

Measuring the Impact Factor of Agents within an Organization Using Communication Patterns

Ignacio Palacios-Huerta¹ Andrea Prat²

22 September 2010

Abstract: Organizational economics predicts that communication patterns within an organization should reflect the relative value of their members to the organization. We propose to measure the impact factor of an agent by applying the Invariant Method—also known as Google’s PageRank algorithm—to electronic communication data. To explore the validity of this measure, we analyze email exchanges among the top executives of a large retail company. We construct their individual impact factors based only on email patterns and we compare them to standard economic measures of organizational importance. We find that: (i) The impact-factor ranking of executives mirrors perfectly their hierarchical ranking; (ii) Impact factor variability is significantly correlated with salary differences; (iii) Subsequent promotions (dismissals) affect executives with unusually high (low) impact factors. We conclude that simple communication-based impact factors may be a useful tool to measure the relative importance of agents in organizations.

The great mass of economic activity and much of social activity takes place not in the market but within the internal environments of organizations. As Kenneth J. Arrow (1) emphasized in his classic work on the economics of organizations, “the purpose of organizations is to exploit the fact that many (virtually all) decisions require the participation of many individuals for their effectiveness.” The decisions that individuals take are a function of the information that they have, and the acquisition of information is itself the result of their own decisions. This means that the actual structure and behavior of an organization depends heavily upon its internal structure of information and communication. That is, the value of creating organizations of a scope more limited than the market as a whole is partially determined by the characteristics of the network information and communication flows.

Thus internal communication is a central activity of most organizations. As information processing and transmission requires time, energy, and resources, communication patterns between agents should reflect the goals and the values of the organization. To optimize communication, organizations tend to rely on a hierarchical structure: raw data is processed at the bottom, while agents in charge of high-level decisions receive more synthetic information (2, 3, 4, 5).

¹ Department of Management, London School of Economics, Houghton Street, London WC2A 2AE, UK. ² Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, UK.

Organizations should then display two related features: the familiar hierarchical structure and a pattern of communication that reflects the chosen hierarchy. This insight occupies a central role in the study of the economics of organizations and has led researchers to hypothesize that “if we record the frequency of communication between different nodes, we [will] find that the pattern is not uniform but highly structured. In fact, the pattern of communication *frequencies* [should] reflect, approximately, the pattern of authority. ... We should not imagine an even flow of messages from all nodes of the network to all others, but instead a highly patterned flow that is characteristic of nearly decomposable structures.” (6, p. 151, italics added). Thus agents who make more important decisions on behalf of the organization should occupy more central positions in the organization’s communication network.

To test this hypothesis, we need to measure the “value” of individual agents within the organization and their importance within the communication network. The former can be related to standard organizational observables, like formal rank, salary, and career trajectory (e.g. promotions and dismissals). For the latter, we draw inspiration from the literature on link analysis in networks. Among all possible ranking methods, only the Invariant Method satisfies four natural properties: invariance to reference intensity, weak homogeneity, weak consistency and invariance to splitting of nodes (7). The Invariant Method, also known as PageRank (8, 9) forms the basis of Google’s search engine. This method is also increasingly used to compute the impact factor of scientific journals (10).

It is useful to explore this application because our organizational setting bears a parallel to academic publishing. The importance of a journal depends not only on how often other publications cite it, but also on how “important” those publications are. The Invariant method is the unique fixed point of a specific operator which calculates the positive eigenvector of an appropriately adjusted matrix of citations. It assigns to an entity i a value that is a weighted average of some function of the citations it gets. Here not all citations have the same value: citations by important entities are more valuable than citations by less important entities, and the importance of the entity is determined *endogenously* and *simultaneously* with the importance of all other entities.

In our organizational context, we replace “A cites B” with “A sends an email to B.” The underlying idea is that agents prioritize the time they devote to communication, directing it where they believe it is most useful (11, 12, 13, 14). The importance of an agent is reflected in the effort that other agents put into emailing them, weighted by those agents’ importance which in turn is endogenously determined within the communication network. We propose to measure such importance through impact factors obtained through the application of the Invariant Method to email traffic patterns.

To validate the use of this measure of agent value, we apply it to a database of email communication between members of a particular organization. We then compare our impact factors with other, independently obtained standard economic indicators of individual productivity.

Communication databases have already been used to understand the underlying interactions among agents both within organizations (15) and in other social contexts (16). The novel

contribution of the present work is to propose and validate the use of impact factors using communication data to understand organizations.

Impact factors are informative only when applied to a set of comparable entities. In the case of publishing, such set is often identified with a scientific discipline. In an organizational setting, it is natural to study a group of agents who belong to the same organization and have a relatively similar job. In our application, we focus our attention on all the agents in a managerial position who have an executive contract in a specific company.¹

Our data contains information on wages, the precise hierarchical structure and the volume of email communications – but not their content – for 15 years (1995-2009) for all the executives of a large European company. This company is one of the largest retailers in the world in terms of size, revenue and profit, and is active in three types of retail distribution: hypermarkets, supermarkets and hard discounters. The data corresponds to its operations in one European country and includes the information for all the company's managers in that country.

The company is managed by 10 high-ranking chief executives and 42 top senior managers (directors). The ten C-level executives are the following:

1. The CEO (Chief Executive Officer) or Managing Director is the most important person in the company, reporting to the Chairman of the Board and board members.

The next two executives in terms of responsibilities are:

2. The COO (Chief Operating Officer), who is the leading corporate officer with responsibility for the daily operation of the company (in some countries it carries the title of President), and
3. The CPO (Chief Product Officer), who is responsible for the product purchases of the company;

The rest of the chief executives are:

4. The CFO (Chief Financial Officer) is the corporate official in charge of the company's finances;
5. The CIO (Chief Information Officer) is responsible for the company's internal information systems;
6. The CMO (Chief Marketing Officer) is responsible for the company's marketing strategy;
7. The CHRO (Chief Human Resources Officer) is responsible for the company's human resources policy;
8. The CEXO (Chief Expansions Officer) is responsible for the expansion of the company within the country;
9. The CLO (Chief Logistics Officer) is responsible for the logistics of the company;

¹ Like most continental Europe, the country where our company operates uses collective labor contracts. This creates a clear and stable distinction between employees who have a managerial contract and those who do not.

10. The CSO (Chief Supermarket Officer) is responsible for the operations of the supermarkets.

In addition to these top executives, the organization has 42 Directors, or senior managers of managers, who are typically responsible for a major business function. They directly report to the corresponding C-level executive.

We compute the impact factor of each executive in our sample by applying the Invariant Method to the matrix of email communications among these corporate officers. An entry in the matrix is simply the number of emails that individual i received from individual j . When an email is sent to more than individual, all the individuals that receive the email are assigned the same share. If an individual is present in the sample at a given position in the firm for only part of the period, we compute her impact factor based on email traffic during that period. This methodology guarantees that, if the relative importance of an individual relative to her colleagues is stable over time, her impact factor is the same independent of the subperiod it is computed on. Finally, the measures of impact factor are normalized by assigning a value of 100 to the individual with the top impact factor.

We compare these impact factor measures, obtained solely from email data, with three sets of organizational measures: the formal hierarchy, salary data, and promotions and dismissals.

Figure 1 depicts the organigram of the company. Next to each executive, we report his or her impact factor. We find that there is a striking correspondence between the ranking derived from email-based impact factors and the hierarchical ranking. In fact, the former never contradicts the latter: for all the 52 employees there is no single case in which an agent with a superior ranked has a lower impact factor than an agent with an inferior rank. More precisely:

1. The individual with the top impact factor is the CEO;
2. The two main executives that report to the CEO have a much greater impact factor (81 for the COO and 69 for the CPO) than the remaining executives whose impact factors range from 6 to 18;
3. In *each* of the divisions of the organization, the C-level executive *always* obtains a greater impact factor than all of the Directors that report to him.

Individual impact factors are stable over time: less than 3% of the total year/individual variance is due to within-individual variance. There appears to be permanent differences between individual managers (See Supplementary Material).

Given the stability of individual impact factor, this one-to-one correspondence between rank and impact factor holds also for subperiods. For instance, the three points above are also true if one restricts attention to each half of the sample or to the last five years.

Second, we study the connection between impact factor and financial remuneration. Compensation should reflect individual contributions to the organization. If the hypothesis

that communication-based impact factors capture the intrinsic value of an agent, managers with higher impact factors should be paid more.

Table 1 reports various correlation coefficients between impact factors and compensation (salaries plus bonuses) both for the executive positions and within each division. The correlations are positive and significant ($p < 0.01$) within the subset of the C-level executives. And the same is true within the divisions for each of the divisions of the company. This result is confirmed by additional analysis reported in the SOM, where compensation is regressed on rank and impact factor. Everything else equal, a 10-point impact factor increase is associated with a 5% pay increase.²

Finally, we turn to career progression and study whether the decision to promote or dismiss an executive or a director is related to her impact factor. As noted above, within-person impact factors are stable over time. Under the null hypothesis, therefore, the impact factor of the dismissed/promoted employee prior to the decision should not differ significantly from the impact factor of the other managers. (This hypothesis does not exclude, for instance, that the impact factor of a manager increases after her promotion solely because she now has a higher rank.)

The simple quantile analysis in Table 2 shows that promotions and dismissals appear to be strongly related to impact factors: *All* individuals who are promoted were in the top quartile of distribution of impact factors before the promotion decision, and *all* individuals who were fired were in the bottom quartile.

In Table 3 we test our null hypothesis more formally. We find that the impact factor is a significant determinant of whether an agent will be promoted or dismissed. In particular, the results indicate that on average a 10-point increase in the impact factor increases the likelihood of promotion by 64 to 67 percent, whereas a 10-point decrease in the impact factor increases the likelihood of dismissal by around 55 percent.

In conclusion, in our dataset individual impact factors appear to be excellent predictors of a range of standard economic indicators of individual value, such as rank, compensation, and career development within organizations. While our findings apply only to the firm that we study, our simple methodology can be implemented in any organization with an email database. It can also be extended to other forms of electronic communication, such as social networks. The methodology is simple and there is a wealth of electronic communication databases. We hope that our findings will stimulate other researchers to analyze the link between intra-firm communication and the way firms allocate, organize, develop, and reward human capital, as first hypothesized by Kenneth Arrow (1).

While the ability of impact factors to predict promotions and dismissals is of interest to researchers, it may also create risks for firms. At first sight, this predictive power may tempt companies into analyzing email patterns as part of their human resources policy, in order to identify promising and problematic cases early on. On second thought, however,

² There also appears to be a positive interaction between rank and impact factor. The link between impact and pay is stronger for higher-rank executives.

this policy may lead to unwanted consequences. Agents who know that their actions are observed may engage in inefficient activities, so much so that the organization might actually prefer to commit not to observe those actions (17). If communication data is used to decide salaries and promotions, groups of agents may collude to generate email traffic. An interesting question – which we leave to future research – is how an organization can design an informative but non-manipulable way to use communication data to select and motivate its members.

Finally, the identification and understanding of patterns of human activity have important consequences beyond organizations, reaching areas as diverse resource allocation, disease spread and different social systems (18). Impact factor measures computed using data on the intensity of human interactions represent a promising avenue for future research.

References

1. K. J. Arrow, *The Limits of Organization*, Norton, 1974.
2. R. Radner, The Organization of Decentralized Information Processing. *Econometrica* 62, 1109-1146, 1993.
3. T. Van Zandt, Real-Time Decentralized Information Processing as a Model of Organizations with Boundedly Rational Agents. *Review of Economic Studies* 66, 633-658, 1999.
4. P. Bolton, M. Dewatripont, The Firm as a Communication Network. *Quarterly Journal of Economics* 109, 809-834, 2004
5. L. Garicano, Hierarchy and the Organization of Knowledge in Production. *Journal of Political Economy* 108, 874-904, 2000.
6. H. Simon “The Impact of Electronic Communication on Organizations.” In R. Wolff (Ed.), *Organizing Industrial Development*. Berlin: Walter de Gruyter, 1986.
7. I. Palacios-Huerta, O. Volij, The measurement of intellectual influence. *Econometrica* 72, 963-977, 2004.
8. G. Pinski, F. Narin. Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics *Information Processing and Management* 12, 297–312, 1976.
9. L. Page, S. Brin, R. Motwani, T. Winograd. The Pagerank Citation Ranking: Bringing Order to the Web. Technical Report. Stanford InfoLab, 1999.
10. C. T. Bergstrom. Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News* 68 (5), May 2007.
11. T. Van Zandt. Information Overload in a Network of Targeted Communication. *RAND Journal of Economics* 35(3): 542-560, Autumn 2004.
12. J. Cremer, L. Garicano, and A. Prat, Language and the Theory of the Firm. *Quarterly Journal of Economics* 122, 373-407, 2007.
13. W. Dessein and T. Santos, Adaptive Organizations. *Journal of Political Economy*, 114(5): 956--995, 2006.
14. A. Calvó-Armengol, J. de Martí, A. Prat, Endogenous Communication in Complex Organizations. Working paper, London School of Economics.

15. Guimera, R, Danon, L, Diaz-Guilera, A, Giralt, F, Arenas, A. Self-similar community structure in a network of human interactions. *Phys. Rev. E* 68,. 065103 (2003)
16. R. Fisman, S. Iyengar, E. Kamenica, and I. Simonson, Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment. *Quarterly Journal of Economics* 121(2): 673-697, May 2006.
17. A. Prat. The Wrong Kind of Transparency. *American Economic Review*, 95(3): 862-877, June 2005.
18. R. D. Malmgren, D. B. Stouffer, A. S. L. O. Campanharo, L. A. N. Amaral, On universality in human correspondence activity. *Science* 325, **1696** (2009)

Table 1 - Correlation Coefficients between Impact Factors and Compensation (Salaries plus bonuses)

Among C-level Executives: 0.83 ($p < 0.01$)

Within each Division:

Operations	0.52 ($p < 0.01$)
Product	0.58 ($p < 0.01$)
Finance	0.45 ($p < 0.01$)
Information	0.37 ($p < 0.01$)
Marketing	0.60 ($p < 0.01$)
Human Resources	0.52 ($p < 0.01$)
Expansion	0.73 ($p < 0.01$)
Logistic	0.25 ($p < 0.01$)
Supermarkets	1.00 ($p < 0.01$)

Note: In parentheses the p -value of the test of whether it is significantly different from zero.

Table 2 - Distribution of Normalized Impact Values and Source of Promotions and Dismissals

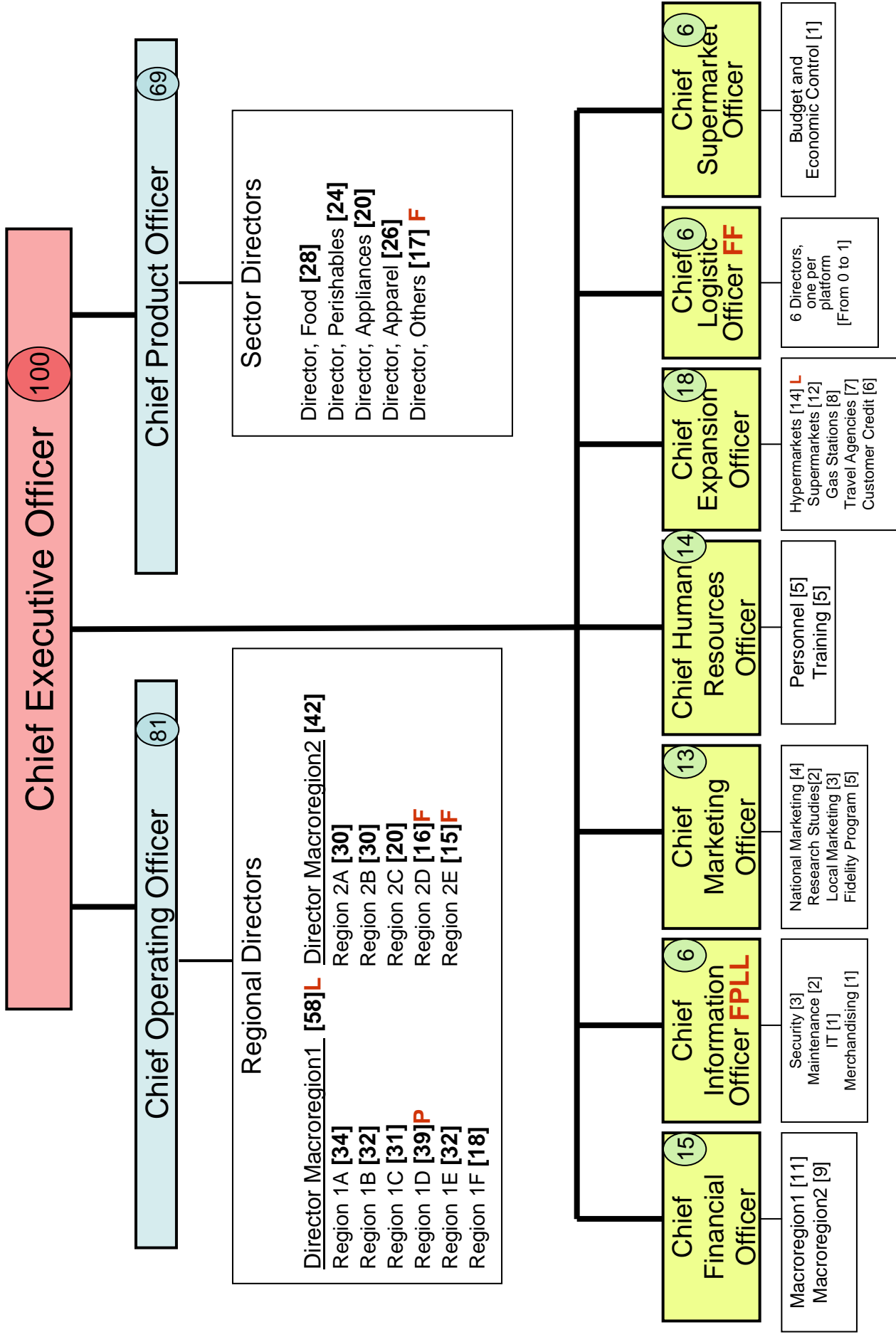
	<i>N</i>	Bottom quartile			Top quartile
		Q1	Q2	Q3	Q4
Operations Division: Regional Directors	13	3	4	3	3
Product Division: Sector Directors	5	1	2	1	1
Rest of Divisions: Other Directors	24	6	6	6	6
Source of:		Q1	Q2	Q3	Q4
Promotions	4	0	0	0	100%
Firings	6	100%	0	0	0

Note: Impact values are normalized by subtracting the mean impact factor and dividing by the standard deviation within each division. The Directors are then grouped into four quartiles from the top quartile (Q4) to the bottom quartile (Q1). Promotions include both within the company and to other companies.

Table 3 - Probit and Logit Regressions for Promotions and Dismissals

	Promotions		Dismissals	
	Probit	Logit	Probit	Logit
Constant	-2.438*** (0.00)	-4.537*** (0.00)	-1.294*** (0.00)	-2.256*** (0.00)
Impact Factor	0.038** (0.02)	0.073** (0.02)	-0.013** (0.56)	-0.023** (0.61)
Department and Ranking Controls	Yes	Yes	Yes	Yes
Akaike Information Criterion	29.77	29.70	35.48	35.67

Notes: *p*-values in parenthesis, ** denotes significant at the 5% level, and * significant at the 1% level. All regressions include interactions between impact factors and ranking level.



Note: **F** denotes an agent who was fired, **P** one who was promoted and **L** one who left voluntarily to another firm. Multiple letters indicate that multiple employees occupied that position during the period of analysis. The impact factor that is reported for each position corresponds to the average impact factor that all the agents that ever occupied that position had during the period of analysis. See SOM.

Table S1

This table reports the average annual impact factor from 0 to 100 and, in parentheses, the standard deviation for each executive position. For each year during the period of analysis 1995-2009, the impact factor is computed using the matrix of email communications among the executive positions during that year. † denotes those positions that were occupied by more than one executive during the period of analysis; for these positions the average and standard deviation for each executive is reported in Table S2. An average impact factor below 1 is simply reported as 1.

Chief Executive Officer	100 (0)		
Chief Operating Officer	81 (3.38)	Chief Product Officer	69 (3.00)
Director, Macroregion 1	58†	Director, Food	28 (1.22)
Region 1A	34 (1.24)	Director, Perishables	24 (1.16)
Region 1B	32 (0.75)	Director, Other	17 †
Region 1C	31 (0.62)	Director, Appliances	20 (0.32)
Region 1D	39 †	Director, Apparel	26 (2.33)
Region 1E	32 (0.81)		
Region 1F	18 (0.71)	Chief Strategy Officer	18 (1.29)
Director, Macroregion 2	42 (2.33)	Hypermarkets	14 †
Region 2A	30 (0.90)	Supermarkets	12 (0.80)
Region 2B	30 (2.34)	Gas Stations	8 (0.00)
Region 2C	20 (0.24)	Travel Agencies	7 (0.44)
Region 2D	16 †	Customer Credit	6 (0.33)
Region 2E	18 †		
Chief Financial Officer	15 (0.35)	Chief HR Officer	14 (3.24)
Financial Dir. Macroregion 1	11 (0.41)	Personnel Director	5 (0.50)
Financial Dir. Macroregion 2	9 (0.40)	Training Director	5 (0.62)
Chief Information Officer	6 †	Chief Logistics Officer	6 †
Security Director	3 (0.01)	Director 1	1 (0.00)
Maintenance Director	2 (0.00)	Director 2	1 (0.00)
IT Director	1 (0.00)	Director 3	1 (0.02)
Merchandising Director	1 (0.00)	Director 4	1 (0.01)
		Director 5	1 (0.00)
		Director 6	1 (0.00)
Chief Marketing Officer	13 (0.59)		
National Marketing Director	4 (0.02)		
Research Director	2 (0.00)		
Local Marketing Director	3 (0.00)		
Fidelity Program Director	5 (0.03)		
Chief Supermarkets Officer	6 (0.78)		
Director Control Economico	1 (0.00)		

Table S2

This table considers the positions that were occupied by more than one person during the period of analysis. It reports the average impact factor (IF), and in parenthesis its standard deviation, of each of the employees in that position during the period of time they were employed at the firm. For the type of separations: “F” denotes that the executive was dismissed and “P” that he or she was promoted within or to another firm. The symbol “-” denotes that he or she still works at the firm at the end of 2009.

Position: Director Macroregion 1	IF (std.dev)	Separation type
#1. Period: 1/1992-3/2001	61 (3.58)	P
#2. Period: 4/2001- -	55 (3.21)	-
Position: Director Region 1D	IF (std.dev)	Separation type
#1. Period: 9/1994-3/2001	43 (2.88)	P
#2. Period: 4/2001- -	36 (2.02)	-
Position: Director Region 2D	IF (std.dev)	Separation type
#1. Period: 1/1995-12/2003	10 (1.33)	F
#2. Period: 1/2004- -	21 (2.05)	-
Position: Director Region 2E	IF (std.dev)	Separation type
#1. Period: 10/1991-9/1999	10 (2.10)	F
#2. Period: 10/1999- -	19 (3.01)	-
Position: Sector Director Others	IF (std.dev)	Separation type
#1. Period: 9/1994-12/2004	12 (0.33)	F
#2. Period: 1/2005- -	21 (2.15)	-
Position: Chief Information Officer	IF (std.dev)	Separation type
#1. Period: 1/1993-12/1996	5 (1.00)	F
#2. Period: 1/1997-9/2002	12 (3.02)	P
#3. Period: 10/2002-12/2007	6 (0.22)	F
#4. Period: 1/2008- -	3 (0.08)	-
Position: Director Hypermarkets	IF (std.dev)	Separation type
#1. Period: 1/1999-12/2003	16 (1.32)	P
#2. Period: 1/2004- -	10 (0.55)	-
Position: Chief Logistic Officer	IF (std.dev)	Separation type
#1. Period: 9/1993-12/2000	6 (0.69)	F
#2. Period: 1/2001-12/2007	6 (1.34)	F
#3. Period: 1/2008- -	8 (2.02)	-

Table S3

This table reports two OLS regressions of compensation (salary and bonuses) on Impact Factors (IF) and Rank (1 for C-level executives, and 0 for Directors). t-statistics reported in parentheses.

Variable	(1)	(2)
Intercept	30.86*** (0.000)	28.76*** (0.000)
IF	0.50** (0.042)	0.53** (0.038)
Rank	20.03*** (0.005)	18.33*** (0.004)
IF*Rank		1.31*** (0.004)