

Internet Appendix for
“What a Relief: How do Firms Respond to Competitors’ Listing Delays?”
(Not to be published)

This Internet Appendix describes how we identify a firm’s product market competitors based on its product descriptions in the annual report (in the business description section) using text-based analysis, the results of a number of validation tests, and examples of firms and their close competitors.

IA.1 Identification of Product Market Competitors

Our identification of product market competitors is based on the degree of product similarity between two firms, as proposed by Hoberg and Phillips (2016). The core idea is that closeness in the product space reflects the degree to which a given firm is similar to its competitors and the level of competition between them.

Following Hoberg and Phillips (2016), we begin by downloading the annual reports of all public firms (excluding financial firms and special treatment firms) between 2003 and 2017 from the disclosure platform designated by the China Securities Regulatory Commission (CSRC) (<http://www.cninfo.com.cn/new/index>). The “Report of the Directors” section of the annual report provides detailed business descriptions of the firm. If this section is not available, we use the “Overview of Business” section instead. The CSRC stipulates that firms must provide accurate and up-to-date business descriptions in the annual report that reflect the business reality of the year. This enables us to capture the dynamic changes in the firm’s product composition and industry presence across the years.

After identifying the text related to the firm’s business descriptions, we parse all words using the THULAC tool and exclude common words as well as words that are not nouns or proper nouns. Then we manually select and keep words that are related to the firm’s business and products, which results in a database (i.e., product market space) of 12,802 unique words. We vectorize the product market vocabulary so that each year a given firm i ’s vocabulary is represented by a vector, with each element being populated by one if firm i uses the given word, and zero if it does not. We then assign pairwise similarity scores based on the cosine similarity between two firms’ vectorized product descriptions. The cosine similarity between two firms (i and j) is higher when two firms’

product market descriptions are more similar. This similarity measure ranges from 0 (no similarity) to 1 (perfect similarity).

$$\text{Product Cosine Similarity}_{i,j} = (V_i \cdot V_j) / (|V_i| \cdot |V_j|)$$

Following Hoberg and Phillips (2016), we label firms i and j as a “membership pair” if, for a given classification, they are in the same industry. We determine the threshold-generating industries with the same fraction of membership pairs as Wind-3 industries (the Chinese equivalent of the SIC3 industries), so that we can compare our industries to Wind-3 industries in an unbiased fashion. The mean threshold value for our sample is 0.2893. Firm j is classified as a competitor of firm i if the similarity score between them is higher than the threshold value.

IA.2 Validation Tests

To ensure our measure properly identifies a firm’s product market competitors, we conduct a series of validation tests. We construct a variable *LnCosMean*, which measures the average level of competition facing a firm in a year and is calculated as the natural logarithm of one plus the average pairwise similarities between the firm and all of its rivals in a given year. The average value of *LnCosMean* is 0.296 with a standard deviation of 0.019 for our sample.

IA.2.1 Competition and disclosure

Hoberg and Phillips (2016) show that firms with more rivals, especially rivals with high similarity, are more likely to disclose information about competitive pressure in the Management’s Discussion and Analysis section of their 10-Ks. We test whether this relationship holds for our competition measure (i.e., *LnCosMean*). Following the spirit of Hoberg and Phillips (2016) and for each firm-year, we construct a measure of competition disclosure, *CompFrequency*, defined as the frequency of the word “competition” in the annual report as a percentage of total words.

Table IA.1 reports the results of the relationship between *LnCosMean* and *CompFrequency*. The coefficient on *LnCosMean* is positive and significant, which is consistent with the findings of Hoberg and Phillips (2016) that firms facing higher competitive pressures are more likely to disclose such information in their annual reports.

IA.2.2 Competition and cross-industry variation

Hoberg and Phillips (2016) show that text-based industry classification captures more across-industry variation and is more informative than traditional industry classifications such as the SIC codes. We test whether this proposition holds for our competition measure (*LnCosMean*).

We compare the informativeness of industry classifications based on the extent to which they generate higher levels of cross-industry variation in profitability (*ROA*, *ROE*), sales growth (*SalesGrowth*), and investment activities (*AssetGrowth*, *FixedInvestGrowth*). Because we hold fixed degree of granularity in the classifications we compare, we conclude that a classification generating a higher degree of across-industry variation is more informative. Following Hoberg and Phillips (2016), we first compute a given firm's industry value of a given characteristic as the mean of the given characteristic among its industry peers for each year and then compute across-industry variation as the standard deviation of these industry characteristics across all firm-year observations in our sample.

We report the results in Table IA2.2. As shown, the standard deviation of profitability (*ROA*, *ROE*), sales growth (*SalesGrowth*), and investment activities (*AssetGrowth*, *FixedInvestGrowth*) based on text-based industry classifications are higher than those in Wind-3 industry classifications, which is consistent with the findings of Hoberg and Phillips (2016) that text-based industry classifications are more informative than traditional industry classifications.

IA.2.3 Competition and firm performance

Porter (2008) shows that a firm can maintain a competitive advantage in the product market after successfully controlling costs, and obtaining a long-term, stable excess profit. Lieberman and Montgomery (1988) also note that firms with competitive advantages in the product market generally have better performance. We test whether this relationship holds for our competition measure (i.e., *LnCosMean*).

We examine two measures of firm performance—*Sales* (defined as the natural logarithm of firm revenue) and *MarketShare* (defined as firm revenue as a percentage of industry revenue based on Wind-3 Industry Classifications). Based on the findings of the prior studies, if the product market is highly competitive, firms are expected to have lower sales and market share.

We report the results in Table IA2.3. Consistent with our conjecture, the coefficient estimate on *LnCosMean* is negative and significant in both columns, which indicates that firms facing higher competitive pressure on average have worse firm performance.

IA.2.4 Firm innovation and competition

Hoberg and Phillips (2016) show that firms' innovation activities can bring them competitive advantages in the product market. Lieberman and Montgomery (1988) also find that firms can

obtain a competitive advantage in the product market through technology innovation. We test whether this relationship holds for our competition measure (*LnCosMean*).

We construct two measures of innovation activities—*Patent_Apply* (defined as the number of patent applications (divided by 1,000)) and *Patent_Grant* (defined as the number of patents granted (divided by 1,000)).

Table IA2.4 reports the regression results. We multiply the dependent variable by 100 for ease of presentation. The coefficient estimate on *Patent_Apply* and *Patent_Grant* is negative and significant, which is consistent with the findings of prior studies that firms with more innovation obtain product market competitive advantage and face less competition.

IA.3 Examples of Firms and Product Market Competitors

In Table IA2.5, we provide selected example of focal firms and their competitors in our sample that are identified based on our product market similarity method.

References in Internet Appendix

- Hoberg, G., and Phillips G., 2016, Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124, 1423–1465.
- Lieberman, M.B., and Montgomery, D.B., 1988, First-mover advantages, *Strategic Management Journal* 9(S1), 41-58.
- Porter, M.E., 2008, *Competitive strategy: Techniques for analyzing industries and competitors*. Simon and Schuster.

Table IA.1 Competition and Disclosure

Dependent Variable	<i>CompFrequency</i>
	(1)
<i>LnCosMean</i>	0.133*** (0.041)
<i>Age</i>	-0.003** (0.001)
<i>Size</i>	-0.005*** (0.001)
<i>Leverage</i>	-0.014*** (0.004)
<i>MB</i>	0.004*** (0.001)
<i>SOE</i>	0.002 (0.003)
<i>BoardSize</i>	0.001*** (0.000)
Firm and Year FE	Included
Observations	23,317
R ²	0.35

Note: ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table IA2.2 Competition and Across-Industry Variation

	<i>ROA</i>	<i>ROE</i>	<i>SalesGrowth</i>	<i>AssetGrowth</i>	<i>FixedInvestGrowth</i>
Text-Based Industry Classification	0.021	0.038	0.181	0.133	0.793
Wind-3 Industry Classification	0.019	0.033	0.150	0.116	0.529

Table IA2.3 Competition and Firm Performance

Dependent Variable	<i>Sales</i>	<i>MarketShare</i>
	(1)	(2)
<i>LnCosMean</i>	-0.526** (0.250)	-0.025* (0.013)
<i>Age</i>	0.068*** (0.010)	-0.000 (0.001)
<i>Size</i>	0.817*** (0.007)	0.013*** (0.000)
<i>Leverage</i>	0.187*** (0.028)	0.005*** (0.002)
<i>MB</i>	-0.110*** (0.007)	-0.001*** (0.000)
<i>SOE</i>	0.042** (0.017)	-0.000 (0.001)
<i>BoardSize</i>	0.007** (0.003)	0.000 (0.000)
Firm and Year FE	Included	Included
Observations	23,292	23,293
R ²	0.92	0.86

Note: ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table IA2.4 Firm Innovation and Competition

Dependent Variable	<i>LnCosMean</i>	
	(1)	(2)
<i>Patent_Apply</i>	-0.501*** (0.168)	
<i>Patent_Grant</i>		-0.584*** (0.226)
<i>Age</i>	0.006 (0.032)	0.007 (0.032)
<i>Size</i>	0.168*** (0.030)	0.166*** (0.030)
<i>Leverage</i>	-0.183* (0.102)	-0.185* (0.102)
<i>MB</i>	-0.014 (0.024)	-0.015 (0.024)
<i>SOE</i>	-0.044 (0.076)	-0.044 (0.076)
<i>BoardSize</i>	-0.009 (0.010)	-0.009 (0.010)
Firm and Year FE	Included	Included
Observations	18,844	18,844
R ²	0.67	0.67

Note: ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table IA2.5 Examples of Firms and Their Product Market Competitors

Incumbent Firm	Year	Rival Firm	Product Similarity Value
Kangmei Pharmaceutical Co., Ltd (Manufacture and distribution of medications, medication ingredients, health foods, and medical equipment, primarily focusing on traditional Chinese medicine)	2006	Shuanglu Pharmaceutical Co., Ltd (Development and marketing of genetic engineering drugs)	0.372
	2010	Qizheng Tibetan Medicine Co., Ltd (Research, development, and production of Tibetan medicine and herbal decoction pieces)	0.308
Tongfang Co., Ltd (Operations in various sectors including consumer electronics, information technology, and energy and environmental industries)	2010	Jiuqi Software Co., Ltd (Operations in the computer software industry)	0.358
	2016	China National Nuclear Co., Ltd (Primarily focus on nuclear power projects, also ventures into the IT sector)	0.316