–366.
Dyreng, S., Lindsey, B., Thornock, J., 2013. Exploring the role Delaware plays as a domestic tax haven. Journal of Financial Economics 108 (3), 751–772.
Easterwood, J., Nutt, S., 1999. Inefficiency in analysts' earnings forecasts: systematic misreaction or systematic optimism? The Journal of Finance 54 (5), 1777–1797.
Elton, E., Gruber, M., Gultekin, M., 1984. Professional expectations: accurary and diagonosis of errors. Journal of Financial and Quantitative Analysis 19, 351–363.
Francis, J., LaFond, R., Olsson, P., Schipper, K., 2004. Costs of equity and earnings attributes. The Accounting Review 79 (4), 967–1010.
Francis, J., Nanda, D., Olsson, P., 2008. Voluntary disclosure, earnings quality, and cost of capital. Journal of Accounting Research 46 (1), 53–99.
Franco, F., Urcan, O., Vasvari, F., 2015. The value of corporate diversification: a debt market perspective. Working papers, London Business School.
Frankel, R., Kothari, S., Weber, J., 2006. Determinants of the informativeness of analyst research. Journal of Accounting and Economics 41 (12), 29–54.
Freeman, R., 1987. The association between accounting earnings and security returns for large and small firms. Journal of Accounting and Economics 9 (2), 195–228.
Garcia, D., Norli, O., 2012. Geographic dispersion and stock returns. Journal of Financial Economics 106 (3), 547–565.
Giroud, X., 2013. Proximity and investment: evidence from plant-level data. The Quarterly Journal of Economics 128 (2), 861–915.
Gow, I., Ormazabal, G., Taylor, D., 2010. Correcting for cross-sectional and time-series dependence in accounting research. The Accounting Review 85 (2), 483–512.
Grossman, S., Hart, O., 1980. Disclosure laws and takeover bids. The Journal of Finance 35 (2), 323–334.
Florian, H., Ljungqvist, A., 2015. As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes. Journal of Financial Economics 118 (3), 684–712.
Hau, H., 2001. Location matters: an examination of trading profits. The Journal of Finance 56 (5), 1959–1983.
Hoberg, G., Phillips, G., 2010. Product market synergies and competition in mergers and acquisitions: a text-based analysis. Review of Financial Studies 23 (10), 3773–3811.

Hope, O.-K., Ma, M., Thomas, W., 2013. Tax avoidance and geographic earnings disclosure. Journal of Accounting and Economics 56 (23), 170–189.
Hwang, L., Jan, C., Basu, S., 1996. Loss firms and analysts' earnings forecast errors. Journal of Financial Statement Analysis 1, 18–30.
Jennings, J., Tanlu, L., 2014. The effect of organizational complexity on earnings forecasting behavior. Working papers, Washington University in St. Louis.
Kandel, E., Pearson, N., 1995. Differential interpretation of public signals and trade in speculative markets. Journal of Political Economy 103 (4), 831–872.
Ke, B., Yu, Y., 2006. The effect of issuing biased earnings forecasts on analysts' access to management and survival. Journal of Accounting Research 44 (5), 965–999.
Kim, O., Verrecchia, R., 1994. Market liquidity and volume around earnings announcements. Journal of Accounting and Economics 17 (12), 41–67.
King, R., Pownall, G., Waymire, G., 1990. Expectations adjustment via timely management forecasts: review, synthesis, and suggestions for future research. Journal of Accounting Literature 9, 113–144.
Landier, A., Nair, V.B., Wulf, J., 2009. Trade-offs in staying close: corporate decision making and geographic dispersion. Review of Financial Studies 22 (3), 1119–1148.
Leuz, C., Nanda, D., Wysocki, P.D., 2003. Earnings management and investor protection: an international comparison. Journal of Financial Economics 69 (3), 505–527.
Litov, L., Moreton, P., Zenger, T., 2012. Corporate strategy, analyst coverage, and the uniqueness paradox. Management Science 58 (10), 1797–1815.
Malloy, C., 2005. The geography of equity analysis. The Journal of Finance 60 (2), 719–755.
Matsumoto, D., 2002. Management's incentives to guide analysts' forecasts. The Accounting Review 77 (3), 483–514.
Nagarajan, N., Sridhar, S., 1996. Corporate responses to segment disclosure requirements. Journal of Accounting and Economics 21 (2), 253–275.
Panzar, J., Willig, R., 1981. Economies of scope. The American Economic Review 71 (2), 268–272.
Petersen, M., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Review of Financial Studies 22 (1), 435–480.
Petersen, M., Rajan, R., 2002. Does distance still matter? The information revolution in small business lending. The Journal of Finance 57 (6), 2533–2570.
Reeb, D.M., Kwok, C.C.Y., Baek, H.Y., 1998. Systematic risk of the multinational corporation. Journal of International Business Studies 29 (2), 263–279.
Richardson, V., 2000. Information asymmetry and earnings management: some evidence. Review of Quantitative Finance and Accounting 15 (4), 325–347.
Rothfeld, M., Patterson, S., 2013. Traders Seek an Edge With High-Tech Snooping. Wall Street Journal.
Shapiro, A., 1978. Financial structure and the cost of capital in the multinational corporation. The Accounting Review 13, 211–266.
Thomas, S., 2002. Firm diversification and asymmetric information: evidence from analysts forecasts and earnings announcements. Journal of Financial Economics 64 (3), 373–396.
Thompson, S., 2011. Simple formulas for standard errors that cluster by both firm and time. Journal of Financial Economics 99 (1), 1–10.
Trueman, B., Titman, S., 1988. Studies on management's ability and incentives to affect the timing and magnitude of accounting accruals. Journal of Accounting Research 26, 127–139.
Veenman, D., 2012. Disclosures of insider purchases and the valuation implications of past earnings signals. The Accounting Review 81 (1), 313–342.
Verrecchia, R., 1990. Information quality and discretionary disclosure. Journal of Accounting and Economics 12 (4), 365–380.
Verrecchia, R., 2001. Essays on disclosure. Journal of Accounting and Economics 32 (1–3), 97–180.
Wan, W., Hoskisson, R., 2003. Home country environments, corporate diversification strategies, and firm performance. Academy of Management Journal 46 (1), 27–45.
Webb, K., Cahan, S., Sun, J., 2008. The effect of globalization and legal environment on voluntary disclosure. The International Journal of Accounting 43 (3), 219–245.
Wooldridge, Jeffrey M., 2009. Introductory Econometrics: A Modern Approach. South Western, Cengage Learning, Mason, OH.