Assessing the Impact of Executive Diversity/Centrality on Firm Performance

Authors: Ailin Sun, Luke Liu, Eric Luong, Sindhu Raghavendra Moktali

Project Guidance: Dan Trepanier **Project Sponsorship:** VANGUARD

Abstract

In the present era of globalization and interconnectivity, diversity has become a prominent focal point, especially within the corporate realm. Some studies explored the link between diversity and a company's performance, utilizing organizational-level statistics; others turned their focus on leadership-level statistics (see References). In our study, we explored the correlation between leadership level diversity and its impact on the stock market performance of corporations.

The main objective of this study is to uncover what diversity and centrality metrics within a firm's leadership have significant correlations to firm performance, and if it's possible to outperform a baseline market return with an understanding of those metrics. In this study, diversity refers to the variances in ethnicity, gender, and age within a firm's leadership, and centrality refers to the relative importance of a leader within a firm, based on their network relationships and influence within the organization.

Guided by the Quant Equity Group (QEG) of Vanguard, which oversees about \$50 billion in active systematic strategies primarily invested in public equities, our team conducted data collection to identify and gather relevant datasets on executive directors, financial performance metrics, and market performance. Using metrics such as the Herfindahl–Hirschman index for diversity, age range standard deviation, and centrality measurements, we generated a dataset to represent a firm's leadership diversity, centrality, and then paired with its performance on the stock market over a period of 15 years.

Our analysis suggests that certain leadership diversity and centrality measures did have statistically significant correlations with the forward returns of firms. However, due to the scope of the project and time constraints, these results merit further studies on the multifaceted phenomena of diversity vs performance.

Project Scope and Objectives

The main objective of our project was to explore the potential correlations between a firm's overall share price returns and its board of directors/c-suite diversity and centrality.

The scope of the project was limited to two main parts. The first, and also the more challenging and time consuming part, is defining, quantifying, and collecting data on centrality and diversity metrics for each firm's leadership circle. The second part was in analyzing the potential correlations between these diversity and centrality metrics and firm financial performance.

This study therefore requires data from both financial and biographical sources.

Financial data needed to consist of a given firm's performance on the stock market, measured in the change in price, as well as a suite of financial metrics, ratios, and general information.

Biographical data, on the other hand, were much more complex. We needed to derive not only leadership centrality information, but also information in relation to ethnicity, age, and gender. Furthermore we also had to decide how to define and quantify multiple and diverse diversity metrics.

Due to the nature of corporate reporting, the major difficulty in data collection is in accurately determining the demographic makeup of both C-suite executives and the governing members of the board over time. While on the other hand, thanks to the nature of corporate reporting, our team is able to accurately determine the financial performance and make up of any given publicly traded corporation.

This dichotomy provides us with two different avenues of data collection: financial and biographical data.

Finally, we need a dataset large enough to provide statistically significant results. This means the project has to encompass as many firms as possible and the longest period of time within the limitations of our chosen dataset.

Our data was collected from different sources. Biographical data was collected via the ISS: Institutional Shareholder Services Directors database, while financial data was collected via the CRSP and COMPUSTAT databases, and then subsequently joined together.

Using this data, our team was able to create a target variable, expressed through stock returns; and a feature set, consisting of firm financial performance metrics, diversity and centrality measures, as well as other information that is determined to be important to our analysis.

This dataset is inclusive of 1500 firms over a 15 year period of time. See the data dictionary located in the appendix [Table A].

Work plan and timeline

January-February

 Collect and preprocess data into a unified dataset on executive directors, financial performance, and market performance to ensure consistency and accuracy

March-Mav

- Augment the prepped data to allow better fine-tuning of diversity and centrality metrics.
- Define and create metrics for centrality, diversity, and explore other potential metrics
- Train and test models using appropriate modeling techniques, such as regression analysis
- Conduct benchmarking/relative performance analysis by analyzing significance of diversity/centrality metrics relative to established financial ratios and fundamentals

May

- Interpret the model coefficients
- Prepare recommendations and next steps based on the insights and implications of the project
- Identify areas for further research or exploration that can add value to Vanguard's Quant Equity
 Group
- Include an executive summary that provides a concise summary of the project objectives, methods, and key findings

• Ensure that the final deliverables are tailored to their needs and provide insights that can help them make informed decisions about their investment strategies

Key Project Findings

Our project goal was to conduct a preliminary correlation analysis of the impact of a firm's board of directors and c-suite diversity and centrality on stock price returns, using similar methods from prior studies. And in order to determine the impact of leadership diversity on returns, we analyzed the statistical relevance of our chosen diversity metrics in relation to the percentage portfolio returns. As for a point of reference, our team initially looked at the share returns in relation to specific metrics; taking the top and bottom 5% number of firms and comparing their average returns.

Utilizing Python's NetworkX package, we constructed a network graph representing each company's executive leadership. Individual directors served as nodes, identified by a unique 'director_detail_id', with edges established between directors from the same company, forming a comprehensive graph for each firm's executive board. This design reflects real-life board member interconnections and presents the leadership network as a complex system. Four centrality metrics - betweenness, closeness, degree, and eigenvector centrality - were computed for each director, capturing various aspects of their relative importance within the network. See appendix for more information regarding their definitions and how we calculated them [Table A].

We first built out a workflow pipeline for a smaller dataset, more specifically the historical Dow Jones Industrial Average (DJIA). After recreating the index with an R-squared of 0.88, we funneled that data into our centrality and diversity analysis to visualize our feature importance based on a long short strategy. After debugging and testing our strategy on the DJIA, we extrapolated our analysis to a bigger dataset, more specifically the S&P 1500. See the appendix for the alpha plot of forward returns for the bottom vs. top 5% of central companies [Section One].

From the alpha plots, it seems that although the discrepancy between the performance of high and low centrality companies is inconsistent in the short term (in some years one has better returns than the others and vice versa), there seems to be a measurable difference with the most central companies outperforming in the long-term [Table E]. Possible explanations for this could be that centralized boards may be more prone to groupthink and conformity but have better strategic coherency, with them able to adapt quickly to market conditions. While those firms with less central boards are more adaptable, and have less governance issues due to a deconcentration of power, or less conflicts of interests.

In order to measure feature importance and collinearity biases, our team created a correlation matrix relative to the percentile returns for each given feature. The results of which determined our chosen featureset. See the appendix for the correlation matrix of our features [Table B].

Based on our correlation heatmap, there is high correlation between each of the centrality values. This implies that board members who are more central in one aspect of the network tend to be central in other areas as well. This could mean that influential individuals within the network often have multiple dimensions to their influence. However, it could also indicate a level of redundancy or overlapping influence within our calculated network. This may represent a high level of risk and large dependencies within the network; negative events could potentially have cascading effects on multiple dimensions of a

board, possibly impacting stock performance. Understanding these relationships can help identify evolving power structures and power dynamics within a company's leadership.

Following this, our team ran a regression analysis to determine the statistical significance of the given features. This regression analysis took the form of an ordinarily least squares model (now referred to as OLS), in which we analyzed the statistical significance of each of the features. See appendix for the full regression model [Table C].

Our initial regression model consisted of all the original features, but since most of the centrality values were correlated with one another. We decided to take a look at the features that were most prominent in the original regression.

We implemented two separate Ordinary Least Squares (OLS) models: one incorporating control variables (z_pe, z_roce), and the other including both biographical and financial control metrics. The control model yielded an R-squared value of 0.029, while the biographical metrics model showed an R-squared value of 0.035. See the appendix for all regressions run and their subsequent r-squared values [Table F]. Although most of the centrality measures exhibited statistical significance at the 95 to 99% confidence level, there were some that were surprisingly not statistically significant.

Statistically Significant Features

- Price/Equity Ratio (P-value << .05, coefficient = 0.165)
 - Price Equity ratio adjusted for industry.
 - An average positive coefficient represents a correlation of a larger P/R ratio and better returns.
- Return on Capital Employed (P-value << .05, coefficient = 0.585)
 - Return on capital employed adjusted for industry.
 - A strong positive coefficient represents a correlation of larger ROCE and better firm returns
- Betweenness Centrality (P-value << .05, coefficient = -0.323)
 - Betweenness centrality adjusted for industry, its value measures the number of shortest paths between a given node and other nodes in the network. In this case, measures how "centrally located" a given leader is based on their fellow board members. So the lower the value, the more closely connected a leader is to their most distantly connected edge case. This path is calculated using Djikstra's algorithm.
 - A strong negative coefficient represents a correlation between less centrally focused firms and better performance.
- Eigenvector Centrality (P-value = 0.015, coefficient = -0.07)
 - Eigenvector Centrality adjusted for industry, its value measures the axis of greatest commonality/centrality amongst the elements of the feature space. Proportional to the sum of the eigenvector centrality values of its neighbors, weighted by the strength of the connections. In this case, measures how strongly connected a given board is to one another.
 - A weak negative coefficient represents a small correlation between less centralized companies and better firm performance.
- Closeness Centrality (P-value = 0.002, coefficient = 0.116)

- Measure of how easily a node can reach all other nodes on the network. It is calculated
 by taking the sum of the shortest path lengths. Measures how closely connected a leader
 is to all other members of the board/c-suite.
- An average positive correlation represents a correlation between more closely related firm leadership and higher returns.
- Number of Board Cohorts (P-value << .05, coefficient = 0.197)
 - Number of Board Cohorts is a value based on the number of boards each member of the firm sits on other than the current firm, split into bins. Measures the amount of other leadership obligations each given firm's leadership needs to share.
 - An average positive correlation represents a correlation between firms with board groups with more seats on other boards and higher returns.

Non-Statistically Significant Features

- Degree Centrality
- % Gender Diversity
- Age Mean
- Age Range
- Number of Age Cohorts
- Age Entropy
- Age Standard Distribution
- Ethnicity HHI
- % Former Employees
- Relatives on the Board

See appendix for a full result of the regression [Table D].

These findings suggest that centrality nuances play a role in shaping firm returns, while the impact of diversity metrics on performance requires further investigation. These results highlight the significance of leadership centrality and imply that higher centrality within a firm's leadership may potentially lead to greater returns [Table D].

Benefits to the Sponsoring Company

Although our results were preliminary in showing that some centrality metrics were statistically significant in determining firm forward returns, we believe that these results may still have some merit in benefiting Vanguard's QEG.

Some previous studies (see References) have shown that firm diversity does in fact lead towards better firm performance, whether this is the result of better employee retention, more creativity and viewpoints, or other reasons is unsure; but the correlation exists. But our study did not measure overall firm diversity on an operational level, instead it measured these metrics on a strategic one: focusing more so on the leadership and direction of the firm rather than those working in its day-to-day.

Our results don't disprove the importance of diversity and wide reaching centrality on firm boards and c-suite level cabinets, but instead demonstrates a new area of focus for future research and a myriad of questions to be answered.

Is it more important to have a firm's strategic vision be centralized rather than be covered with multiple points of view? Are firms who focus on the benefits of adaptability and problem solving abilities at the operational level instead of the strategic better performing than those who don't? How well can firms leverage their diverse human capital in a tumultuous market environment, and can they do better?

Next Steps

As our results have demonstrated a complex relationship between diversity/centrality and firm performance, our team believes there are still steps to be taken for further analysis.

One of the limitations of this project is in the issues of data quality and quantity within the biographical dataset. As mentioned previously, between financial and biographical information, biographical data on publicly traded firms is at best shaky. Information on board membership and the accompanying personal information for members isn't commonly reported, and any data gathered relies either on best-guess extrapolation or information derived from machine learning prediction models. To improve the quality of biographical data, collaboration with companies directly could be sought to ensure accurate and comprehensive information about board members and executive teams.

There was also a distinct lack of a "voting power" feature that could've be used to calculate the amount of true influence each executive member has on the direction of the firm, meaning that we were only able to detect the effects of surface level diversity and centrality, not the true "operating" diversity of the firm. Another potential avenue for future research could be to develop a proxy for 'voting power' within firms to gain a deeper understanding of the dynamics of diversity and centrality.

It's also important to note the potential for market capitalization bias when studying centrality and diversity in firms. Market cap, which represents the total dollar market value of a company's outstanding shares, is often closely tied to company size and public perception. Therefore, larger companies tend to have more resources for their leadership structure and diversity, and could significantly affect the results of the analysis. For future research, correcting for this bias is highly recommended.

Finally, our analysis was contained within the scope of ISS' director's database; which only maintains the board makeups of the firms publicly traded on the S&P 1500. This means all our biographical data is based on the top 1500 firms traded by the S&P 1500, and not inclusive of possibly smaller firms or industries. As with any data analysis, and especially in a biased set like ours; a larger and more *diverse* sample size may be beneficial to fully capture the true effects of diversity on returns. Further analysis into biases could help get a more holistic view into the centrality of company executives.

Appendix A

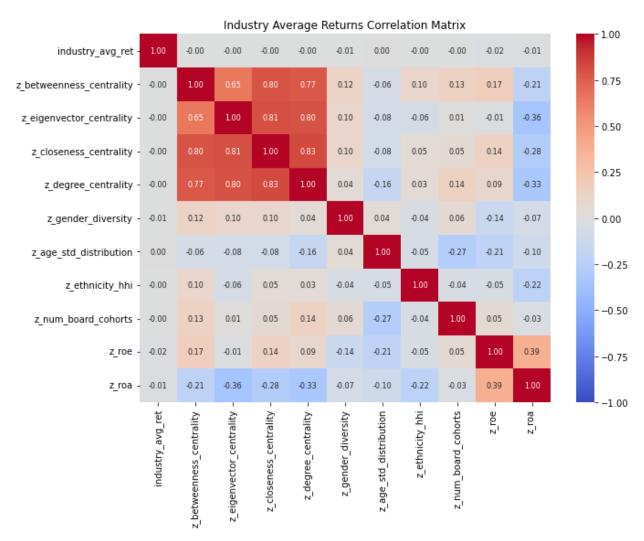
[Table A]

Data Dictionary

Feature Name	Description
Betweenness Centrality	Value of importance based on the number of shortest paths between that node and other nodes in the network. Shortest path is calculated using Djikstra's algorithm.
Eigenvector Centrality	Value from the axis of greatest commonality/centrality amongst the elements of our feature space. Its value is proportional to the sum of the eigenvector centrality values of its neighbors, weighted by the strength of the connections.
Degree Centrality	Number of connections a node has regardless of the specific nature of the connection. It measures the importance of the node within a network
Closeness Centrality	Measure of how easily a node can reach all other nodes on the network. It is calculated by taking the sum of the shortest path lengths.
Ethnicity Herfindahl-Hirschman Index	The calculated Herfindahl–Hirschman index of the ethnic diversity, measures overall
Gender Diversity	Percentage of non-male identifying individuals present
Number of Age Cohorts	The number of age cohorts split by the weighted average age range
Number of Board Cohorts	The number of cohorts split by the weighted average number of boards other than the current firm each board member is sitting on.
Age Entropy	Measures the level of uncertainty or randomness in the age distribution.
Age Mean	Mean age of a firm's board for a given year
Return on Capital Employed	Financial ratio that measures the profitability and efficiency of a company's capital investments.

[Table B]

Correlation Matrix



[Table C]

Regression Results with Selected Featureset

	Regressor	P-Value	t-Statistic
0	z_betweenness_centrality	0.447102	-0.760256
1	z_eigenvector_centrality	0.214710	-1.240722
2	z_closeness_centrality	0.307093	1.021344
3	z_degree_centrality	0.455983	-0.745479
4	z_gender_diversity	0.010520	-2.558277
5	z_age_mean	0.005389	-2.782846
6	z_age_range	0.890165	0.138095
7	z_num_age_cohorts	0.584026	0.547515
8	z_age_entropy	0.171617	-1.367031
9	z_age_std_distribution	0.462943	0.734010
10	z_ethnicity_hhi	0.184725	1.326352
11	z_former_emp_diversity	0.770519	0.291697
12	z_num_board_cohorts	0.592892	-0.534650
13	z_relatives	0.949575	0.063240

[Table D]

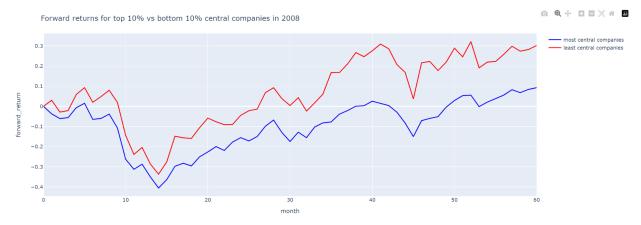
OLS Regression Results							
Dep. Variable:	forward_return	R-square	d:	=======	0.035		
Model:	OLS		•		0.034		
Method:	Least Squares			_	40.20		
	on, 05 Jun 2023	•	statistic):		75e-124		
Time:	03:41:53	0	linood:		-47658.		
No. Observations:	17712	AIC:			535e+04		
Df Residuals:	17695	BIC:		9.	548e+04		
Df Model:	16						
Covariance Type:	nonrobust						
	coef	std err	t		[0.025	0.975]	
const	1.0681	0.027	39.797	0.000	1.015	1.121	
z_pe	0.1650	0.024	6.738	0.000	0.117	0.213	
z_roce	0.5849	0.027	21.547	0.000	0.532	0.638	
z_betweenness_centrali	ty -0.3230	0.050	-6.489	0.000	-0.421	-0.225	
z_eigenvector_centrali	ty -0.0674	0.028	-2.443	0.015	-0.121	-0.013	
z_closeness_centrality	0.1164	0.037	3.107	0.002	0.043	0.190	
z_degree_centrality	0.0334	0.060	0.557	0.577	-0.084	0.151	
z_gender_diversity	-0.0117	0.029	-0.398	0.691	-0.069	0.046	
z_age_mean	0.0461	0.028	1.676	0.094	-0.008	0.100	
z_age_range	0.0018	0.078	0.024	0.981	-0.151	0.155	
z_num_age_cohorts	-0.0013	0.041	-0.033	0.974	-0.081	0.079	
z_age_entropy	0.0577	0.034	1.685	0.092	-0.009	0.125	
z_age_std_distribution	-0.0382	0.084	-0.455	0.649	-0.203	0.126	
z_ethnicity_hhi	0.0285	0.027	1.037	0.300	-0.025	0.082	
z_former_emp_diversity	0.0125	0.027	0.461	0.645	-0.041	0.066	
z_num_board_cohorts	0.1965	0.035	5.563	0.000	0.127	0.266	
z_relatives	-0.0396	0.028	-1.407	0.159	-0.095	0.016	
Omnibus:	39451.087	Durbin-W	atson:		0.362		
Prob(Omnibus):	0.000	Jarque-B	era (JB):	437184	254.104		
Skew:	20.532	Prob(JB)	:		0.00		
Kurtosis:	771.573	Cond. No			7.92		

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[Section One]

DOW 30 Forward returns for 2007 - 5 Years

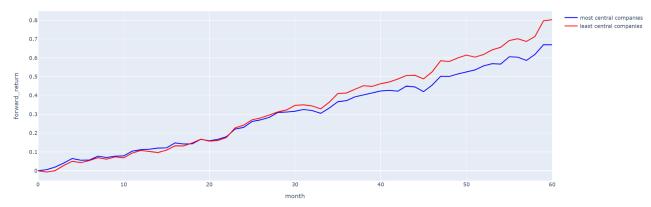


[Table E]

ISS Universe Average Forward Returns across 5 Years for Yearly Portfolios



Average Forward Returns for top 10% vs bottom 10% central companies (2008-2018)



Average Forward Returns for top 10% vs bottom 10% diverse companies based on board mean age (2008-2018)



[Table F]
Regression P-values

+	4	L	L	L
 	Regressors	t-values	p-values	R-squared
4	roce	22.0495	0.435538	0.0267187
2	pe	7.90531	2.59537e-106	0.00351633
6	roce+pe	6.80463	0.0202766	0.0292569
3	centrality	3.32017	0.000901441	0.00450589
10	roce+pe+cent	2.35131	0.0187185	0.0327556
7	roce+cent	2.3214	0.155282	0.0302845
1	diversity	-0.779768	0.105266	0.00162911
8	roce+pe+div	-1.17223	0.159475	0.0306693
5	roce+div	-1.18344	0.236651	0.0281047
11	roce+pe+cen+div	-1.4069	0.24112	0.0350708
9	roce+cen+div	-1.42118	1.04555e-11	0.0325953
0	de	-1.61993	2.82865e-15	0.000148152
T	T			

References

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. Journal of Financial Economics, 94(2), 291-309.

Center for Research in Security Prices. (2023). CRSP stocks database. Wharton Research Data Services. https://wrds-web.wharton.upenn.edu/

Gompers, P. A., Mukharlyamov, V., & Xuan, Y. (forthcoming). The Cost of Friendship. Journal of Financial Economics. https://doi.org/10.1016/j.jfineco.2022.05.001

ISS: Institutional Shareholder Services. (2023). Compustat database. Wharton Research Data Services. https://wrds-web.wharton.upenn.edu/

Larcker, D. F., So, E. C., & Wang, C. C. Y. (forthcoming). Boardroom Centrality and Firm Performance. Journal of Accounting and Economics.

Lu, Y., Naik, N. Y., & Teo, M. (2015). Diverse Hedge Funds. Review of Financial Studies, 28(6), 1718–1758. https://doi.org/10.1093/rfs/hhu109

Standard & Poor's Compustat. (2023). Compustat database. Wharton Research Data Services. https://wrds-web.wharton.upenn.edu/

Trash Can Reborn

This project's objective is a correlation analysis of the impact of a firm's board of directors and c-suite diversity and centrality on overall share price returns. This objective is achieved by a biographical analysis of firm leadership, which requires a dataset and further feature analysis to quantify multiple diversity metrics. This "biographical data" was paired alongside the forward returns for each given firm. Due to the nature of corporate reporting, the major difficulty in data collection is in determining the demographic makeup of both C-suite executives and the governing members of the board over time. While on the other hand, thanks to the nature of corporate reporting, our team is able to accurately determine the financial performance and make up of any given publicly traded corporation.

This dichotomy provides us with two different avenues of data collection, split into the two groups of "financial" and "biographical" data. Financial data needed to consist of a given firm's performance on the stock market, measured in the change in price, as well as a suite of financial metrics, ratios, and general information. Biographical data on the other hand was much more complex; with the ability to derive not only leadership centrality information, but also information in relation to ethnicity, age, and gender. Furthermore we needed to decide how to define and quantify multiple diversity metrics.

Finally, it will need to be a dataset able to provide statistically significant results. This means the project will encompass as many firms as possible and the longest period of time within the limitations of our chosen dataset.

Biographical data was collected via the ISS: Institutional Shareholder Services Directors database, while financial data was collected via the CRSP and COMPUSTAT databases and subsequently joined together on the ticker of the stock and the adjoining date.

Using this data, our team was able to create a target variable, expressed through stock returns; and a feature set, featuring firm financial performance metrics, diversity and centrality measures, as well as any other information that is determined to be important to the prediction. This dataset is inclusive of 1500 firms over a 15 year period of time. See the data dictionary located in the appendix [Table A].

We implemented two separate Ordinary Least Squares (OLS) models: one incorporating control variables (z_pe, z_de, z_roce), and the other including both biographical and financial control metrics. The control model yielded an R-squared value of 0.029, while the biographical metrics model showed an R-squared value of 0.035. Taking into account lagging effects, the ideal time horizon to lag the model features came out to 4 years, giving us an R-squared of 0.071. Although most of the centrality measures exhibited statistical significance at the 1% level, with the exception of degree centrality, the diversity measures did not demonstrate statistical significance, except for z_num_age_cohorts. These findings suggest that centrality nuances play a role in shaping firm returns, while the impact of diversity metrics on performance requires further investigation. These results highlight the significance of leadership centrality and imply that lower centrality within a firm's leadership may potentially lead to higher returns [Table D].

OLS Regression Results						
Dep. Variable: Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	forward_return OLS Least Squares Mon, 05 Jun 2023 03:41:53 17712 17695 16 nonrobust	R-square Adj. R-s F-statis Prob (F- Log-Like	quared: tic: statistic):	9.	0.035 0.034 40.20 75e-124 -47658. 535e+04 548e+04	
=======================================						_
	coef	std err	t	P> t	[0.025	0.975]
const z_pe z_roce z_betweenness_central z_eigenvector_central z_closeness_centrality z_degree_centrality z_gender_diversity z_age_mean z_age_range z_num_age_cohorts z_age_entropy z_age_std_distributio z_ethnicity_hhi z_former_emp_diversit z_num_board_cohorts z_relatives	ity -0.0674 y 0.1164 0.0334 -0.0117 0.0461 0.0018 -0.0013 0.0577 n -0.0382 0.0285	0.027 0.024 0.027 0.050 0.028 0.037 0.060 0.029 0.028 0.078 0.041 0.034 0.084 0.027 0.027 0.035 0.028	6.738	0.000 0.015	1.015 0.117 0.532 -0.421 -0.121 0.043 -0.084 -0.069 -0.008 -0.151 -0.081 -0.009 -0.203 -0.203 -0.205 -0.041 0.127 -0.095	1.121 0.213 0.638 -0.225 -0.013 0.190 0.151 0.046 0.100 0.155 0.079 0.125 0.126 0.082 0.066 0.266 0.016
Omnibus: Prob(Omnibus): Skew: Kurtosis:	39451.087 0.000 20.532 771.573	Durbin-W Jarque-B Prob(JB) Cond. No	atson: era (JB): :	437184	0.362 254.104 0.00 7.92	0.010

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

with specific centrality measures being correlated with firm performance,