African Journal of Science, Technology, Innovation and Development

ISSN: 2042-1338 (Print) 2042-1346 (Online) Journal homepage: http://www.tandfonline.com/loi/rajs20

Multivariate analysis to research innovation complementarities

Hernán Alejandro Morero & Pablo Ortiz

African Journal of Science, Technology, Innovation and Development, 2017
https://doi.org/10.1080/20421338.2017.1355586

Multivariate analysis to research innovation complementarities

Hernán Alejandro Morero[1*] and Pablo Ortiz[2]

[1]*Centro de Investigaciones y Estudios sobre Cultura y Sociedad (CIECS) – CONICET y UNC, Córdoba, Argentina*
[2]*Facultad de Ciencias Económicas, Departamento de Economía y Finanzas, UNC, Córdoba, Argentina*
**Corresponding author email: hernanmorero@eco.uncor.edu*

It is widely recognized that orthodox economics is obsessed with econometrics tools. However, econometrics techniques have a limited capacity to deal with qualitative variables coming from surveys. This paper presents a defence of the use of statistical methods, in particular multivariate analysis, which is the overall objective of the paper. Multivariate analysis is a set of methods that can be used when the problem that arises implies multiple dependent or interdependent variables of a qualitative nature. We considered an issue in the literature to probe multivariate analysis in a particular topic, namely: the question of innovation complementarities. We analyzed the presence of complementarities between internal and external innovation activities in 257 software firms from Argentina during the period 2008–2010, comparing the consideration of the problem of complementarities with the more modern complementarity econometrical tests, super and sub modularity tests arising from diverse firm-innovation function estimations (OProbit, Tobit and Probit), with the engagement of the same issue with multiple factor analysis and cluster techniques. The results show not only that the same results obtained by the econometrical tools can be reached by multivariate analysis techniques, but also that multiple factor analysis and cluster techniques allow for better exploitation of the richness of qualitative data.

Keywords: innovation complementarities, multivariate analysis, software sector, supermodularity, plurality

Introduction

More than a decade and a half has passed since Parisian students complained about the narrowness of their economics education, the lack of realism in economics teaching, the uncontrolled use of mathematics as an end in itself, with the result that economics became an 'autistic science', and the post-autistic economics movement was founded in 2000 (Fullbrook 2003). Several similar petitions had been launched before, for example at Cambridge University (UK), Harvard (US), and the international open letter 'The Kansas City proposal'. Several heterodox journals were born or grew over these years, many networks of students in economics seeking critical thinking arose, and even a global association, the World Economic Association, was born in 2011.

However, besides that the fact that some heterodox and interdisciplinary approaches have gained terrain over the past years, the current dominance of orthodox economics has not been changed and the original claims have not lost their relevance. Scholars still claim the need for a plurality of approaches adapted to the complexity of the object analyzed. That comprises pluralism, both in theories and in methods; particularly, methods that allow critical thinking to take place, and to apprehend the qualitative and open nature of the social realm.

Econometrics has a prominent place in economics. It is considered the 'rigorous' empirical work in the discipline. Usually, in economics, particularly mainstream economics, the 'scientific status' of empirical research is sought as systematic explanation arrived at by a narrow set of methods, characterized mainly by hard testing and restrictive econometric regression exercises. In this manner, economics is typically perceived to be closer to the 'hard' sciences; and, in this sense, economics is isolated from other methodological approaches, a whole branch of social research and, indeed, from the advances and practices from other social disciplines (Downward and Mearman 2009).

Modern orthodox economics puts a particular, even obsessive, emphasis on mathematical modelling and on econometric testing practice. Lawson (1997) argues that this obsession is a symptom of a widespread deductivism in the discipline, assuming the ubiquity of covering laws of explanation, and the adherence to seeking strict regularities, whether or not these are derived from formally deductive or indicative premises (Downward and Mearman 2009).

This abrogated supremacy of econometrics to the empirical study of economic phenomena has been accused of leading to a systematic effort of the economists to explain and to response continuously to their persistent failure performance (Lawson 1997). Some authors even point out that the rise of the econometrics domain of economics empirical work is in part responsible for the fall of some heterodox paradigms in the discipline, oppressing the appearance and development of alternate theoretical voices (Colander and Landreth 2008). Furthermore, econometrics has other epistemological problems: the feasibility of the necessary conditions of quantification, the necessary non-necessary additive nature of the 'material' from social sciences (Downward and Mearman 2002), and the open nature of human agency, among others.

However, empirical methods seem to have an irreplaceable, or maybe an inevitable, role in economic inferences, as long as social reality still involves some empirical side as well. In that sense, there appears a necessity for an extended tool kit of methods for economists' empirical scientific work. More properly, there is a need, for example, for an extended acceptance of many other statistical techniques, and qualitative and interpretative methods, as 'scientific' as econometrics. Moreover, many econometricians now are somewhat skeptical as to whether the use of formal econometric methods has been more fruitful than less formal techniques (Backhouse 1998).

The main objective of this paper is to present a critique of the abrogated supremacy of econometrics and its supposed superior status in the empirical work of economics through a defence of alternate quantitative methods applied to a specific issue in the literature (innovation complementarities). In order to do so, we will compare the performance of typical hard econometrical tools against other alternate statistical strategies.

The quantitative arm of other social sciences, for example psychometrics or envirometrics, has developed other methods to deal with qualitative data, known as multivariate analysis techniques. In this paper, we make a comparison of the performance of two quantitative strategies – hard econometric estimation and testing vs. multivariate analysis – to address a problem widely established in the economic literature: the existence of complementarity or substitutability between internal and external innovative efforts of firms. Our working hypothesis is that the same results obtained from econometrical analysis can be obtained by multivariate techniques, with additional richness and details, suggesting a better performance of the latter.

This will be the main contribution of the paper. It also contributes to closing a particular gap in the empirical literature on innovation complementarities: the concern about the determinants of these complementarities. Traditional econometrical strategies have failed to clarify this gap. We will show how multivariate analysis techniques could contribute to a better understanding of this issue.

Firstly, we present the particular topic of application. The question of innovation complementarities is properly presented in the next section: to begin with, the diverse theoretical positions, and the dominant empirical studies and their results; and, afterwards, our proposal of the recognition of the minor studies that followed an alternate quantitative strategy combining statistical techniques.

The following section presents our case. There, we analyze the presence of complementarities between internal and external innovation activities in 257 software firms from Argentina during the period 2008–2010. We compare the results, applying the more modern complementarity econometrical tests – super and sub modularity tests arising from diverse firm-innovation function estimations – with multiple factor analysis and cluster techniques. Afterwards, there is a comparison of the performance of both methods. The outcomes show that not only can multivariate analysis techniques obtain the same results as the hard econometrical tools, but also that multiple factor analysis and cluster techniques allow for a better exploitation of the richness of qualitative data.

Innovation complementarities. How far can econometrics go?

In this section, we introduce the specific topic of application to compare both empirical techniques. Our study of innovation complementarities will show that hard econometric testing is a limit to our understanding. Here we present a very brief state of the art of this issue, theoretically and empirically.

Corporate innovation can be sourced by knowledge internally and/or externally. The effective integration of both sources is recognized as a key factor explaining a successful innovative performance, which is the reason there is a tradition in the literature of industrial economics studying the ways and degree of complementation or substitution of external and internal innovative activities.

Theoretically, in economics opposite arguments around the question of complementarities between innovative activities can be found. The most common explanations to sustain the prevalence of substitutability relations come from transaction costs (Coase 1937; Arrow 1962; Williamson 1985) and property right theories (Grossman and Hart 1986).

Rival theories, such as the resource-based view of the firm approach, mostly present arguments to sustain the hypothesis of complementarity, but are, to some extent, ambivalent about the merits of each type of knowledge source. Their stress on heterogeneous and inimitable assets and resources of firms seems an emphasis on the superiority of in-house activities (Teece 1986). Conversely, this approach highlights the benefits of knowledge sharing and cooperation, and has arguments grounded in the concept of absorptive capacity. For instance, complementarity could arise because it is necessary to have internal competences that allow effective absorption of external knowledge (Cohen and Levinthal 1989). In sum, this kind of argument from authors who have resource-based views or approaches of the firm mainly argue that internal knowledge creation activities usually reduce the inefficiencies of external acquisition and allow the modification and improvement of knowledge absorption from outside the firm. In that sense, complementarity relations between internal and external knowledge sources for innovation are more likely to arise.

Finally, from properly evolutionary theories and learning economics there are arguments that emphasize the importance of complementing external and internal knowledge for innovation. A differential characteristic of the innovation strategies of firms, and their performance, is the diverse combination of internal and external knowledge sources (Freeman and Soete 1997). Moreover, from learning economics, the distinction between diverse types of knowledge (know-how, know-what, know-why and know-who) (Lundvall and Johnson 1994; Foray and Lundvall 1998) establishes that internal and external sources should be combined in different learning and innovation modes: the science, technology and innovation (STI) mode, focused mainly on codified and technical knowledge management; and the doing, using and interacting (DUI) mode, more focused on the daily learning processes and the knowledge created in informal and formal interactions (Jensen et al. 2007).

Thus, it is a controversial topic because opposing arguments can be found in the literature that allow us to expect both substitutability and complementarity relations, and there is no unequivocal guidance to the best mix between internal and external sourcing.

The quantitative studies of innovation complementarities

Regarding the empirical study of this issue Mohnen and Röller (2005) recognized basically two quantitative strategies in the literature: the correlation and the direct approaches.

The most common econometric strategy has been the so-called correlation approach in which simple correlations between the variables, with or without controls, are analyzed. Table 1 summarizes the studies in this line, underlining their results around the question.

There it can be seen that there are studies that found that internal and external innovative activities tend to be substitutes both in developed economies (Pisano 1990; Blonigen and Taylor 2000; Love and Roper 2001) and in emerging and developing ones (Mytelka 1978; Fikkert 1994; Basant and Fikkert 1996). There are also studies that found complementarity relations (Arora and Gambardella 1990, 1994; Veugelers 1997; Braga and Willmore 1991; Deolalikar and Evenson 1989). Further, there are studies in this line that show ambivalent results, arguing that complementarity is sensitive to technical sectoral specificities (Audretsch, Menkveld, and Thurik 1996) or to structural aspects of the firm, such as size (Veugelers and Cassiman 1999). Thus, the empirical literature in the line of the correlation approach does not reach conclusive results.

The relevance of the results of the studies in this line has been challenged, since it has been argued that the correlation approach just accounts for the co-occurrence of external and internal knowledge sources, and does not test directly their complementarity in relation to innovation results (Mohnen and Röller 2005).

The second econometric strategy tries to cover this failure by adopting a direct approach and including empirical studies concerned with the study of complementarities in direct relation to the performance effects. It is characterized by hard testing for complementarities, using coefficients coming from estimations of the innovation function (testing in terms of innovation output) (Cassiman and Veugelers 2006; Hagedoorn and Wang 2012; Hou and Mohnen 2013; Morero, Ortiz, and Wyss 2014), or from a firm production function (testing in terms of productivity) (Lokshin, Belderbos, and Carree 2008; Schmiedeberg 2008; Hou and Mohnen 2013).

The results in this line tend to be more conclusive towards complementarity itself. A series of studies points out complementarity relations (Cassiman and Veugelers 2006; Álvarez, Morero, and Ortiz 2013; Hou and Mohnen 2013; Morero, Ortiz, and Wyss 2014) and others show ambivalent or contingent results (Lokshin, Belderbos, and Carree 2008; Schmiedeberg 2008; Hagedoorn and Wang 2012).

The most influential paper in the direct approach is the work of Cassiman and Veugelers (2006) that analyzes complementarity between external knowledge purchase and internal R&D activities in Belgian firms. They explicitly introduced complementarity tests (super and sub modularity tests) in relation to the innovation performance of firms using discrete data, arising from diverse estimations of the innovation function. Their results point out that these activities are complementary to innovation, and this is sensitive to contextual aspects and other aspects of the strategic environment of the firm.

However, the direction and the degree in which contextual variables affect complementarity is not well captured by their method and remains an open research path. In response to the fact that selection of innovation strategy is endogenous, which may cause biased estimates, Cassiman and Veugelers (2006) proposed a two-step procedure to construct predicted values for the innovation strategy of the adoption approach and to use them as instrument variables for the firm's innovation strategy in the innovation regression (direct approach). The adoption of each strategy is approached through a bivariate probit model, which regresses the non-exclusive innovation activities on assumed exogenous control variables. To elucidate the contextual variables that affect innovation, in pursuit of variables that can explain the joint occurrence of innovation activities – i.e. variables that affect complementarity between innovation activities – a multinomial logit model is used. This type of model is useful when one tries to explain choices of several mutually exclusive alternatives, in this case the exclusive combinations of performed internal R&D (Make) and acquired technology externally (Buy) decisions (the dependent dummy variables). Thus, there are firms that have no innovation activities; they only have their own R&D activities, external technology acquisition, and combine their own R&D activities and external technology acquisition. However, like other econometric models, the multinomial logit model has some assumptions that can often be a limitation that restricts the findings.

One of the most serious assumptions is the Independence of Irrelevant Alternatives (IIA): the relative odds between any innovation activity are independent of the number and nature of other ones being simultaneously considered. This property cannot be sustained as long as decisions on the choice between internal or external activities cannot be considered independent, and then the results might not be very realistic.

Another consideration that must be taken into account is having sufficient observations in each exclusive category for the multinomial logit estimates. As noted by Agresti and Kateri (2011), when some category occurs

Table 1: Empirical evidence on internal and external innovative activities complementarities. The correlation approach.

Results		
Substitutability	Complementarity	Ambivalent
Blonigen and Taylor (2000); Mytelka (1978); Fikkert (1994); Basant and Fikkert (1996); Pisano (1990); Love and Roper (2001)	Arora and Gambardella (1994); Braga and Willmore (1991); Deolalikar and Evenson (1989); Veugelers (1997); Schmiedeberg (2008)	Audretsch, Menkveld, and Thurik (1996); Veugelers and Cassiman (1999)

Source: Own elaboration

relatively few times, this limits the number of predictors for which effects can be estimated precisely: multinomial logit estimates may be quite biased and estimates of standard errors may be poor.

Additionally, models with several predictors often suffer from multicollinearity, i.e. correlations among predictors. Deleting such a redundant predictor can be helpful, for instance reducing standard errors of other estimated effects (Agresti and Kateri 2011). Nevertheless, correlations among predictor variables can have wealth for themselves and for the whole context of the innovative activities; even direct and indirect effects are sidestepped, and may be too valuable to be considered merely as redundant variables.

Under these considerations, although this two-step procedure performed by Cassiman and Veugelers (2006) is valuable because it considers the endogeneity of the selection of innovative activities, the accumulation of requirements of the models in each step impose limitations on results and underlying conclusions on the complementarity of innovative strategies. These limitations are recognized by the authors themselves when noting the poor predictive power of the adoption rates and the weak statistical significance of the variables, reflecting the need to search for more informative firm characteristics that explain the adoption of individual innovation activities, including good potential instruments (Cassiman and Veugelers 2006; Álvarez, Morero, and Ortiz 2013; Hou and Mohnen 2013; Morero, Ortiz, and Wyss 2014).

Association approach

An additional empirical quantitative strategy exists in the literature that can be characterized as the association approach. In relation to the last two lines, this approach is a kind of intermediate approach, in the sense of being stronger than the correlation approach in probing complementarity, but weaker to hardly test it respect to the direct approach. There are investigations that resort to combining diverse statistical techniques, such as multivariate analysis, to establish and explore complex associations between variables, qualitative in nature, but engaged through quantitative tools.

As Johnson, Wichern, and Education (2014) point out, the complexities of most phenomena require an investigator to collect observations on many different variables, and the body of statistical methodology concerned with this is multivariate analysis. The objectives of investigations for which multivariate methods may be useful include the following (Johnson, Wichern, and Education 2014):

- Data reduction or structural simplification. The phenomenon being studied is represented as simply as possible without sacrificing valuable information. This is done by building a few indicators as transformation of the original variables.
- Grouping in defined groups. Finding groups of ‘similar’ observations or variables is possible, based upon measured characteristics.
- Investigation of the dependence among variables. The nature of the relationships among variables is of interest. Are all the variables mutually independent or are one or more variables dependent on the others? If so, how?
- Prediction. Relationships between variables must be determined for predicting the values of one or more variables based on observations of the other variables.
- Hypothesis construction and testing. Specific statistical hypotheses, formulated in terms of the parameters of multivariate populations, are tested. This may be done to validate assumptions or to reinforce prior convictions.

Figure 1 shows some of the most common multivariate methods.

According to the objectives for which it is useful, indicated in the previous paragraph, the multivariate methods indicated in Figure 1 have developed in varying degrees in different disciplines, particularly within the *metrics* discipline: econometrics, psychometrics, chemometrics, environmetrics, etc. These *metrics* are but applications of statistics exclusively focused on developing theory and methods for problems in a particular discipline, i.e. the purpose of econometrics is fitting and testing economic models, rooted, in general, in economic theory. In this sense, beyond the benefits of having their own *metrics* area, further development of methods goes in the same direction as the mainstream. At this point, it becomes important to note the progress of other disciplines, mainly in the field of social sciences, not only in their development of proper statistical methods, but also in the context of their problems, how they address them, and their similarities with economic problems. For instance, factor analysis was nurtured and developed, primarily, by scientists interested in constructs, such as intelligence. The purpose of factor analysis is to describe, if possible, the covariance relationships among many variables in terms of a few underlying (latent), but unobservable, random variables called factors, pursuing the description of a set of variables observed by a small number of latent variables (Johnson, Wichern, and Education 2014).

These techniques have been occasionally applied in economics, even to the problem of complementarities. However, the results from these investigations are often disregarded in the usual empirical literature reviews (see, for example, Mohnen and Röller 2005; Cassiman and Veugelers 2006; Schmiedeberg 2008). This is certainly a bias in the discipline, which gives to these methods a lower status.

Four particular papers can be mentioned in line with the association approach that directly address the question of the use of internal and external innovative activities. Particularly Doloreux (2015) explored the link between the use of a whole series of internal and external knowledge sources with the innovation performance of Canadian wine firms. His empirical work involved a principal component analysis of 16 variables of knowledge sourcing (including various internal sources, as well as diverse market, institutional and specialized sources), and a posterior cluster analysis with innovation performance variables. The results confirm the absence of complementarity between internal and external sources, which suggests the possibility of a substitutability relation. The study reveals that, in this sector in particular (a ‘supplier dominated’ one), the commercial knowledge

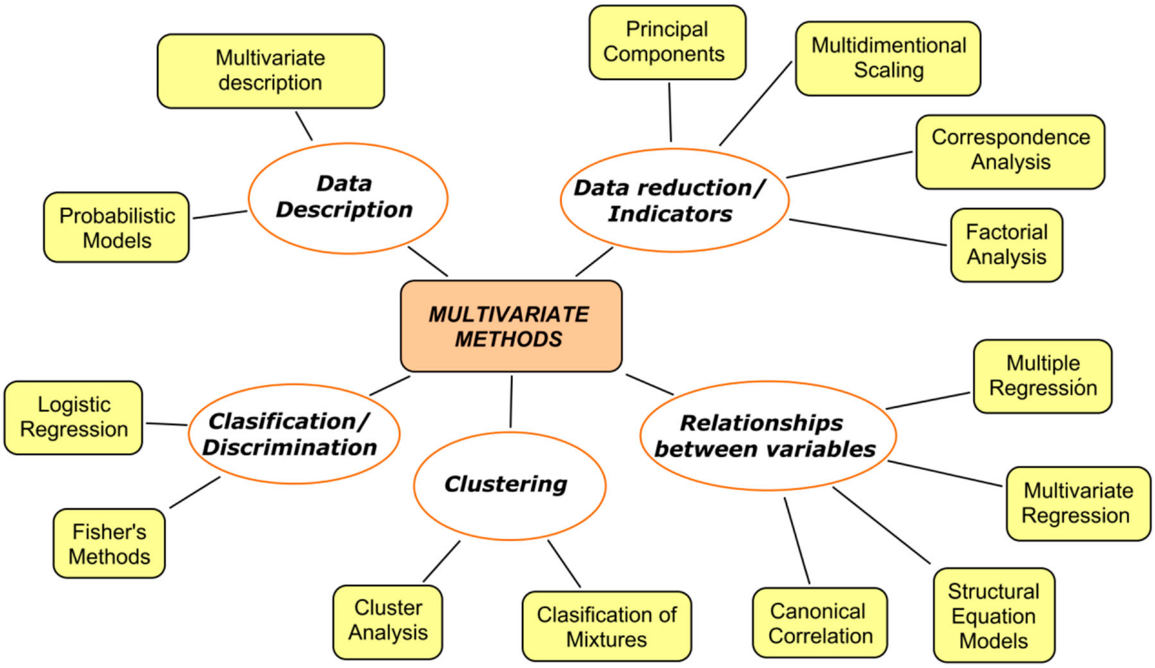


Figure 1: Usual multivariate techniques.

coming from clients and suppliers strongly influences the innovation performance of the firms. It should be remarked that his analysis enables a multiple examination of the complementarity between sources.

In the context of an emerging economy, Motta, Morero, and LLinás (2007) followed a similar path, analyzing the knowledge sourcing of automobile firms in Argentina. They performed a multiple factor analysis of internal innovative efforts, diverse external linkages and internal competences of the firm, and afterwards applied a hierarchical cluster analysis with innovation and structural supplementary variables building a typology of firms. Their analysis confirms and states the existence of complementarity between internal and external sources in relation to the innovation performance[1] of the automobile firms, and allows an appreciation of the degree of the intensity of this complementarity across groups (namely, the clusters). The work of Morero (2010; 2015) followed a similar path with similar results. Milesi (2006), also using these methods, identified six patterns of innovation in Argentinean manufacturing firms in the 1990s. Their results show that the composition of the efforts of the firms has a relation with the type of innovation: firms with efforts biased by internal technology development tend to innovate in products, whereas firms with efforts concentrated on external technology acquisition tend to innovate in process. Suarez (2015) analyzed the changes in the innovation strategies of Argentinean manufacturing firms during 1998–2006, adding a dynamic perspective. She performed a discriminant cluster analysis, and found that the firms that continuously innovated over the study period (practically a decade) presented high levels of complementary investments and capabilities.

The studies in this line have the potential to complement the direct approach, which is so concerned with a simple test to prove the existence of complementarity that it missed the whole comprehensive analytical richness of the nature of the multiple relationships complexities arising between the diverse knowledge sourcing of the organizations. Also, these kind of studies could be more fruitful for engaging the issue of studying the particular determinants of complementarity, that has been shown as a multi causal and very complex relation, which failed to be approached by hard econometric regressions. In the next section, we illustrate this point for the case of the software sector of Argentina.

However, traditional reviews of the empirical work in economics usually disregard the efforts made in this line. This means that it is not a line properly recognized in the discipline, due to the supposed superiority of econometric procedures.

An application case: The software sector in Argentina

First, a direct approach is performed to test for complementarities and seek their determinants. Second, an association approach is carried out, to compare their results and explanatory power.

The data came from 257 Argentinean software firms and it covered the period 2008–2010. These firms constitute a representative sample according to size, and the available information allowed the construction of appropriate indicators to generate pertinent models, useful for later testing the supermodularity and submodularity between knowledge sources. These data are based on a technological survey carried out over 2011 under the research project '*Capacity of Absorption and Production Systems Connectivity and Local Innovation*' from the Carolina Foundation, and provided us a cross-sectional data set to perform all the quantitative exercise.

A direct approach application

Complementarity tests in the software sector of Argentina[2]

We apply the method of Cassiman and Veugelers (2006) to test the existence of complementarity in the innovation strategies of Argentinean firms from the software sector. Testing for complementarities between two variables when the nature of the available data regarding the key variables are discrete implies testing if the objective function is supermodular in these arguments. The condition of supermodularity between two arguments implies that the function shows complementarity between these arguments, and the condition of submodularity shows substitutability[3] (Milgrom and Roberts 1990; Topkis 1998).

It can be assumed that innovation is measured like an ordinal variable, which considers the introduction of new products, new processes, improved products, significant improved processes, organizational changes, or the development of new commercial channels; each weighted according to the novelty degree of the innovation. For instance, the innovation depends on recourse to internal and external knowledge sources, represented by dummy variables and denoted by KS^i_{int} and KS^i_{ext}, respectively. Additionally, traditional structural factors (*Size, Property of Capital, Age, Specialization, Exports, Linkages and Competences*) influence innovation; these variables are incorporated into the model by Z_i vector. Thus, we carried out an Ordered Probir model:

$$I^{*i} = (1 - KS^i_{int})(1 - KS^i_{ext})\delta_{00} + KS^i_{int}(1 - KS^i_{ext})\delta_{10} + KS^i_{ext}(1 - KS^i_{int})\delta_{01} + KS^i_{int} \cdot KS^i_{ext}\delta_{11} + \mu \cdot Z^i + \varepsilon^i \quad (1)$$

where I^* is a latent variable: an index underlying the innovation ordinal responses observed, while δ's and μ are coefficients of knowledge sources and control variables.

The coefficients of knowledge are relevant to perform the test of complementarities. The way they were introduced to the model reflect the recurrence in knowledge sources: neither (*Not Internal; Not External*), only one (*Only Internal* or *Only External*), and both simultaneously (*Internal & External*). Their components are detailed in the Appendix. Table 2 shows the frequency of these indicators for the complete sample. It can be seen that the majority of the firms performed some kind of innovative effort, as just only 4% of the sample resorted neither to internal nor to external sources. Besides that, more than 22% of the firms concentrated their efforts on internal sources, up to 7% on external, and almost two-thirds of the sample combined some degree of internal and external innovation activities.

A description of control variables involved in the model can be seen in Table 6.

The Wald test for inequality restrictions to test for complementarities use the estimated coefficients for the use of knowledge sources as follows:

$$H_0: \delta_{00} + \delta_{11} - \delta_{10} - \delta_{01} \leq 0$$

$$H_1: \delta_{00} + \delta_{11} - \delta_{10} - \delta_{01} > 0$$

In addition to the Probit Model, considering alternative dependent variables, Probit and Tobit models were used as a robustness check. Table 3 shows the models estimated with control variables that showed greater significance. In terms of performance, the Ordinal Probit model shows a proportion of correct prediction of 0.52, while the Probit and Tobit models showed 0.70 and 0.70, respectively. In all cases, a positive relationship between *innovation* and *linkages and competences* is evident, which would mean that higher levels of *linkages and competences* make it more likely to get higher levels of innovation.

Table 4 shows the results of the test. When the Wald statistic is below 1.642, the null hypothesis is accepted, whereas when the statistic is above 7.094 it is rejected (Kodde and Palm 1986).

Thus, the results point to the existence of a relation between internal and external knowledge sources for innovation: the hypothesis of supermodularity test is accepted, while the submodularity test for all models considered is not. Thereby, internal and external knowledge sources are complementary, inducing, when applied jointly, a higher level of innovation. It can also be seen that with the different models the conclusions are maintained, and Wald tests are robust to the variations of the different models, indicating the robustness of the quantitative analysis.

The determinants of the complementarity in the software sector[4]

In this section we intend to determine how certain control variables affect the joint adoption of innovation activities, thus constituting drivers of complementarity: those variables that are significant for the *Int&Ext*, but not for other innovative strategies (Cassiman and Veugelers 2006). Table 5 shows the results of a multinomial logit model regression, taking as the benchmark case the *Not Internal; Not External*, so all coefficients are expressed as relative to a non-innovative strategy. The model has a good fit, showing a proportion of correct prediction of 66.29 and it was selected by the lower AIC and BIC from a series of alternate regressions.

This relation of complementarity is true for the importance of *Financial Obstacles*, which shows a positive relation with the probability of the firm combining innovative activities (*Int&Ext*), relative to not doing any innovation activity; and is also true for the perception of the *Human Resources Environment* quality, which shows a negative relation with the probability of the firm combining innovative activities.

Furthermore, as Table 5 shows, the multinomial logit model also shows that the policy instrument *Fonsoft* affects significantly and positively the probability of the use of both internal and external sources jointly, and also the probability of resorting only to internal knowledge

Table 2: Recurrence of internal and external knowledge sources. Complete sample.

	Frequency (%)
Not Internal Not External	3.89
Only Internal	6.61
Only External	22.18
Internal & External	66.93

Note: 1 case missing.

Table 3: Estimates of the models specified.

	Ordinal Probit			Probit			Tobit		
Knowledge Sources dummies	Coefficient(1)		Sign.(2)	Coefficient(1)		Sign.(2)	Coefficient(1)		Sign.(2)
(intercept)	–	–		−1.594	0.6282	**	–	–	
Not Internal Not External	–	–		–	–		–	–	
Only Internal	0.4124	(0.5554)		0.2456	(0.5931)		1.5596	(2.0494)	
Only External	0.3317	(0.4873)		0.2330	(0.5105)		1.0895	(1.7646)	
Internal & External	0.9093	(0.4764)	*	0.8233	(0.5022)		3.8383	(1.7235)	
Controls									
Size	0.0002	(0.0007)		−0.0009	(0.0008)		0.0015	(2.5045)	
Origin of Capital	−0.3529	(0.2920)		−0.2040	(0.3514)		−1.7944	(1.1346)	
Export Profile	0.0020	(0.0025)		0.0027	(0.0030)		0.0077	(9.8069)	
Specialized in Services	−0.1568	(0.2118)		−0.3059	(0.2574)		−1.5291	(8.2992)	*
Specialized in Products	−0.1622	(0.1982)		−0.2647	(0.2454)		−1.0068	(7.7073)	
Age	0.0146	(0.0103)		0.0078	(0.0125)		0.0468	(3.9297)	
Linkages	0.2525	(0.0972)	***	0.3123	(0.1211)	***	1.0767	(3.7906)	***
Competences	0.4044	(0.1949)	**	0.4685	(0.2305)	**	2.6442	(7.5831)	***
/cut 1	1.6316	(0.5824)	***						
/cut 2	3.109	(0.5996)	***						
Log-likelihood	−227.85								
R^2	–			–			–		
R^2 adj	–			–			–		
AIC	481.71			–			–		
Prob > chi^2	–			0.0001			0.0000		
Corr (obsv'd and pred'd values)	–			–			0.4939		
Perc. of Correct Predictions	0.5247			0.7037			–		

(1) Standard error in parentheses.
(2) *** Significant at 1%; ** Significant at 5%; * Significant at 10%.
Note: The estimated coefficients of the models are expressed as the deviations of the coefficient of Not Internal and Not External knowledge sources recurrence.

Table 4: Complementarity and substitutability tests. Wald statistics.

	Ordinal Probit	Probit	Tobit
Supermodularity Test	7.79E-20	8.81E-26	1.12E-24
Submodularity Test	2.04163	1.918098	2.613098

Note: The test is accepted if the Wald statistic is below the lower bound at 10% of significance (1.642), and it is rejected if the statistic is above the upper bound (7.094) (Kodde and Palm 1986).

Table 5: Multinomial logit model.

	Knowledge Sources Recurrence		
	IntAndExt	OnlyExt	OnlyInt
Financial obstacles	0.8246* (0.4766)	0.6168 (0.4792)	0.7857 (0.5494)
Internal knowledge-skills obstacles	−1.0306 (0.7428)	−0.7510 (0.7458)	−1.2716 (0.7837)
Uncertainty of the innovation demand	−0.8070 (0.6867)	−0.8392 (0.6893)	−1.0657 (0.7340)
Human Resources Environment	−0.8984** (0.4505)	−0.6097 (0.4514)	−0.8181 (0.5090)
Size	0.0184 (0.0269)	0.0196 (0.0269)	0.0145 (0.0277)
Specialized in Products	−0.1629 (1.4447)	−0.3434 (1.4544)	0.8277 (1.6005)
Diversified	−2.1230 (1.3494)	−2.2762* (1.3498)	−2.8551 (1.7793)
Work Organization	0.6946 (0.7314)	0.1483 (0.7351)	0.2580 (0.8549)
Competences	1.6149 (1.3281)	0.4213 (1.3309)	0.9567 (1.5782)
Policy Fontar	13.7715 (1422.5)	13.6051 (1422.5)	14.3454 (1422.5)
Policy Fonsoft	2.3258* (1.2878)	2.0576 (1.2925)	3.0404** (1.4530)
_cons	4.8599 (5.3896)	6.7069 (5.4258)	5.6537 (5.8345)
Pseudo R^2	0.1661		
Prob > chi^2	0.0101		
AIC	346.79		
BIC	461.34		
Perc. of Correct Predictions	0.6629		
N	178		

(1) Standard error in parentheses.
(2) *** Significant at 1%; ** Significant at 5%; * Significant at 10%.
Note: Specialized in Services omitted to avoid collinearity problems.

sources. Finally, the *diversified* firms show a negative relation with the probability of resorting exclusively to external knowledge sources, relative to sources *specialized in services* firms.

However, these results show the same limitations as those in Cassiman and Veugelers (2006) that we have already mentioned. First, the results are very few: the significant positive sign of financial obstacles and the significant negative sign of human resources environment. Second and more importantly, those variables that are significant, *i.e. Human Resources Environment*, do not have a clear economic interpretation. These results, although significant, make little economic and social sense. Third, additional robustness checks and complementary estimations that can be done, only go in the direction of restricting more and more the very few results that we have. In that direction, we can mention the path pointed out by Cassiman and Veugelers (2006). It implies that the econometrical exercise should be completed with Bivariate Probit estimations of the Internal and External recourse. This finalization of this process would result in the construction of predicted innovation strategy decisions (from multinomial logit) or predicted innovation activities (from bivariate probit), and in the use of these as instruments in the innovation performance regression. However, in line with what was mentioned in the previous section regarding limitations of this econometric strategy, it is unlikely to improve the performance of the innovation function estimation (direct model), but it may, surely, further distort the results and conclusions about complementarities. Econometrics has its limits.

An association approach application: Complementarities in the software sector of Argentina

Method and variables

The variables involved in our analysis are qualitative in nature. Neither innovation performance, the organizational and work structure of the firms, their R&D structure, nor innovative efforts could be exclusively reduced in continuous terms without losing a significant degree of explanatory richness.

All the above are qualitative variables, and for their quantitative analysis, it is convenient to apply techniques from multivariate data analysis. Multivariate analysis is a set of methods that can be powerful when the problem that arises implies multiple dependent or interdependent variables. One particular multivariate analysis technique is multiple factor analysis. It is a data reduction technique that allows us to summarize a large number of heterogeneous variables (called the active variables) in a new space, drawing (in fact, projecting) the observations (in this case, the firms) into a new set of variables called the factors. The factors are new variables that maximize the variability of the active variables selected. That is, the factors are a smaller number of variables, more manageable than the original variables, but are also homogeneous and continuous variables composed of a combination of the originals. In the case of a series of ordinal variables or categorical multidimensional variables, the factors are combinations of the categories of all the active variables involved in the analysis. The technique constructs factors until all the variability (also called the inertia) of the active variables is summarized by the factors, which implies that all the factors together explain the same information that the original variables.

Therefore, to characterize the recourse to internal and external innovation activities in our sample, we will apply a factor analysis as a way to reduce dimensions between diverse categories of qualitative variables. This analysis will enable us to use a new set of variables for each firm (the factors) that summarizes the knowledge sources recourse in a homogeneous way to compare all the cases. In this paper, we will apply the multiple factor analysis to reduce dimensions of two ordinal indicators: internal innovation activities and external innovation activities, which are the active variables of the factor analysis. Thus, a set of factors will be generated in terms of the active variables (in fact, in terms of all their categories), and we will project the cases (the firms) onto those new dimensions. This projection of each observation in new 'homogeneous' dimensions (made up of heterogeneous qualitative variables) allows us to calculate distances between the cases, specifically in these terms.

In short, the procedure for carrying out a multiple factor analysis can be summed up by the next points:

i. Considering the simplest case, with two categorical variables, the dataset can be represented as a table of contingency for dimension I × J, where I and J represent the categories of each variable, like binary variables, and arrange it as a matrix X.[5]
ii. The table of relative frequencies, F, is calculated from the matrix X, representing the joint relative frequencies as f_{ij}, and the marginal relative frequencies as $f_{i.}$ And $f_{.j}$.
iii. The standardized table Z is calculated from F, whose elements are: $z_{ij} = \dfrac{f_{ij}}{\sqrt{f_{i.}f_{.j}}}$.
iv. The dimension reduction and the consequent determination of factors requires a projection that keeps the relative distances of original data, maintaining its structure and promoting the wanted factors to be uncorrelated, which implies maximizing the variability of the projected points. That is, max $a'Z'Za$ subject to the vector a has norm 1: $a'a = 1$.

In this paper, we apply the multiple factor analysis to reduce dimensions of two ordinal indicators: internal innovation activities and external innovation activities, which are the active variables of factor analysis. Thus, a set of factors will be generated in terms of the active variables (in fact, in terms of all their categories), and we will project the cases (the firms) onto those new dimensions. This projection of each observation in new 'homogeneous' dimensions (made up of heterogeneous qualitative variables) allows us to calculate distances between the cases, specifically in these terms (i.e. how far/close is one firm from another related to their internal/external knowledge sourcing).

The goal of this analysis is to build relatively homogeneous groups of firms, related to the relative importance of the external and internal components of their innovation activities, and then to evaluate the associated innovative performance, in homogeneous dimensions, to analyze if there is an association suggestive of complementarities. The factors calculation could be used to perform a cluster hierarchical analysis, clustering the firms in such a way that the cases in the same group are more similar (in relation to the factors) to each other than to those in other groups.

Moreover, it is also possible to project in the new dimensions not only the observations, but also their other characteristics not involved in the factor construction, as supplementary variables. Thereby, the groups could also be characterized by the level of the other variables, involved in neither the factor analysis nor in the cluster analysis, through a proportion 'Valeur Test' (Morineau 1984) for each category of all the variables. In that sense, not only will we be particularly be interested in the variables of innovation, to characterize the innovation performance of the groups and to evaluate innovation complementarities, but also to see the association of other variables, such as structural and organizational, with the recourse to diverse policy instruments and obstacles to innovation.

In sum, in applying cluster techniques to the factor analysis, we will elaborate typologies of firms according to internal/external balance of innovation activities, and we will constitute homogeneous groups of firms in these terms. Also, the innovation indicators will be projected as supplementary variables, as well as structural, policy and obstacles to innovation variables.

From the previous section, the variables of innovation activities and the variables in Lickert scales (as obstacles to innovation) were converted into ordinal terms as Appendix details. The other binary illustrative variables were used as such. Table 6 presents the sample levels of all the variables used to perform the multiple factor analysis. Regarding our key variables, the software firms showed a wide variety of internal and external knowledge sources mix. Around 48% of the sample showed a high level of internal efforts and a similar percentage of them had a low indicator of external sourcing. However, one in five firms made low external efforts and almost 23% of firms in the sample showed high levels of external innovation activities. Innovation performance in the sample was dominated by medium levels, which represent around half of the sample, and more of 19% of the firm presented high innovation levels. We will try to analyze which knowledge combination are assisted to groups dominated by these firms, whose with better innovation performance; but also, the other extreme, with the less innovator firms.

Results

The factor analysis reduced dimensions of two variables (Internal and External Innovation Activities) with three categories each (Low, Medium, High), which are the active variables. The analysis gave four factors which collect 100% of the inertia. When we use the four factors, it implies using the same information as the active original variables, as they accumulate their whole variability. These new dimensions allow us to calculate distances between the firms in these terms, and to apply cluster techniques to build relatively homogeneous groups of firms on the main features of their recourse to innovation activities. Then, additional variables could be projected into these new dimensions to characterize the groups. The detail of inertia accumulation of the factors as the results of the hierarchical cluster analysis and the cut considered, which present the best fit, are presented in Box 1. There it can be seen that four factors accumulate all the inertia (the variability) of the active variables. That means that those factors summarize the same information of the original variables, but in new homogeneous dimensions. That allows us to apply cluster techniques and to calculate 'proximities' between the cases (firms) in terms of their internal and external knowledge sourcing, and to group the cases as the hierarchical cluster analysis tree shows. The tree presented shows various possible

Table 6: MFA variables. Sample levels.

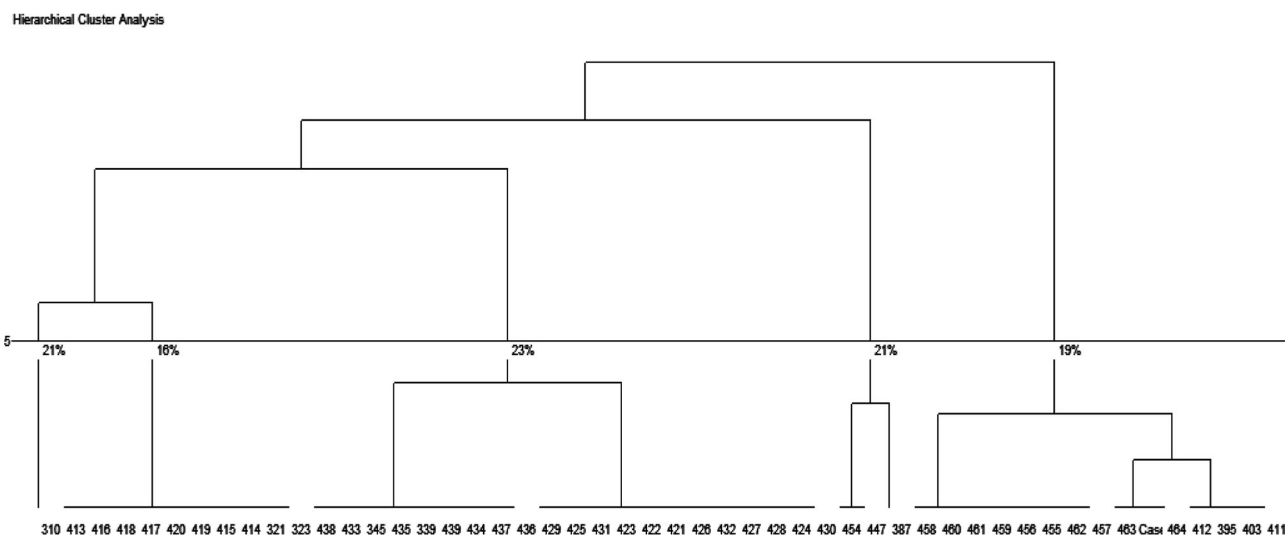
	All Sample					
	Low	Medium	High	Yes	No	NA
Innovation Indicator	30.85%	49.46%	19.68%			–
Internal Innovation Activities	19.07%	31.91%	48.64%			0.39%
External Innovation Activities	48.25%	28.40%	22.96%			0.39%
Competences	29.18%	25.29%	21.79%			23.74%
Work Organzation	13.23%	32.30%	28.02%			26.46%
Financial Obstacles	23.35%	24.51%	50.97%			1.17%
Internal knowledge-skills Obstacles	12.84%	21.01%	64.59%			1.56%
Uncertainty of the demand Obstacles	39.69%	32.68%	26.07%			1.56%
Appropiability Obstacles	61.09%	21.79%	15.56%			1.56%
Age PreConvertibility				8.56%	91.05%	0.39%
Age Convertibility				35.41%	64.20%	0.39%
Age Post Convertibility				54.47%	45.14%	0.39%
Specialized in Products				47.86%	50.58%	1.56%
Specialized in Services				28.02%	70.43%	1.56%
Not Specialized – Diversified				22.57%	75.88%	1.56%
Policy Fonsoft				37.35%	38.52%	24.12%
Policy Soft Law				23.74%	79.26%	–
Policy Fontar				17.51%	58.75%	23.74%
	Small	**Medium**	**Big**	**National**	**Foreign**	
Size	68.48%	17.51%	13.23%			0.78%
Property of Capital				91.05%	8.17%	0.78%

grouping cuts. We selected the more illustrative one and with the best fit, but the results presented stand in the other cuts, and can be provided by the authors upon request.

indicator of innovation, but another performance output can serve as well (for example, sales growth, productivity growth, exports performance, etc.).

Box 1. Multiple factor analysis and cluster analysis details.

Number	Eigenvalue	%	% Accumulated
1	0,5895	29,47	29,47
2	0,5006	25,03	54,50
3	0,4994	24,97	79,41
4	0,4105	20,53	100,00



Hence, we proceed to the analysis of the clusterization. Table 7 presents the over- and underrepresented categories of the indicators of the cluster analysis. Redundant variables were omitted. The cut with the best fit gave us a classification in five groups/clusters, which could be characterized by their recourse to internal and external innovation activities. Cluster I is the group of firms with outstanding internal and external innovation efforts, and comprises 21% of the sample. At the opposite extreme is Cluster V, which represents 19% of the firms of the sample and is characterized by low internal and external innovation activities. Clusters II to IV present intermediate behavior. Cluster II (23% of the sample) shows medium efforts in external activities, but medium/high internal efforts. Cluster III (21% of the firms of the sample) presents low external activities together with high in-house efforts, and Cluster IV is also characterized by external efforts, but with a medium level of internal activities, representing the 16% of the software firms.

Table 8 highlights the significant modalities of our key variables to analyze complementarities. We are looking for comparable results from the previous ones from a direct approach analysis, focused on the internal/external complementarities in relation to innovative performance. That is the reason we specially look for the illustrative

Our first result is that complementarity holds. Focused on the extreme groups, we can see that Cluster I, with high levels both of internal and external activities, also has a significant overrepresentation of the firms with a high level of innovation output. On the contrary, Cluster V presents an overrepresentation of the firms with a low level of innovation performance, which is the group with low levels of innovation activities too.

Moreover, we can see that other variables affect the complementarity status or, at least, show an association with these groups that evidences the complementarity relations. We should see which variables appear significantly in both extremes, but with 'opposite' modalities (i.e.: 'low' vs. 'high', 'yes' vs. 'no', etc.). Size, specialization and policy variables are contextual variables that appear significantly associated with innovation efforts complementarity. Larger and specialized product software firms are associated with the best overrepresented innovative performance. At the other extreme, small firms, born in the convertibility period, and diversified or specialized in services, characterize the cluster with low internal and external innovative activities and an even worse relative innovation performance. The results also show that policy variables are crucial in the extremes (as well as non-significant at all in the intermediate groups): Cluster

Table 7: Cluster analysis: over and underrepresented variables.

Cluster I	Variable	Category	% of category in group	% of category in set	Sign.([1])
Active Variables	External Innovation Activities	High	100.00	22.96	***
	Internal Innovation Activities	High	68.52	48.64	***
	Innovation	**High**	**40.74**	**19.07**	***
Structural and organizational variables	Size	Big	24.07	13.23	**
	Specialized in Products	Yes	57.41	47.86	*
Policy variables	Policy Soft Law	Yes	35.19	23.74	**
	Policy Fontar	Yes	25.93	17.51	*
	Underrepresented Categories				
Structural and organizational variables	Size	Small	55.56	68.48	**
Obstacles to innovation	Financial Obstacles	Low	14.81	23.35	*
Cluster II	**Variable**	**Category**	**% of category in group**	**% of category in set**	**Sign.** ([1])
Active Variables	External Innovation Activities	Medium	100.00	28.40	***
	Internal Innovation Activities	Medium	41.38	31.91	*
	Internal Innovation Activities	High	58.62	48.64	*
	Innovation	**Medium**	**56.90**	**46.30**	**
Structural and organizational variables	Property of capital	Foreign	17.24	8.17	***
	Age Preconvertibility	Yes	13.79	8.56	*
	Size	Medium	27.59	17.51	**
	Underrepresented Categories				
Obstacles to innovation	Appropiability Obstacles	Low	51.72	61.09	*
Structural and organizational variables	Size	Small	58.62	68.48	**
	Property of capital	National	82.76	91.05	**
Cluster III	**Variable**	**Category**	**% of category in group**	**% of category in set**	**Sign.**([1])
Active Variables	External Innovation Activities	Low	100.00	48.25	***
	Internal Innovation Activities	High	100.00	48.64	***
	Innovation	***Not significantly different from sample levels***			
Structural and organizational variables	Age Post Convertibility	Yes	70.37	54.47	***
	Age Convertibility	No	77.78	64.20	**
	Specialized in Products	Yes	57.41	47.86	*
	Work Organization	High	37.04	28.02	*
Obstacles to innovation	Appropiability Obstacles	Low	72.22	61.09	**
	Underrepresented Categories				
	none variable relevant				
Cluster IV	**Variable**	**Category**	**% of category in group**	**% of category in set**	**Sign.**([1])
Active Variables	Internal Innovation Activities	Medium	100.00	31.91	***
	External Innovation Activities	Low	100.00	48.25	***
	Innovation	***Not significatly different from sample levels***			
Structural and organizational variables	Age Post Convertibility	No	56.10	45.14	*
	Not Specialized – Diversified	Yes	31.71	22.57	*
	Underrepresented Categories				
Obstacles to innovation	Appropiability Obstacles	High	7.32	15.56	*
Structural and organizational variables	Size	Big	4.88	13.23	*
Cluster V	**Variable**	**Category**	**% of category in group**	**% of category in set**	**Sign.**([1])
Active Variables	Internal Innovation Activities	Low	98.00	19.07	***
	External Innovation Activities	Low	58.00	48.25	*

(*Continued*)

Table 7: Continued.

Cluster I	Variable	Category	% of category in group	% of category in set	Sign.([1])
	Innovation	**Low**	**48.00**	**33.46**	**
Structural and organizational variables	Size	Small	82.00	68.48	**
	Age Convertibility	Yes	46.00	35.41	*
	Specialized in Products	No	72.00	50.58	***
	Not Specialized – Diversified	Yes	32.00	22.57	*
	Specialized in Services	Yes	40.00	28.02	**
	Work Organization	Low	22.00	13.23	**
	Competences	Low	38.00	29.18	*
Policy variables	Policy Fonsoft	No	50.00	38.52	**
	Policy Soft Law	No	86.00	76.26	**
	Policy Fontar	No	72.00	58.75	**
	Underrepresented Categories				
	Innovation	**High**	**10.00**	**19.07**	**
Structural and organizational variables	Work Organization	High	16.00	28.02	**

Table 8: Summary of significant key variables – cluster analysis([1]).

Key Variables	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
Internal Innovation Activities	High***	Medium*/High*	High***	Low***	Low***
External Innovation Activities	High***	Medium***	Low***	Medium***	Low*
Innovation	High***	Medium**	–	–	Low**

([1]) *** Significative at 1%; ** Significative at 5%; * Significative at 10%

I, where high efforts, both internal and external, led to high innovation output, has overrepresented firms that resort to policy instruments, as Fontar and the benefits of the Software Law. The opposite is overwhelming in Cluster V, where the firms that not only do not resort to these policy benefits, but also to Fonsoft, are significantly overrepresented.

Other weaknesses characterize the less innovative group, not necessarily associated with complementarity. Firms with low levels of competences and less agile work organization structures are overrepresented in Cluster V. Obstacles in general seem not to have a relevant role affecting complementarities. The low importance of financial obstacles in Cluster I show a one-way importance. This helps to understand why this variable was relevant in the previous Multilogit analysis, but their sign is disturbed by the fact that it does not operate in the same direction for less innovative firms. Age is another variable that does not have a clear role: except for those firms from the convertibility period, characterized by the neoliberal reforms in Argentina, which are overrepresented in the poor group, while no period presents overrepresentation in the more innovative Cluster. That implies the presence of firms from the three periods.

Third, we can see that the diverse mix between internal/external activities from intermediate groups allows appreciation of the sensitivities in relation to complementarities in the middle.

Cluster II is characterized by medium firms, with foreign overrepresentation. It is a group with medium to high internal innovative activities combined with medium external efforts, as well as a medium innovation performance. It reinforces the importance of size affecting complementarity. Appropriability obstacles for innovation could mitigate, in this case, the innovation output, but not the internal efforts. Cluster III is characterized by new firms and those specialized in products, with high work organization, but without an effective combination of internal (high) with external (low) efforts. This led to a worse innovative performance in relation to Cluster I. In addition, appropriability obstacles for innovation appear relevant in this case, but do not affect the recourse to in-house activities. Cluster IV is characterized by predominantly diversified SMEs that have started to make some internal efforts. They have an average innovative performance that seems to not be affected by appropriability obstacles.

It must be highlighted that these last two points (the other variables significantly associated to complementarity and the intermediate internal/external mixing associations) remained mainly absent in the econometric regressions and testing.

Final remarks

The paper compares two different quantitative strategies to address a particular relevant issue in the industrial economics literature. In particular, we address the question of internal/external innovation complementarities with the more modern complementarity econometrical tests, super- and submodularity tests, arising from diverse firm-innovation function estimations; and we compare this with addressing the issue with multiple factor analysis and cluster techniques.

The results show that not only can multivariate analysis techniques obtain the same results as hard econometrical tools, but also that multiple factor analysis and cluster techniques allow for better exploiting of the richness of qualitative data. In fact, the econometrical results seem rather trivial in comparison to the multivariate analysis results, which allow us both to establish the existence of complementarity relations and to appreciate the intensity of these relations and their association to other organizational and structural aspects of the firms, as well as their association with policy promoting instruments.

In relation to the theoretical literature, firstly it must be pointed out that every empirical strategy adopted in this paper stands for internal and external complementarity for innovation, against the orthodox view of transaction cost and property right theories.

Secondly, our empirical comparison shows that where traditional hard econometrical testing failed (leaving a gap in the literature), multivariate analysis has succeeded. The question of the determinants of innovation complementarities was properly addressed by the association approach. Size, the specialization of the firm and policy promotion instruments affect the relevance of the complementation between internal and external efforts to reach innovation results. It is not trivial because it means that there is room for industrial policy to boost the innovation possibilities of firms, which is especially critical for emerging economies. It should be considered that the impact of every policy will differentially affect large, small and medium firms, as well as diversified firms and firms specialized in services or products.

These last aspects that remained obscured by the econometric analysis are enlightened by the multivariate analysis performed. This highlights the relevance of a wider acceptance of multivariate techniques in economics. In the same way as it is accepted in other social sciences, economics must learn that there is no universal method and that we have a need for multi methodological approaches. The case presented in this paper reinforces the claim for plurality in economics, not just for a greater tolerance and democratization of the discipline, but also for the progress of science.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. In fact, this complementarity was later confirmed through supermodularity tests in Álvarez, Morero, and Ortiz (2013).
2. This section relies partially on Morero, Ortiz, and Wyss (2014).
3. A detailed explanation of the mathematical and econometrical issues regarding the methodology can be seen in Morero, Ortiz, and Wyss (2014) and Morero, Ortiz, and Motta (2015).
4. This section relies partially on Morero, Ortiz, and Motta (2015).
5. For more than two qualitative variables, it can be seen as a contingency table where categories of all the variables are placed in columns and observations, in rows.

References

Agresti, Alan, and Maria Kateri. 2011. *Categorical Data Analysis*. Berlin: Springer.

Álvarez, I., H. A. Morero, and P. Ortiz. 2013. "Knowledge Sources Complementarities in Argentina's Production Networks." XVIII reunión Anual Red Pymes Mercosur, Resistencia, Argentina.

Arora, Ashish, and Alfonso Gambardella. 1990. "Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology." *The Journal of Industrial Economics* 38 (4): 361–379.

Arora, Ashish, and Alfonso Gambardella. 1994. "Evaluating Technological Information and Utilizing it: Scientific Knowledge, Technological Capability, and External Linkages in Biotechnology." *Journal of Economic Behavior & Organization* 24 (1): 91–114.

Arrow, K. 1962. "Economic Welfare and the Allocation of Resources for Invention." In *The Rate and the Direction of Inventive Activity*, edited by R. Nelson, 609–626. Nueva Jersey, USA: Princenton University Press.

Audretsch, D. B., A. J. Menkveld, and A. R. Thurik. 1996. "The Decision Between Internal and External R & D." *Journal of Institutional and Theoretical Economics* 152 (3): 519–530.

Backhouse, Roger E. 1998. "If Mathematics is Informal, Then Perhaps We Should Accept That Economics Must be Informal Too." *The Economic Journal* 108 (451): 1848–1858.

Basant, Rakesh, and Brian Fikkert. 1996. "The Effects of R&D, Foreign Technology Purchase, and Domestic and International Spillovers on Productivity in Indian Firms." *Review of Economics and Statistics* 78 (2): 187–199.

Blonigen, Bruce A., and Christopher T. Taylor. 2000. "R&D Intensity and Acquisitions in High-Technology Industries: Evidence From the US Electronic and Electrical Equipment Industries." *The Journal of Industrial Economics* 48 (1): 47–70.

Braga, Helson, and Larry Willmore. 1991. "Technological Imports and Technological Effort: An Analysis of Their Determinants in Brazilian Firms." *The Journal of Industrial Economics* 39 (4): 421–432.

Cassiman, Bruno, and Reinhilde Veugelers. 2006. "In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition." *Management Science* 52 (1): 68–82.

Coase, R. H. 1937. "The Nature of the Firm." *Economica* 4 (16): 386–405.

Cohen, W., and D. Levinthal. 1989. "Innovation and Learning: The Two Faces of R&D." *The Economic Journal* 99 (397): 569–596.

Colander, David, and Harry Landreth. 2008. "Pluralism, Formalism and American Economics." In *Pluralist Economics*, edited by Edward Fullbrook, 26–43. London: Zed Books.

Deolalikar, Anil B., and Robert E. Evenson. 1989. "Technology Production and Technology Purchase in Indian Industry: An Econometric Analysis." *The Review of Economics and Statistics* 71 (4): 687–692.

Doloreux, David. 2015. "Use of Internal and External Sources of Knowledge and Innovation in the Canadian Wine Industry." *Canadian Journal of Administrative Sciences/ Revue Canadienne des Sciences de L'Administration* 32 (2): 102–112.

Downward, Paul, and Andrew Mearman. 2002. "Critical Realism and Econometrics: Constructive Dialogue with Post Keynesian Economics." *Metroeconomica* 53 (4): 391–415.

Downward, Paul, and Andrew Mearman. 2009. "Reorienting Economics Through Triangulation of Methods." In *Ontology and Economics: Tony Lawson and his Critics*, edited by Edward Fullbrook, 130–141. New York: Routledge.

Fikkert, Brian Todd. 1994. *An Open or Closed Technology Policy?: India's Regulation of Technology Licenses, Foreign Direct Investment, and Intellectual Property*. New Haven: Yale University.

Foray, Dominique, and B. Lundvall. 1998. "The Knowledge-Based Economy: From the Economics of Knowledge to the Learning Economy." In *The Economic Impact of Knowledge*, edited by Dale Neef, G. Anthony Siesfeld and Jacquelyn Cefola, 115–121. Boston: Butterworth-Heinemann.

Freeman, C., and L. Soete. 1997. *The Economics of Industrial Innovation*. 3rd ed. London: Pinter.

Fullbrook, Edward. 2003. *The Crisis in Economics: the Post-Autistic Economics Movement: The First 600 Days*. Vol. 22: Cambridge, UK: Psychology Press.

Grossman, S. J., and O. D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *The Journal of Political Economy*, 94: 691–719.

Hagedoorn, John, and Ning Wang. 2012. "Is There Complementarity or Substitutability Between Internal and External R&D Strategies?" *Research Policy* 41 (6): 1072–1083.

Hou, Jun, and Pierre Mohnen. 2013. "Complementarity Between In-House R&D and Technology Purchasing: Evidence From Chinese Manufacturing Firms." *Oxford Development Studies* 41 (3): 343–371. doi:10.1080/13600818.2013.807910.

Jensen, M. B., B. Johnson, E. Lorenz, and BÅ Lundvall. 2007. "Forms of Knowledge and Modes of Innovation." *Research Policy* 36 (5): 680–693. doi:10.1016/j.respol.2007.01.006.

Johnson, Richard Arnold, Dean W. Wichern, and Pearson Education. 2014. *Applied Multivariate Statistical Analysis*. New Jersey, USA: Pearson.

Kodde, D. A., and F. C. Palm. 1986. "Wald Criteria for Jointly Testing Equality and Inequality Restrictions." *Econometrica: Journal of the Econometric Society*, 54: 1243–1248.

Lawson, Tony. 1997. *Economics and Reality*. New York: Routledge.

Lokshin, B., R. Belderbos, and M. Carree. 2008. "The Productivity Effects of Internal and External R&D: Evidence From a Dynamic Panel Data Model." *Oxford Bulletin of Economics and Statistics* 70 (3): 399–413.

Love, James H., and Stephen Roper. 2001. "Location and Network Effects on Innovation Success: Evidence for UK, German and Irish Manufacturing Plants." *Research Policy* 30 (4): 643–661. doi:10.1016/S0048-7333(00)00098-6.

Lundvall, BÄ, and B. Johnson. 1994. "The Learning Economy." *Journal of Industry Studies* 1 (2): 23–42.

Milesi, D. 2006. "Patrones de innovación en la industria manufacturera argentina." In *Buenos Aires*, 1–80. Argentina: LITTEC, UNGS.

Milgrom, Paul, and John Roberts. 1990. "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economic Review* 80 (3): 511–528.

Mohnen, Pierre, and Lars-Hendrik Röller. 2005. "Complementarities in Innovation Policy." *European Economic Review* 49 (6): 1431–1450. doi:10.1016/j.euroecorev.2003.12.003.

Morero, H. A. 2010. "Internacionalización, Tramas Productivas y Sistema Nacional de Innovación." *Journal of Technology Management & Innovation* 5 (3): 142–161.

Morero, H. A. 2015. "Domestic and Foreign Knowledge Sources for Innovation in Internationalized Production Networks: The Automotive and the Iron and Steel Cases." *Revista Brasileira de Inovação* 14 (1): 193–216.

Morero, H. A., P. Ortiz, and J. Motta. 2015. "The Determinants of Innovation Complementarities in the Software Sector. Evidence From Argentina." 13th Globelics International Conference, La Habana, Cuba, de Septiembre de 23–25.

Morero, H. A., P. Ortiz, and F. Wyss. 2014. "Make or Buy to Innovate in the Software Sector." *Pymes, Innovación y Desarrollo* 2 (3): 79–99.

Morineau, Alain. 1984. "Note sur la caractérisation statistique d'une classe et les valeurs-tests." *Bulletin Technique Centre Statistique Informatique Appliquées* 2 (1-2): 20–27.

Motta, J., H. A. Morero, and I. LLinás. 2007. "Procesos de aprendizaje y de acumulación de conocimiento en las empresas autopartistas argentinas." XII red PyMes MERCOSUR, Campinas, Brazil.

Mytelka, Lynn Krieger. 1978. "Licensing and Technology Dependence in the Andean Group." *World Development* 6 (4): 447–459. doi:10.1016/0305-750X(78)90094-3.

Pisano, Gary P. 1990. "The R&D Boundaries of the Firm: An Empirical Analysis." *Administrative Science Quarterly*, 35: 153–176.

Schmiedeberg, Claudia. 2008. "Complementarities of Innovation Activities: An Empirical Analysis of the German Manufacturing Sector." *Research Policy* 37 (9): 1492–1503.

Suarez, D. 2015. "Innovative Strategies: When Path Dependence Turns Into Path Creation. Innovation and Performance in the Argentinean Manufacturing Sector." *globelics Working Paper Series No. 2015-04, ISBN: 978-87-92923-09-7*. http://www.globelics.org/wp-content/uploads/2015/08/GWP-2015-04.pdf

Teece, David J. 1986. "Profiting From Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy." *Research Policy* 15 (6): 285–305.

Topkis, D. M. 1998. *Supermodularity and Complementarity*. New Jersey, USA: Princeton University Press.

Veugelers, Reinhilde. 1997. "Internal R&D Expenditures and External Technology Sourcing." *Research Policy* 26 (3): 303–315. doi:10.1016/s0048-7333(97)00019-x.

Veugelers, Reinhilde, and Bruno Cassiman. 1999. "Make and Buy in Innovation Strategies: Evidence From Belgian Manufacturing Firms." *Research Policy* 28 (1): 63–80.

Williamson, O. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Gabler.

Appendix: Variables

Control variables

Specialization_Products

0.**No.** Less than sixty per cent of turnover comes from product sales.

1.**Yes.** Sixty per cent or more of turnover comes from product sales

Specialization_Services

0.**No.** Less than sixty per cent of turnover comes from services sales.

1.**Yes.** Sixty per cent or more of turnover comes from services sales

Specialization_Diversified

0.**No.** Other cases distinct of category Yes (1).

1.**Yes.** The firm's turnover is composed between 40–60% of sales of products and between 40–60% of sales of services.

Property_of_Capital.

0.**National.** More than a half per cent of the capital property is national.

1.**Foreign.** Less than a half per cent of the capital property is national.

Exports_Continuous. Percentage of turnover exported in 2010.

Linkages_Ordinal.
1. **Low.** If the firm has only one or no kind of interactions related to the objectives listed in the category high below.
2. **Medium.** If the firm has only two kinds of interactions related to the objectives listed in the category high below.
3. **High.** If the firm interacts with other actors at least in three of these objectives: technical assistance, R&D collaboration, joint commercial initiatives or quality assurance collaboration.

Size_Continuous. Quantity of workers in 2010.
Size_Ordinal.
1. **Small.** Firm under 30 employees.
2. **Medium.** Firm between 30 and 100 employees.
3. **Big.** Firm with more than 100 employees.

Age_Continuous. Years to 2011 from foundation of the firm.
Age_PreConvertibility.
0.**No.** Other cases distinct of category Yes (1).
1.**Yes.** The firm was founded before 1991.

Age_Convertibility.
0.**No.** Other cases distinct of category Yes (1).
1.**Yes.** The firm was founded in 1991–2001.

Age_PostConvertibiliy.
0.**No.** Other cases distinct of category Yes (1).
1.**Yes.** The firm was founded after 2001.

Work_Organization_Ordinal.
1. **Low.** No use of agile methods.
2. **Medium.** Uses agile methods sometimes.
3. **High.** Uses agile methods permanently.

Policy_Fontar.
0.**No.**
1.**Yes.** The firm received a FONTAR aid.

Policy_Fonsoft.
0.**No.**
1.**Yes.** The firm received a FONSOFT aid.

Policy_Software_Law.
0.**No.**
1.**Yes.** The firm receives the Law Software benefits.

Financial_Obstacles_Ordinal. Regarding the relevance of the obstacle for innovation on a scale from values 1, 2, 3, 4 and 5.
1. **Low.** Values 1 or 2.
2. **Medium.** Value 3.
3. **High.** Values 4 or 5.

Internal_Knowledge_Skills_Obstacles_Ordinal. Regarding the relevance of the obstacle for innovation on a scale from values 1, 2, 3, 4 and 5.
1. **Low.** Values 1 or 2.
2. **Medium.** Value 3.
3. **High.** Values 4 or 5.

Appropiability_Obstacles_Ordinal. Regarding the relevance of the obstacle for innovation (the copy of a novelty) on a scale from values 1, 2, 3, 4 and 5.
1. **Low.** Values 1 or 2.
2. **Medium.** Value 3.
3. **High.** Values 4 or 5.

Uncertainty_Innovation_Demand_Obstacle_Ordinal. Regarding the relevance of the obstacle for innovation on a scale from values 1, 2, 3, 4 and 5.
1. **Low.** Values 1 or 2.
2. **Medium.** Value 3.
3. **High.** Values 4 or 5.

Human_Resources_Environment. Likert variable (1 to 5), regarding the perception of the quality of human resources.

Competences_Continuous. Ponders 0,2 the following five ordinal variables: Quality standards, R&D structure, Worker's qualification, Quality management and training structure. These variables have three categories each, from low to high. Detailed information about sub indicators is available in Morero et al. (2014).

Competences_Ordinal.
1. **Low.** If the firm has only one or none of the structures listed in the category high below.
2. **Medium.** If the firm has only two of the structures listed in the category high below.
3. **High.** If the firm has a training structure, an R&D structure, and quality standards certifications.

Independent and dependent variables
Innovation_Continuous. Summing up the values in parenthesis = innovation in products new to the firm (1) + innovation in services new to the firm (1) + introduction of improved products new to the firm (1) + introduction of improved services new to the firm (1) + innovation in processes new to the firm (1) + organizational innovation new to the firm (1) + innovation in commercialization new to the firm (1) + innovation in products new to the market (3) + innovation in services new to the market (3) + innovation in processes new to the market (3) + organizational innovation new to the market (3) + innovation in commercialization new to the market (3).

Innovation_Ordinal.
1. **Low.** If Innovation_Continuous assumes values until 5.
2. **Medium.** If Innovation_Continuous assumes values 6 to 11.
3. **High.** If Innovation_Continuous assumes values 12 to 18.

Innovation_Binary.
0.**No.** If Innovation_Ordinal is Low.
1.**Yes.** If Innovation_Ordinal is Medium or High

Not_Internal_Not_External.

0.**No.**

1.**Yes.** If the firm has carried out neither internal nor external R&D activities, purchases of licenses, consultancies, or buying of specific software for innovation.

Only_Internal.

0.**No.**

1.**Yes.** If the firm has carried out internal R&D activities, but not considered any external innovation activity.

Only_External.

0.**No.**

1.**Yes.** If the firm has carried out any external innovation activity (external R&D, purchase of particular software for innovation or licenses, or paid for consultancies).

Internal_and_External.

0.**No.**

1.**Yes.** If the firm has carried out both internal R&D activities and any of the external innovation activity considered.

Internal_Activities_Ordinal.

1. **Low.** If the firm does only one or none of the activities listed in the category high below.
2. **Medium.** If the firm does only two of the activities listed in the category high below.
3. **High.** If the firm does development of internal software, internal R&D, and design of new products or processes.

External_Activities_Ordinal.

1. **Low.** If the firm does at most one of the activities listed in the category high below.
2. **Medium.** If the firm does only two of the activities listed in the category high below.
3. **High.** If the firm does at least three of these activities: purchase of software for innovation, consultancies, external R&D and purchase of licenses.