Inferring Firm Performance through Analytics in Latent Strategy Spaces

Ashish S. Galande, Prakash Satyavageeswaran, Sudhir Voleti

"What explains firm performance?" is a pivotal question in management research. Firm performance (FP) is some function of characteristic factors internal to the firm (e.g., firm demographics, capital structure etc.), of environmental factors external to the firm (e.g., industry structure, regulatory controls, etc.), and of firm strategy, which connects and reconciles the two. Despite various available metrics, measuring and thereafter modeling this complex, contextual and evolving firm strategy object has remained a challenge. Previous approaches have used metric variables to proxy for particular aspects of firm strategy or have tested theoretical conceptualization using primary data collected from managers. We take an alternative, general and scalable approach. We mine for and directly transform strategic content in relevant textual sources into usable variables. The two elements of strategy focal to our study are readily available in regulatory filings in the form of descriptive text. However, three obstacles stand in the way of practical implementation of this approach - (a) extracting objective measures of firm strategy from text descriptions alone; (b) modeling cross-firm *externalities* in the firm's strategic elements; and (c) estimating, interpreting and statistically inferring the marginal effects of firm strategy on FP.

Our solution approach is to innovatively combine three existing methods, viz. bag-of-words text mining and analysis, latent space modeling, and conventional spatial econometrics, to overcome the three obstacles mentioned. We use text analysis procedures to *atomize* a firm's strategic vocabulary into a vector of frequency-weighted phrase-tokens in some latent space of firm strategy. Next, we use relative firm locations in this latent space as an input to a class of spatial autoregressive models from the spatial econometrics literature, and thereby enable estimation, inference and interpretation of strategy externalities on FP. We label this approach as *text-spatial* modeling.

We demonstrate our proposed approach empirically on firms featuring in the 2013 rankings of the Fortune 1000 firms. Both our FP measures, Tobin’s Q and Return on Assets (RoA), are derived from firm financial information for financial year 2012 taken from S&P Compustat. To operationalize the strategy variables, we extract the text of firms' 2012 Form 10-K filings with the SEC. We label as 'BD' (for

'business description') the text content of 'Item 1 - Business' in the 10-K. BD describes the business of the company and reflects the strategic choice of product-markets and product-portfolios: company's business, its' subsidiaries, its' markets. Similarly, 'RF' (for 'Risk factors') forms the text content of 'Item 1A - Risk Factors', wherein firms describe potential risks, likely external effects, contingencies, and other threats disclosed to inform (potential) investors.

We find evidence that there exist externalities in firm strategy in both product portfolio choice and risk factor exposure. We also find that much of the strategy externality effect acts through other firms' outcomes rather than their factor inputs, and (expectedly) that in firms which are neighbors in both product and risk space, the externality effects of firm strategy are the strongest.

As illustration, consider Intel Corp. which belongs to the *Technology* sector. Figures 1a and 1b show Intel's 30 closest neighbors in product space and risk space respectively. The vertices in "Blue" are firms sharing the same sector as Intel and the ones in "Red" are from other sectors, suggesting that strategy externalities exert influence across sector and industry.

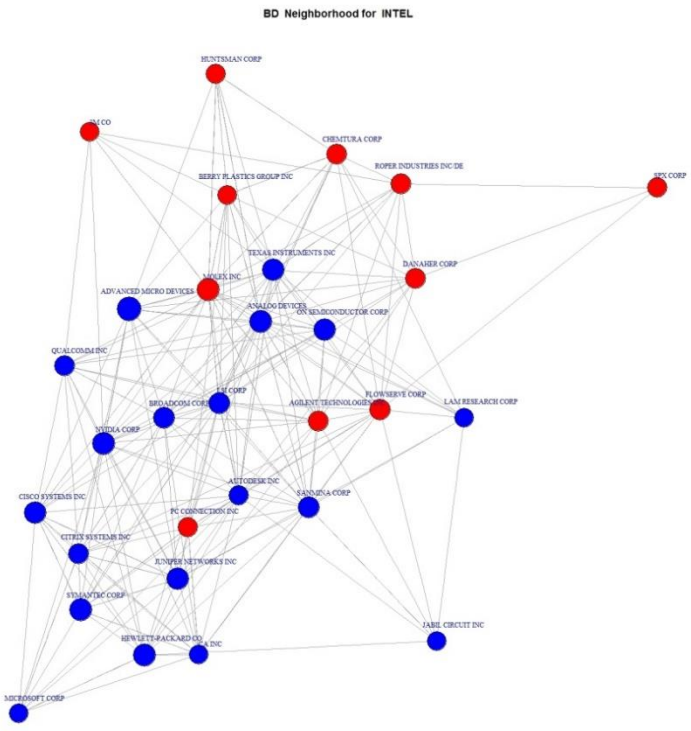

Figure 1a: BD based neighborhood for Intel

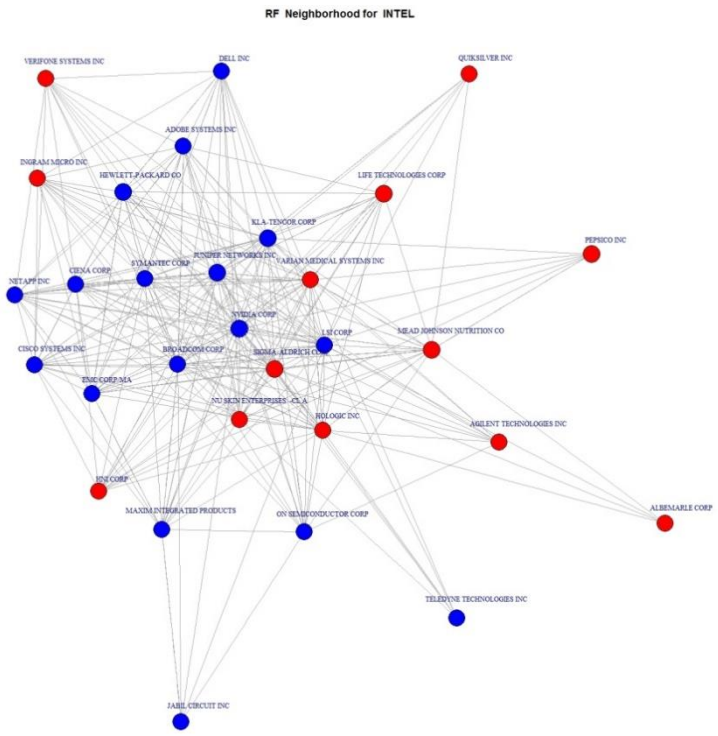

Figure 1b: RF based neighborhood for Intel

We explore three ways in which firm strategy could influence FP: through firm characteristics, through firm's own outcomes, or through an interaction of the two strategy spaces. We refer to these three possibilities as "modeled effects", "SAR" for simultaneous auto regressive and "SDM" or spatial Durbin Model, respectively. These claims are evaluated against the benchmark "IID" or identically and independently distributed model which assumes a simple linear regression with zero strategy externality effects. Twelve different model specifications (four predictor subsets times three predictor types) emerge. We use the adjusted R squared (adj R^2) measure of fit following the use of a generalized regression framework for model estimation. Panel A in Table 1A displays the adj R^2 values of the baseline IID model and the twelve model specifications we estimate for Tobin's Q. Panel B shows the % increase in penalized model fit for each model over the baseline IID model. Table 1b displays similar comparisons for RoA as performance measure. Together, these results provide answers to three important questions: First, "Do cross-firm externalities in product portfolio and risk factors explain FP outcomes?" Second,

"How much additional FP is explained due to cross-firm externality effects?" And finally, "How do these externalities act? Through neighbors' inputs? Neighbors' outcomes? Both?"

Table 1a: Model Fit across Different Specifications for Tobin's Q

Panel A	**Adjusted R squared for Model Fit**				
Model	**Baseline model**	BD Alone	RF Alone	BD & RF Simultaneous	BD, RF & BDRF Simultaneous & Interacted
IID	**0.453**				
Modeled		0.47	0.475	0.479	0.48
SAR		0.46	0.457	0.46	0.49
SDM		0.478	0.511	0.535	**0.639**

Panel B	**% Change over IID in Model Fit**				
Model	**Baseline model**	BD Alone	RF Alone	BD & RF Simultaneous	BD, RF & BDRF Simultaneous & Interacted
IID	0%				
Modeled		3.75%	4.86%	5.74%	5.96%
SAR		1.55%	0.88%	1.55%	8.17%
SDM		5.52%	12.80%	18.10%	**41.06%**

Table 1b: Model Fit across Different Specifications for RoA

Panel A	**Adjusted R squared for Model Fit**				
Model	**Baseline model**	BD Alone	RF Alone	BD & RF Simultaneous	BD, RF & BDRF Simultaneous & Interacted
IID	**0.433**				
Modeled		0.435	0.45	0.456	0.453
SAR		0.438	0.435	0.446	0.447
SDM		0.485	0.478	0.518	**0.563**

Panel B	**% Change over IID in Model Fit**				
Model	**Baseline model**	BD Alone	RF Alone	BD & RF Simultaneous	BD, RF & BDRF Simultaneous & Interacted
IID	0%				
Modeled		0.46%	3.93%	5.31%	4.62%
SAR		1.15%	0.46%	3.00%	3.23%
SDM		12.01%	10.39%	19.63%	**30.02%**

While the information in Table 1 suggests that externalities in strategy inform FP, it is of limited utility in explaining how the effect acts and what might drive it. Cross-firm strategy externalities in product and risk spaces could have a positive, a negative or a non-significant effect on any focal firm's performance. The working papers carry more detail on the signs, magnitudes and relative importance of different firm characteristics in interacting with strategy variables.

In sum, we address the problem of (i) extracting strategic content from textual descriptions of firm actions, perceptions and expectations, (ii) transforming this content into econometric measures of firm strategy, (iii) developing a general model of FP as a function of firm strategy, (iv) empirically implementing the proposed model on a large and diverse sample of firms. Through this, we provide a novel manner of better understanding and inferring the effects of strategy on FP.