

Discussion of A Measure of Competition Based on 10-K Filings

GORDON M. PHILLIPS*

1. Introduction

Li, Lundholm, Minnis (henceforth LLM) examine one of the most central and important areas of economics: competition. Competition is what is viewed as promoting efficiency and what keeps the price of a product closer to its marginal cost. The topic dates back to Adam Smith [1776] in the *Wealth of Nations*. George Stigler won the Nobel prize in 1982 for his “seminal studies of industrial structures, functioning of markets and causes and effects of public regulation.”

Competition has its own Wikipedia entry. According to Wikipedia “Competition in biology, ecology, and sociology, is a contest between organisms, animals, individuals, groups, etc., for territory, a niche, or a location of resources, for resources and goods.”¹ Competition has an important policy focus as society wishes to ensure that the contest or game is “fair.” One of the chief parts of competition enforcement in business is to ensure companies do not take actions that restrain competition and result in harm to consumers (or firms) through prices that are systematically higher (or lower) than the cost of production.

Competition has many attributes and extensive research and policy interest. An examination of competition typically begins by determining the market or location over which competitors are competing. Second, the identity of current or potential competitors competing in the product space is determined. Third, regulators or analysts would determine the type of competition and whether it is unidimensional: competition through prices (Bertrand competition), quantities (Cournot competition), or whether

*Marshall School of Business, University of Southern California and National Bureau of Economic Research.

I acknowledge very helpful comments from Gerard Hoberg, Hernan Ortiz-Molina, Oguzhan Ozbas, and Giorgo Sertsios.

¹ See <http://en.wikipedia.org/wiki/Competition>.

competition might also involve multiple dimensions such as product variety or product differentiation (Chamberlain [1933]) or quality (Spence [1975]). After these traditional steps, the extent of competition is measured and the actual or potential effects of competition are examined or determined.

From the above process discussion, there are five important aspects of competition examination: (1) market definition, (2) identification of competitors, (3) identification of the type of competition, (4) measurement of competition, and (5) identifying or quantifying effects of competition.

The analysis of competition and these aspects of competition are sufficiently important that there have been groundbreaking new methods and tools developed and used in its analysis. Game theory was developed to analyze how firms and parties interact and affect the environment in which they compete. Given that parties compete for multiple periods, dynamic game theory was developed and used.

On the empirical side, early industrial organization took structure as exogenous and analyzed conduct and subsequent performance.² The analysis was interindustry and did not take into account that industry structure is itself endogenous. Given industry structure in many cases is clearly not exogenous as participants make decisions to influence other participants and structure itself, new empirical industrial organization that emphasizes structural models and intraindustry analysis began to be developed and used (see Bresnahan [1989] and Phillips [1995]), and is the primary way competition is analyzed today in industrial organization (see Einav and Nevo [2007]).

LLM focus on one of these aspects: the measurement of competition. Given the extensive history and increasing complexity of competition analysis, LLM is striking in its simplicity. LLM count the number of times the firm uses the word “competition” and close variants of competition in its 10-K statement filed with the Securities and Exchange Commission. They follow previous work in this area by Hoberg and Phillips [2010, 2011] and Hoberg, Phillips, and Prabhala [2013] (HPP) that also uses linguistic information from 10-Ks to measure competition. What is unique to LLM is that they bypass the other parts of the examination of competition and measure competition without identifying the market or the competitors in that market. Competition is inferred by what the firms say.

Subsequent to the measurement of competition, LLM study the effects of competition to validate their measure by examining the extent that firm accounting returns decrease more when firms mention competition more. They use the measure in an interindustry sense in the tradition of the structure, conduct, and performance literature in that the measure of competition is on the right-hand side of the estimated equations.

²This literature was termed “Structure, Conduct and Performance” (SCP) literature. See Schmalensee [1989] for a review of the issues and see Bresnahan [1989] for an early review of the new empirical organization.

In what follows, I briefly review the LLM method and findings. I discuss how the LLM measure relates to historically used measures of competition. I follow by examining more conceptual issues of how to measure competition and the potential problem of managerial misrepresentation and discuss several points raised by conference participants. I follow by discussing how the LLM measure of competition was used and can be used in the future with the new measures of competition provided by Hoberg and Phillips [2010, 2011] and HPP [2013].³ I conclude by suggesting areas for future research.

2. *The LLM Measure and Traditional Measures of Competition*

In this section I first discuss traditional measures of competition. I then discuss the new LLM measure of competition and the situations where the LLM measure will be useful to researchers. I then discuss the limitations of the LLM measure. I then compare the LLM measure to more traditional industry-based measures of competition and the new text-based market definition methods of Hoberg and Phillips [2011]. My conclusion is that the LLM measure is useful but should be used in conjunction with other industry-based methods of defining market segments and identifying competitors. This combination both addresses some of the limitations and provides new ways of examining interesting research questions.

2.1 TRADITIONAL MEASURES OF COMPETITION

How does the new LLM measure compare to traditional measures of competition? There are three traditional measures of competition that have been used by economists. These are the Herfindahl–Hirschman index (HHI), which is the sum of squared market shares for all firms in an industry product segment, price–cost margins (PCMs) or Lerner indices, and cross-price elasticities of demand. A cross-price elasticity of demand measures the change in own-firm demand with respect to the price change of a rival product. Formally, it is defined as $E_{A,B} = \frac{\partial Q_A}{\partial P_B} \times \frac{P_B}{Q_A}$.

The first thing to note when comparing the LLM measure to traditional measures is that traditional measures of competition are based on a definition of an industry product market and corresponding rival firms within that product market. This is true for the HHI as it is the sum of squared market shares. Other traditional measures such as the cross-price elasticity of demand, which is the proportional change in demand for a given change in a rival firm's price, are firm based but require identification of competitors. This is also true for PCMs as these are computed for firms within or relative to industry benchmarks and are frequently averaged over firms in an industry group to compute an industry price cost margin.

³These new measures of competition along with industry classifications and corresponding firm memberships are available to researchers via the Internet at www.rhsmith.umd.edu/industrydata.

An example that estimates industry supply and demand equations to estimate the type of industry competition is Phillips [1995]. Phillips [1995] follows in the tradition of new-empirical industrial organization of Bresnahan [1989]. He estimates industry supply and demand functions and estimates the degree of competition in four different industries following debt recapitalizations. Equations are estimated at the industry level and identified through the use of a substitute industry price. The conclusion is that industry competition is reduced in three of the industries following debt recapitalizations.

Nevo [2001] determines the type of industry competition and the source of high PCMs in the cereal industry by modeling underlying demand to estimate PCMs without observing the underlying actual costs. Competitor market share data, prices, brand characteristics, and advertising are all needed. The conclusion of the paper is that product differentiation is responsible for the high PCMs.

The traditional measures, however, have their limitations as well. The costs of these methods are related to their precision—a researcher needs well-defined markets, products, and identification of competitors along with time-series of prices and quantities. Thus by their nature, traditional measures are only applicable in well-defined markets where such data exist.

An additional problem with these existing measures is that they do not capture potential competition, which may cause firms to engage in limit pricing when faced with a potential entry threat. The traditional measures also require a definition of the relevant product market. Many industries such as business services are not well defined. For most research purposes the definition of a market is based on Standard Industrial Classification or NAICS codes. Hoberg and Phillips [2011] have shown that these codes are subject to staleness and misclassification and are improved upon by looking for similarities in product descriptions.

Note that the new LLM measure does not have to be used at the exclusion of more traditional measures. They may be complementary in some situations. The ideal use of the new LLM measure will be in studying who wins and loses inside product markets that are determined using either more traditional market segment definitions or within the new text-based market definitions of Hoberg and Phillips [2011]. In the latter half of their paper the authors conduct tests where they combine their new measure with more traditional measures of competition and new market segment definitions provided by Hoberg and Phillips.

2.2 THE LLM MEASURE OF COMPETITION

The LLM measure itself is simple and direct as the authors note. It is firm specific. The LLM measure counts the number of times the firm uses the word competition or variants thereof, (competitor, competitive, compete, competing), including those words with an “s” appended. They remove any case where “not,” “less,” “few,” or “limited” precedes the word by three or fewer words. They then scale this measure to create the variable *PCTCOMP*,

which is the number of competition words per 1,000 words in the 10-K. The mean/median/75th percentile of this variable is 0.58/0.48/0.78. In their analysis they use the variable *COMP*, which is the decile-ranked value of the percentage of times the word competition is used.

The authors conduct several empirical tests to try to ascertain whether their measure captures competition and also to address potential issues that conference participants brought up. I will first discuss the potential issues that were raised and then discuss how the authors addressed these issues.

The main issue, as with any new measure, is whether it measures competition itself. The authors thus begin by validating that their measure is correlated with historical measures of competition and also the new text-based similarity measure that was used in Hoberg and Phillips [2010]. They find that mentions of competition are highly correlated with measures of industry competition. In particular, table 2 shows that the mentions of competition increase with Hoberg and Phillips product similarity. Looking at the *t*-statistics for differences in means across *PCTCOMP* quintiles, similarity is statistically much higher for the highest quintile of mentions of competition versus the lowest quintile.

Next the authors conduct their main tests. In Hypothesis 1, the authors examine the extent to which changes in cash flows are correlated with mentions of competition. They estimate specifications like the following:

$$\begin{aligned} \Delta RNOA_{i,t+1} = & \beta_0 \times Year_t + \beta_1 \times RNOA_{i,t} + \beta_2 \times \Delta NOA_{i,t} \\ & + \beta_3 \times COMP_{i,t} + \beta_4 \times COMP_{i,t} \times RNOA_{i,t} \\ & + \beta_5 \times COMP_{i,t} \Delta NOA_{i,t}. \end{aligned} \quad (1)$$

In this equation the dependent variable is the change in the operating return on net assets. The authors use their measure of competition, *COMP*, as an exogenous measure on the right-hand side of this equation. In addition to *COMP*, the new proposed measure of competition, the independent variables include the level of *RNOA* to capture mean reversion, the past change in Net Operating Assets (ΔNOA) to capture decreasing ROI interactions between these variables and *COMP* to capture the fact that competition will cause firms investments to have a diminished payoff.

One initial large concern is that *COMP* is treated as exogenous. Clearly *COMP* may be influenced by the competitors' contemporaneous actions and the firm's own actions—even if it is lagged. The reason that just lagging the variable may not suffice is that longer term decisions such as new product introductions and investment in new plant and equipment have long-lasting implications and may affect competition itself. For example, DuPont may build a large new plant to preempt entry by competitors for some time to come. But it may still talk about high competition by potential entrants.

The potential endogeneity of this measure is related to the main concern about this measure that was brought up by several conference participants. That concern is the potential for firms themselves to misrepresent the

extent of competition they face. The potential for such misrepresentation is clear. The authors do recognize the potential for such misrepresentation and thus, as I discuss later, attempt to validate their measure by comparing it to existing measures of competition. Of course, this measure or existing measures of competition will not be used directly by regulators as they conduct microlevel studies of competitive effects. The LLM measure can still be useful in measuring some of the competitive pressures that firms face that may influence financial statement analysis.

I discuss two types of potential misrepresentation here. The first potential type of misrepresentation is that firms that are potentially restraining competition may discuss at length competition in small product areas and inflate the extent of competition they face in order to avoid an inquiry by antitrust authorities. They may face low competition but may also talk about extensive competition in order to get a proposed acquisition approved.

A second type of potential misrepresentation is that firm managers may blame competition for production problems or products that fail to attract consumers. If earnings are low because consumers choose not to buy the firm's products, the managers may try to attribute their poor results to the effect of competition. Firms may produce a new product for which there is low demand and the firm may blame the lack of sales on competition. This attribution would not be due to negligence, but rather to the intent by managers to save their jobs. An early example is Ford's production of the Edsel car. The product's demise may really be managerial miscalculation or mistakes, yet the manager may mention competition as an explanation.

LLM thus test if their measure is robust to industry measures of competition. In table 8, they include multiple measures of industry competition and also interact industry measures of competition with the firm-specific mentions of competition. The results for the firm-specific reporting of competition remain robust. The table also reports negative coefficients on the industry measures of competition interacted with the level of the ROAs and also change in the operating assets, consistent with industry competition having additional negative effect on margins. The most likely interpretation of these results is that managers of firms that expect to fare the worst as a result of competition mention competition the most. These firms are the ones that have higher levels of margins and a higher lagged increase in their operating assets.

2.3 LIMITATIONS OF THE LLM MEASURE IN COMPETITION ANALYSIS

There are three primary limitations of the LLM measure in addition to the potential for misrepresentation discussed earlier.

First, by construction it focuses just on measurement. If we return to the definition of competition on Wikipedia, "competition ... is a contest between organisms, animals, individuals, groups, etc., for territory, a niche, or a location of resources, for resources and goods." LLM document how participants report or describe the "contest" or interaction at the end of the year. This measure does not identify the market in which the firms are

competing, the type of contest, or what the firms are competing over (prices, quantities, type of good).

Thus, the first limitation, and perhaps the most important, is that it does not identify the product market in which the firms are competing. This limitation is particularly important for competition authorities. In the United States, the Department of Justice has extensive guidelines for horizontal mergers (see: <http://www.justice.gov/atr/public/guidelines/>) and a major component is the definition of the market. The Department of Justice guidelines say “market definition allows the agencies to identify market participants and measure market shares and market concentration.” “The measurement of market shares and market concentration is not an end in itself, but is useful to the extent it illuminates the merger’s likely competitive effects.”

Second, related to the question of market definition, the LLM measure does not identify the competitor firms in the product market that the firm may be referring to in its 10-K. Identifying competitors is clearly important to measure if a product faces competition. In economics, in a large set of dynamic models including recent articles at the intersection of industrial organization and finance and accounting (Chod and Lyandres [2011], Spiegel and Tookes [2013]), identifying competitors is key.

Third, we do not know if firms are competing through quantity, competing by first building new plants, or competing through offering product variety. Identifying the type of competition, as well the relevant product market and the competitors, is also important in measurement as one key measure of extent of competition that has been used is the cross-price elasticity of demand. Nevo [2000] estimates own- and cross-price elasticities of demand and their effect on postmerger prices in the ready-to-eat cereal market to determine the extent products compete with each other.

While the new LLM measure has no identification of a market segment and no identification of competitors, the lack of these features has some positives. It is useful in situations where researchers are studying firm-specific managerial decisions that are related to managerial perceptions, and not actual competition. Some decisions such as advertising or capital structure may be taken based on managers’ perceptions of competition. It is also useful to researchers who are interested in how firm-specific earnings evolve over time based on firm-specific measures of competitive intensity in situations where it is reasonable to take competition as fixed in the short run. It thus is more useful for short-run decisions or analysis of shorter-run changes.

The reason that this measure may be more useful for analysis of shorter run changes is that, as discussed earlier, the measure of competition developed here is firm-specific and may suffer from endogeneity concerns even if lagged (see Bresnahan [1989] for a discussion). It is well recognized that competition can be particularly affected by the firms themselves through their longer-term investments in plant and equipment. Thus competition is particularly endogenous with respect to longer-term investments. Given

short-term decisions will have less direct effect on this measure, this endogeneity concern would be less for these decisions.

3. *New Hoberg and Phillips Text-Based Measures of Competition*

Recently, Hoberg and Phillips [2010, 2011] and Hoberg, Phillips, and Prabhala [2013] have also analyzed firm 10-Ks using computational linguistics and provide new measures of competition on their Web site as well.

I will spend a brief part of my comments discussing these new measures and how they can be used and are used by LLM. LLM combine their measure with these new industry definitions to further enhance the tests in their paper. They compute a measure of *COMP* that counts how many times competitors identified using the Hoberg and Phillips TNIC classification mention competition together. They find that there are larger decreases in margins when competitors simultaneously mention competition.

Hoberg and Phillips [2010, 2011] use all the words that the firm has in the product description of its 10-K and determine the relation of these words with all other publicly traded firms to produce a relatedness or similarity score for each firm with all other publicly traded firms (specifically, they produce a cosine similarity score). Their analysis thus produces an $N \times N$ matrix of similarity scores. Given that there are approximately 5,000 firms in each year, their analysis produces approximately a $5,000 \times 5,000$ matrix in each year. They then determine industries or competitors for each firm by grouping firms into industries or competitors based on minimum similarity scores.

They provide two different classification systems—both of which are available on a publicly accessible Web site. Both of these classification systems identify market segments and also identify competitors of each firm. They both allow competition to evolve over time with annual changes. The first is historically motivated while the second allows industry competition to be firm-centric. The first, “fixed industry classifications” (FIC), is analogous to SIC and NAICS industries. In this classification system, firms are grouped into a set of industries that are fixed over time and membership in an industry is required to be transitive. For example, if firms B and C are in firm A’s industry, then firms B and C are also in the same industry. Furthermore, although Hoberg and Phillips FIC industries adhere to the same transitivity restrictions as SIC and NAICS industries, they differ because they use clustering algorithms that maximize total within-industry similarity based on product market word usage in 10-K business descriptions.

The second classification system is more general. In this classification, Hoberg and Phillips allow firm competitors to change every year and relax the membership transitivity requirements of FIC industries and view industries like flexible networks. They name these new generalized network industries “text-based network industry classifications” (TNIC). In this classification system, each firm can have its own set of distinct competitors analogous to a social network. In this system each individual firm has a distinct

set of competitors (friends in a social network), with competitors of one firm not necessarily being competitors of each other. To illustrate why transitivity is restrictive, suppose firms A and B both view firm C as a rival. If A and B each have products with different distinct features or enhancements that C does not have, then A and B may not compete against each other as they may serve different product segments.

Using these sets of competitors, they then produce several different industry measures of competition that begin with the determination of the product market space, including (1) Herfindahl indices based on the new competitors (available on their Web site), (2) a product-fluidity measure that captures product market threats used in HPP [2013], and (3) average similarity of the 10 closest competitors, which is used in Hoberg and Phillips [2010]. Lerner industry PCMs can be calculated if price and quantity data are available from sources such as supermarket scanner databases. Identification of peer firms is key to calculate this measure for an industry segment.

Relative to existing industry classifications, these new text-based classifications offer economically large improvements in their ability to explain managerial discussion of high competition as compared to SIC or NAICS codes, the specific firms mentioned by managers as being competitors, and how advertising and R&D create future product differentiation. These new industry measures also offer econometric gains in explaining the cross-section of firm characteristics. Empirical tests further benefit from information about the degree to which specific firms are similar to their competitors, which cannot be derived from zero-one membership classifications such as SIC or NAICS.

In addition to the new industry product-segment measures, HPP [2013] also provide a new measure of overall product market risk that they term “*Product Fluidity*.” This measure used the relation between the 10-K product descriptions of the firm and its rivals. It measures how rival firms are changing the product words that overlap with firm’s vocabulary. It is also available on their Web site.

Product fluidity thus measures the change in a firm’s product space due to moves made by competitors in a firm’s product markets. It captures the idea that entry by rivals can pose competitive threats to a firm. Further supporting a link to competitive threats, HPP find that fluidity is positively correlated with the business descriptions of entrepreneurial firms receiving venture capital or undertaking IPOs—capturing both actual and potential entrants. It would be interesting to see if the LLM mentions of competition are correlated with this new measure of competitive threats provided by HPP.

4. *The LLM Measure Combined with the Hoberg Phillips Industries*

One of the most interesting parts of the paper and where the paper’s measure is most successful is when both industry measures of competition

and firm measures of competition are included together. The authors use both traditional SIC codes and the Fama–French 48 industries where SIC codes are grouped together as well as competitors based on Hoberg–Phillips text-based network classification identification of competitors. They average how many times a firm’s competitors mention competition using these different industry product markets. They label this new measure *iCOMP*.

Table 9 reports that SIC code–based definitions of competitors fare less well than the Hoberg–Phillips text–based network of competitors. The table shows that when *iCOMP* is calculated based on firms with similar product descriptions in their 10-K (as given in Hoberg and Phillips [2011], the coefficients on $iCOMP \times D_NOA$ and $iCOMP \times RNOA$ become significantly larger, and the coefficients on $COMP \times D_NOA$ and $COMP \times RNOA$ become insignificant. They thus report that their new industry-based average measure works best when calculated with the Hoberg–Phillips industry definitions, a finding consistent with the results in Hoberg and Phillips [2011] that the TNIC industry product segments improve the identification of competitors.

Table 10 of the paper continues to explore industry groupings where the managerial mentions of competition may be stronger. It reports that *COMP* and the *COMP* interaction term are strongest in industries with high existing rivalry and high product similarity within the Hoberg–Phillips TNIC industry groupings. Thus the LLM new measure of industry competition identifies which firms’ earnings within an industry grouping are most sensitive to competitive pressure.

5. Conclusions

LLM provide an interesting new measure of competition. It has the advantage that it is simple. It should be useful in situations where researchers are studying whether firm-specific short-run managerial decisions are related to managerial perceptions of competition. Decisions such as advertising and short-run pricing and quantity decisions may be taken based on managers’ perceptions of competition. It is also useful to researchers who are interested in how firm-specific earnings evolve over the shorter run based on firm-specific measures of competition. It is less useful by itself for longer-term decisions such as investment given that the measure itself can be influenced by the firm’s decisions themselves or new product introductions that require identification of the relevant market and competitors.

When combined with other new text-based measures of competition provided by Hoberg and Phillips, LLM show that their new measure of competition is particularly strong. Their results show that, when the product space has higher measures of similarity and firms talk about competition, the decrease in margins is particularly economically and statistically significant. Combining the LLM measure with other industry-based measures such as the Hoberg and Phillips industry groupings that identify competitors

combines the advantages of competitor and market identification with the LLM measure of who is most affected within an industry group by competition.

Additional applications and research questions moving forward could consider which specific firm loses or gains when faced with shocks that increase competition. One promising avenue would be to identify potential external shocks and see if, in these cases, the firms mention high competition more and then see which firms lose or gain as a result of the increased competition. Natural experiments or external shocks can include industry downturns from demand shocks or supply shocks. One application would be to examine import penetration, a form of a supply shock. If imports spike upwards in some industries, it should result in the weaker firms in these industries mentioning competition more often.

REFERENCES

- BRESNAHAN, T. "Empirical Methods for Industries with Market Power," in *Handbook of Industrial Organization*, Vol. II, edited by R. Schmalensee and R. Willig. Amsterdam: North-Holland, 1989: 1011–57.
- CHAMBERLIN, E. *The Theory of Monopolistic Competition*. Cambridge, MA: Harvard University Press, 1933.
- CHOD, J., AND E. LYANDRES. "Strategic IPOs and Product Market Competition." *Journal of Financial Economics* 100 (2011): 45–67.
- EINAV, L., AND A. NEVO. "Empirical Models of Imperfect Competition: A Discussion," in *Advances in Economics and Econometrics: Theory and Applications*, edited by R. Blundell, W. Newey, and T. Persson, Ninth World Congress Series, 2007: 86–96.
- HOBERG, G., AND G. M. PHILLIPS. "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis." *Review of Financial Studies* 23 (2010): 3773–811.
- HOBERG, G., AND G. M. PHILLIPS. "Text-Based Network Industries and Endogenous Product Differentiation." Working paper, University of Maryland, 2011.
- HOBERG, G.; G. M. PHILLIPS; AND N. R. PRABHALA. "Product Market Threats, Payouts, and Financial Flexibility." *Journal of Finance* (2013): Forthcoming.
- LI, F.; R. LUNDHOLM; AND M. MINNIS. "A Measure of Competition Based on 10-K Filings." *Journal of Accounting Research* 51 (2013): 399–436.
- NEVO, A. "Mergers with Differentiated Products: The Case of the Ready to Eat Cereal Industry." *Rand Journal of Economics* 31 (2000): 395–421.
- NEVO, A. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69 (2001): 307–42.
- PHILLIPS, G. M. "Increased Debt and Industry Product Markets: An Empirical Analysis." *Journal of Financial Economics* 37 (1995): 189–238.
- SCHMALENSEE, R. "Inter-Industry Studies of Structure and Performance," in *Handbook of Industrial Organization*, Vol. II, edited by R. Schmalensee and R. Willig. Amsterdam: North-Holland, 1989: 951–1009.
- SMITH, A. *The Wealth of Nations*. London: W. Strahan and T. Cadell Publishers, 1776.
- SPENCE, A. M. "Monopoly, Quality and Regulation." *Bell Journal of Economics* 6 (1975): 407–14.
- SPIEGEL, M., AND H. TOOKES. "Dynamic Competition, Valuation and Merger Activity." *Journal of Finance* 68 (2013): 125–72.