DISCUSSION PAPER SERIES

No. 4483

INNOVATION COMPLIMENTARITY AND SCALE OF PRODUCTION

Eugenio J Miravete and Jose C Pernias

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP4483.asp**

ISSN 0265-8003

INNOVATION COMPLIMENTARITY AND SCALE OF PRODUCTION

Eugenio J Miravete, University of Pennsylvania and CEPR
Jose C Pernias, Universidad Jaume I

Discussion Paper No. 4483
July 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 4483

July 2004

ABSTRACT

Innovation Complimentarity and Scale of Production*

This Paper is an empirical study on the existence of complementarity between product and process innovation. We present an econometrically feasible model that uses the information contained in the innovation profile of each firm to test for the existence of complementarity among production and innovation strategies. We apply the model to analyse the Spanish ceramic tiles industry where the adoption of the single firing furnace in the 1980s facilitated the introduction of new product designs as well as to opening new ways of organizing production. Our econometric results show that there is significant complementarity between product and process innovation. We are able to separate the nature of complementarity relationships and thus, our results show that both *intrinsic* – technologically driven – and *induced* complementarity – due to firms unobserved heterogeneity – are significant. Small firms tend to be more innovative overall.

JEL Classification: C52, L20 and O32
Keywords: complementarity, process innovation, product innovation, supermodularity and unobserved heterogeneity

Eugenio J Miravete
Department of Economics
University of Pennsylvania
McNeil Building, 3718 Locust Walk
Philadelphia, PA 19104-62971
USA
Tel: (1 215) 898 1505
Fax: (1 215) 573 2057
Email: miravete@ssc.upenn.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=139116

Jose C Pernias
Departamento de Economía
Universitat Jaume I.
Campus Riu Sec.
12071 Castellón
SPAIN
Tel: (34 964) 728 610
Fax: (34 964) 728 591
Email: pernias@eco.uji.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=161036

*For their comments, we thank Christopher Flinn, Neil Gandal, Rebecca Henderson, Bo Honoré, Harumi Ito, Boyan Jovanovic, Ariel Pakes, John Panzar, Lars-Hendrik Röller, Duncan Simester, Wilbert Van Der Klaauw, and Kenneth Wolpin, the editor, Kenneth Hendricks, and the two detailed anonymous referees that triggered a major revision of our Paper. We are also indebted to participants at the WZB conference on ‘Empirical Advances in Industrial Organization,’ Berlin, 1997; the NBER-Universities Research Conference on ‘Competition and Organization in Technology-Intensive Industries,’ Boston, 1997; the CIRANO conference on ‘Innovation and Supermodularity,’ Montreal, 2000; the ‘First International Industrial Organization Conference,’ Boston, 2003; as well as seminar audiences at Arizona State University, Instituto de Análisis Económico, University of Michigan, New York University, Universidad Politécnica de Cartagena, and Rice University. Angel Ortí facilitated our access to the data used in this Paper. Financial support through INSEAD research project 2010-5339R and by the CV Starr Center for Applied Economics at NYU is gratefully acknowledged.

Submitted 11 May 2004

1 Introduction

It has long been recognized that modern firms develop several innovative strategies to adjust to the challenging new conditions of increasingly integrated markets. International globalization of the economy and fast developments in telecommunications, computers, and information technology have revolutionized the way firms are organized, and have immensely increased firms ability to introduce new products, use new technologies, and experiment with new designs and/or manufacturing procedures. An immediate question that arises, both in theory and empirically, is whether these two forms of innovation are in some way related. If product and process innovation are complements these innovative strategies are mutually reinforcing because increasing the level of any of them increases the marginal profitability of the other, *e.g.*, Milgrom and Roberts (1990). Thus, for instance, the design of public policies that give incentives to develop one strategy should be aware of the "externalities" of such policies for other areas of decision of firms. Also, as Milgrom and Roberts (1995) first noticed, the existence of complementarities requires a high degree of coordination among firms' activities and, thus, they favor hierarchical organizational structures over flat ones.

Although the idea of complementarity is intuitively appealing, uncovering whether such complementarity among strategies exist turns out to be a very difficult task. The genuine difficulty arises because most of the time, testing for complementarity relies on measuring correlations among error terms of equations representing the optimal decision rules of firms. These simplified representation of the optimal decision rules may also include the effect of misspecification and/or missing variables in addition to individual unobserved heterogeneity of firms environments and organizational structure. As first noticed by Athey and Stern (1998), the existence of firms unobserved heterogeneity may be responsible for the correlation among strategies even though complementarity may not exist at all. Therefore, this paper uses actual data from the Spanish ceramic tiles industry to evaluate whether complementarities among innovation strategies exist while at the same time controlling for the observed correlation among strategies that might only be induced by firms' unobserved heterogeneity.[1]

To consistently measure the complementarity among innovation strategies, we develop and estimate a structural discrete choice model of production and innovation decisions that is capable of distinguishing between *intrinsic* (true) and *induced* complementarity (due to unobserved heterogeneity). Our estimation approach is based on the innovative profile of firms, *i.e.*, making use of the information revealed by the different combinations of joint innovation decisions that firms may adopt. There are a couple of reasons that favor our approach. First, we can explicitly deal with the existence of unobserved heterogeneity. This overcomes the criticism of Athey and Stern (1998) to the overwhelming majority of models used to test for the existence of complementarity. Second, these commonly used models deal with innovation strategies in isolation, then studying the correlations across error terms. We also show in this paper that dealing with innovations in isolation leads to a misspecified model if *intrinsic* complementarity is present. Furthermore, we show that attempting to accommodate the existence of *intrinsic* complementarity within this "in-isolation approach" leads to incoherent empirical models in the sense of Schmidt (1981).

[1]Our econometric approach also incorporates many contributions of the growing empirical literature on complementarities of innovations. The works of Arora and Gambardella (1990) and Ichniowski, Shaw, and Prennushi (1997) are two good examples of the existing attempts to test the implications of the complementarity hypothesis, although controlling only for observable firm's differences.

The main contribution of this paper is to develop an econometrically feasible, discrete choice, structural model of production and innovation decision that is able to identify the source of the complementarity relation among strategies if they exist. Our model builds, only partially, on the approaches of Athey and Schmutzler (1995) and Athey and Stern (1998). We focus on the existence of complementarity among scale, product, and process innovation. We suggest a very simple model of firms' production and innovation decision making. Firms are assumed to maximize profits non-cooperatively but simultaneously with respect to production and the innovative strategies that they want to pursue. We only distinguish whether these innovation activities are either demand enhancing or cost reducing. Within the framework of the present model, any general strategy aimed to reduce the substitutability of firms' product relative to the competitors' should be considered a demand enhancing innovation. By reducing the degree of substitution between the own's and competitors' products, the firm increases the potential mark-up that may charge to its customers. Similarly, any strategy aimed to give the firm a competitive advantage through unit cost reductions should be considered a cost reducing innovation.

In order to estimate this model and the implications of the complementarity hypothesis, we use data from DIRNOVA, a database of Spanish firms, which includes information on several innovation activities that they conduct. The information contained in this data set allows us to distinguish between demand enhancing and cost reducing oriented innovations, as well as to control for the effect of firms unobserved heterogeneity by means of a structural discrete choice model of production and innovation decision.

Our estimates show that there is complementarity among strategies in the ceramic tiles industry that we cover in this study. In general, we are able to document the existence of complementarity and we always reject model specifications that ignore these complementary relationships. Furthermore, our data allows us to identify the source of such complementarity, *i.e.*, distinguishing whether complementarity is *intrinsic* or *induced*. Our empirical results are consistent with industry configurations where small and medium sized firms have a comparative advantage in the adoption and employment of flexible manufacturing methods. Typical Schumpeterian arguments relating increases in the scale with higher probabilities of adopting or developing an innovation appear to fail because of the nature of the innovations considered here, none of which requires a large scale of production to be successfully implemented. This empirical evidence is also consistent with the view, *e.g.*, Milgrom and Roberts (1990), that the transition to flexible manufacturing methods involves radical and coordinated changes of the firms activities.

In the ceramic tiles industry product and process innovations are complements while smaller firms tend to engage more frequently in demand enhancing innovation. Our framework allows us also to identify the nature of these complementarities. For instance, the complementarity between the inverse of the scale and product innovation is *intrinsic*, *i.e.*, explained by technological relationships. Actually, the major innovation of this industry during the 1980s —the single firing furnace— allowed reducing the minimum efficient scale of production of profitable plants. This technology also allowed integrating product design more easily. However, taking advantage of such design capabilities required an active entrepreneurial will. We interpret that the *induced* nature of complementarity between product and process innovation supports the idea that taking advantages of technology to introduce new products mainly lays on elements of unobserved firm heterogeneity such as the organization of production, access to distribution channels, and/or background and experience of managers.

The paper is organized as follows. In Section 2 we describe the data and the features of the Spanish ceramic tiles industry. We also report some preliminary evidence of correlation among firms' strategic decision variables. This description helps to fix some ideas and support some of the assumptions of the simple model of monopolistic competition that we develop in Section 3. This section also includes a discussion on the effect of dealing with the two types of complementarities to specify a coherent econometric model. Section 4 presents the estimates of the structural model of production and innovation decisions for the Spanish ceramic tiles industry. Section 5 concludes. The Appendices include a discussion on the incoherence problems resulting from the estimation of the determinants of innovation strategies in isolation, and a detailed derivation of the likelihood function used to estimate the suggested structural model.

2 The Data

This section describes the data set used in this study and presents some preliminary evidence in favor of the complementarity between product and process innovation, as well as pointing out the relation between innovative profile and size of the firms. In this section we also present some measures of conditional pairwise correlation between scale of production, product, and process innovation, respectively.

2.1 The Spanish Ceramic Tiles Industry

This Spanish ceramic tiles industry is currently ranked second largest in the world, after the Italian and ahead of the Brazilian industry. At the beginning of the eighties, Spanish ceramic tiles production suffered from technological backwardness as compared to the Italian industry, and it could only compete in international markets targeting low quality niches on a basis of labor intensive production techniques. But by mid eighties, many firms had already adopted the single firing furnace. This was a major innovation. Compared to the existing technology (product specific first firing furnace and full/half cycle double firing furnace), it transformed the production of tiles in a much more energy efficient and automated process. Furthermore, this new technology also allowed the production of high quality, low water absorbtion tiles of larger dimensions and different shapes and colors. Design and application of computerization was easily embedded in the production process, and an increasing variety of new high quality products flooded the market afterwards.

There might be strong arguments in favor of the complementarity of product and process innovation being in this case purely technical as the process innovation actually allowed for new designs that were not possible with the previous technology. However, new designs only became profitable as the domestic market matured. In addition, as wages increased due to economic expansion, the adoption of the single firing furnace became the optimal strategy for firms in this industry because they otherwise could not compete even in the domestic market. Wage increases that accompanied the fast economic growth in Spain during the second half of the 1980s acted as a trigger of a series of manufacturing changes that eventually led to a drastic transformation in the organization of the ceramic tiles firms. The adoption of this process innovation eased the introduction of new designs and thus increased the marginal profitability of carrying some sort of product innovation, which in turn also increased the marginal return of the investment in single firing furnaces. As our results report, product and process innovation appear to go hand in hand although technology alone does not explain such positive correlation

Table 1: Descriptive Statistics

	Mean	Std. Dev.
OUTPUT	5.384	1.938
PRODUCT	0.347	0.476
PROCESS	0.361	0.480
EX	0.253	0.250
EU	0.419	0.493
EX · EU	0.140	0.230
TM	0.646	0.478
TMHI	0.213	0.409
AGE	2.681	0.719
MPROD	0.387	0.487
MPRODHI	0.037	0.189
EXIT	0.069	0.254
ENTRY	0.049	0.215
TIME	0.595	0.491

All variables defined in the text. OUTPUT is measured in logarithm of 1986 million Pesetas and AGE is the logarithm of years since the creation of the firm.

between innovative strategies. Unobserved factors, firm specific characteristics, possibly access to distribution channels, and managerial ability may also explain such complementarity.

The ceramic tiles example is also illuminating in relation with the role of the scale of production plays in the adoption of innovations. The single firing furnace was a major labor saving innovation but required a complete restructuring of the production plant. Furthermore, and contrary to the Schumpeterian rule, it reduced the minimum efficient scale of production. Thus, firms who adopted the new furnace underwent a major transformation that turned profitable low levels of production. Obviously, under these circumstances, using the scale of production as exogenous regressor will lead to simultaneity bias in the estimation and most likely to inconsistent estimates in our nonlinear econometric model. We therefore include the scale among the endogenous decisions variables of our model.

2.2 Scale and Innovative Behavior

Our econometric analysis uses data from DIRNOVA, a database of Spanish firms for 1988 and 1992. This database was collected by IMPIVA, a public agency in charge of promoting international agreements on transfer of technology, commercial distribution, joint-ventures, and subcontracting between Spanish and foreign firms. The DIRNOVA database contains data built from information obtained through direct interviews with managers of the companies applying a systematic methodology for its collection over the years. Furthermore, it covers an interesting period of transformation of the Spanish economy after joining the European Union in 1986, and during a strong period of economic growth that lasted until the end of 1992.

Table 1 presents the descriptive statistics for the ceramic tiles industry. For each firm we know the output level (in logarithm), OUTPUT; whether firms engage in demand enhancing, or cost reducing innovation, PRODUCT and PROCESS, respectively; the percentage of the total production that is exported, EX; whether the European Union is the principal foreign market for foreign sales, EU; whether firms have at least one or two registered trademarks, TM and TMHI, respectively; the number of years that the firm has been in business (in logarithm), AGE; a dummy variable, MPROD, to indicate whether firms produce more than one product as defined at the 7-digits SIC level; and MPRODHI to indicate whether firms produce at least three products at the 7-digit SIC level. In addition we also include a TIME dummy for observations corresponding to year 1992 and define an ENTRY dummy to indicate those firms who are only present at the 1992 sample and an EXIT dummy to identify those firms that are only present in the 1986 sample. According to Table 1, the Spanish ceramic tile industry is characterized by middle sized firms, with an average of about ten years of presence in the industry. About one third of these firms engage in product and/or process innovation. Although nearly 80% of firms export, the typical firm only sells abroad about one quarter of its production, being the European Union the main destination of the ceramic tiles exports. Most firms only manufacture a single product and most of them own registered trademarks to differentiate from competitors.

We observe dummy indicators of the innovative strategies in which firms are engaged in every period. We denote $x_d = 1$ when firms acknowledge that they participate in marketing and advertising projects, which is obviously related to demand enhancing innovation activities. Similarly we define $x_c = 1$ whenever firms participate in the development of new manufacturing projects, which is more related to cost reducing innovation strategies.[2] Note that these innovation indicators identify potentially reversible strategies, *i.e.*, they do not necessarily represent decisions on investments such as particular adoption of capital-embodied innovations. Although it is expected that firms generally make use of these strategies during several periods, it is possible that those innovation strategies may be discontinued later. The advantage of these indicators is that they are unequivocally related to the innovative profile of firms and both belong to the last stages of the innovation management process.[3]

Given our dichotomous indicators of innovation activities, four innovation profiles are possible. Firms may specialize in either product or process innovation, as suggested by Abernathy and Utterback (1978) or Kleeper (1996), they may either not innovate at all, or more interestingly, if the innovate, they do so in both activities, more in accordance with Milgrom and Roberts' (1990) view of the modern manufacturing process. Table 2 reports the proportion of firms that follows each combination of innovation strategies. These results are also stratified by the scale of firms, x_y, distinguishing whether they are above or below the mean scale of each industry. For the whole sample, over 20% of firms innovate simultaneously in product and process. Non-innovative firms amount to 50% of the sample. The remaining 30% of firms either only carry out product or process innovations. This pattern is also observed in the high and low scale samples, although smaller firms appear to innovate slightly more frequently but they also adopt more often those innovation profiles comprised of only one practice. With respect to the scale of production, firms engaged in both innovations are generally smaller than those not involved in

[2]Other available indicators of process innovation are the license and assistance in production, acquisition and/or transfer of technology, and staff training programs. The other product innovation indicator available is the expansion of the commercial structure. We have estimated the model using different combinations of these product and process innovation, but results regarding complementarities are robust to using different indicators of product and process innovation.

[3]An alternative valid interpretation of our results would be to analyze complementarities between marketing and manufacturing strategies, but we prefer the more general setup of Section 3 where marketing and manufacturing strategies are indicators correlated to demand enhancing and cost reducing strategies as defined in equations (2) and (3) below.

Table 2: Innovation Choices and Scale of Production

	N	Freq. (%)	Mean scale	S.D. scale
Whole Sample				
Both	88	20.4	5.29	2.20
Only product	62	14.4	5.02	2.13
Only process	68	15.7	5.56	1.75
None	214	49.5	5.47	1.80
All firms	432		5.38	1.94
High Scale Sample				
Both	63	21.1	6.37	0.78
Only product	39	13.1	6.20	0.76
Only process	46	15.4	6.42	0.86
None	150	50.3	6.30	0.84
All firms	298		6.32	0.82
Low Scale Sample				
Both	25	18.7	2.58	2.29
Only product	23	17.2	3.00	2.19
Only process	22	16.4	3.77	1.78
None	64	47.8	3.53	1.95
All firms	134		3.30	2.08

N is the number of firms adopting each innovation profile and "Freq" is the share of these firms relative to the total number of firms in each sample, Freq. We also report the mean and standard deviation, (S.D.), of the scale of production. High (low) scale sample only includes those firms whose output levels are above (below) the overall mean output.

innovating at all. This difference in size is more pronounced in the low scale sample, while in the high scale sample there are practically no differences of size between both groups of firms. Interestingly, firms engaged in product innovation are smaller than those that do not adopt this strategy —the mean scale is $\bar{x}_y = 5.18$ for those firms where $x_d = 1$ and $\bar{x}_y = 5.49$ when $x_d = 0$— but firms doing process innovation are slightly larger than those that does not carry out this kind of innovation practices —$\bar{x}_y = 5.41$ when $x_c = 1$ and $\bar{x}_y = 5.37$ when $x_c = 0$.

2.3 Unconditional Association

Association among strategies is the direct consequence of the supermodularity of the profit function in the decision variables $\{x_y, x_d, x_c\}$. Thus, the existence of complementarity relationships among strategies leads to pairwise monotone co-movements of the endogenous variables.[4]

Are the apparent complementarity relationships shown in Table 2 significant? To answer this question, Table 3 reports Kendall's τ coefficients of association among decision variables, *i.e.*, production, product, and process innovation.[5] We test the null hypothesis of independence

[4] The theoretical foundations of this comparative static result is shown to hold by Hölmstrom and Milgrom (1994), but was earlier implemented empirically by Arora and Gambardella (1990).

[5] We also computed Pearson's linear correlation coefficient, r, and Spearman's rank-order correlation coefficient

Table 3: Unconditional Association of Strategies

	Whole Sample	High Scale Sample	Low Scale Sample
PRODUCT, PROCESS	0.321 [0.000]	0.350 [0.000]	0.253 [0.000]
PRODUCT, OUTPUT	−0.024 [0.454]	−0.001 [0.983]	−0.178 [0.000]
PROCESS, OUTPUT	0.036 [0.260]	0.066 [0.042]	−0.057 [0.078]
N	432	298	134

Kendall's τ coefficients of association and asymptotic p-values between brackets. High (low) scale sample only includes those firms whose output level is above (below) the overall mean. N is the number of observations in each sample.

between pairs of decision variables. As in Table 2, we compute these association measures stratified by scale of the firms as well as for the whole sample. Results indicate that product and process innovation are positively associated regardless of the scale of production of firms. As for the other relationships, product innovation and scale of production are negatively associated in the case of small ceramic tiles firms, while the association between process innovation and scale is more tenuous, appearing to be positive in the case of large firms but negative in the case of small firms.

If ceramic tiles firms where identical beyond these three decision variables, we could conclude that its profit function would be supermodular only in (x_d, x_c). However, this descriptive analysis does not condition on any observed or unobserved heterogeneity of firms. Furthermore, pairwise association measures are far too weak a tool to distinguish whether the observed association respond to the existence of intrinsic rather than induced complementarity. The following section presents an econometric framework where such distinction is feasible.

3 Production and Innovation in Modern Manufacturing

In this section we present a highly stylized model of vertical and horizontal product differentiation that allows us to study the relation between firms' optimal production and innovation decisions. We discuss in detail this econometrically feasible model that, most importantly, is capable of identifying whether the observed complementarity is *intrinsic* or *induced* by firms' unobserved heterogeneity. This theoretical framework provides us with simple testable hypothesis on the existence of strategic complementarities among the firms' decision variables.

In this section we first review the general theory of supermodular profit functions and its relation to complementary strategies. We later present our econometric specification of the profit function and finally discuss the estimation of such model.

r_s, which lead to the same qualitative results. However Pearson's r is not adequate to measure possibly nonlinear association between two variables. Nonparametric correlation is more robust than linear correlation to the existence of outliers. Spearman's r_s and Kendall's τ are invariant to monotone transformations of variables. However, Kendall's τ is more nonparametric and generally preferred to Spearman's r_s because it uses only the relative ordering of ranks instead of their numerical difference. See Press, Flannery, Teulosky, and Vetterling (1986, §14.5–14.6) for the computation of Kendall's τ and its asymptotic distribution.

3.1 Supermodularity and Complementarity

We contemplate a framework where firm i decides on the output level, $x_{y\,i}$, and wether to implement a demand enhancing or a cost reducing innovation, $x_{d\,i}$ and $x_{c\,i}$ respectively. The vector $\boldsymbol{x}_i = (x_{y\,i}, x_{d\,i}, x_{c\,i})'$ represents the decision variables of the firm. Obviously, the characteristics of each firm and the market where it operates determine the relative profitability of different production and innovation strategies. We distinguish among three separate types of environmental parameters: revenue specific, $\boldsymbol{Z}_{r\,i}$, cost specific, $\boldsymbol{Z}_{c\,i}$, and technology specific characteristics, $\boldsymbol{Z}_{k\,i}$. The economic environment of the firm is therefore represented by the vector $\boldsymbol{Z}_i = (\boldsymbol{Z}'_{r\,i}, \boldsymbol{Z}'_{c\,i}, \boldsymbol{Z}'_{k\,i})'$ of exogenous variables.

Deciding how these environmental variables affect each component of the profit function defines a model of firm behavior. Our general model of production and innovation decision is summarized by the following profit function:

$$\pi(\boldsymbol{x}_i; \boldsymbol{Z}_i) = R(x_{d\,i}, x_{y\,i}; \boldsymbol{Z}_{r\,i}) - C(x_{c\,i}, x_{y\,i}; \boldsymbol{Z}_{c\,i}) - K(x_{d\,i}, x_{c\,i}; \boldsymbol{Z}_{k\,i}). \tag{1}$$

Model (1) is quite general and captures many of the features of a flexible manufacturing system. In addition to production, firms engage in process innovation in order to improve their competitive position in their respective markets. The product-innovative firm introduces new designs to differentiate from competitors. This demand enhancing innovation x_d shifts the firm's residual demand, thus shifting the revenue function $R(\cdot)$ up:

$$R(1, x_{y\,i}; \boldsymbol{Z}_{r\,i}) \geq R(0, x_{y\,i}; \boldsymbol{Z}_{r\,i}). \tag{2}$$

Similarly, the process-innovative firm obtains a competitive advantage by reducing its total cost of production $C(\cdot)$ through the application of better technology and/or more efficient process management methods, x_c:

$$C(1, x_{y\,i}; \boldsymbol{Z}_{c\,i}) \leq C(0, x_{y\,i}; \boldsymbol{Z}_{c\,i}). \tag{3}$$

Finally, both innovations are costly to implement, and thus, the induced increase in demand or unit cost reduction has to compensate the adoption cost of innovations, $K(\cdot)$.

The maximization problem of any firm i consists in choosing the scale of production $x_{y\,i} \in \mathbb{R}$, as well as whether to engage in product and process innovation, $x_{d\,i} \in \{0, 1\}$ and $x_{c\,i} \in \{0, 1\}$, respectively. While the first variable is continuous, the other two are discrete, and thus, firms face a non-convex decision problem.

The discreteness of the decision variables requires at a theoretical level that we define the set of control variables over a lattice. A lattice is defined by the set $\mathbb{X}$ and the partial order $\geq$, where $\forall x, x' \in \mathbb{X}$, the set $\mathbb{X}$ also contains a smallest element under the order that is larger than both x and x', and a largest element that is smaller than both. If $\mathbb{X} = \mathbb{R}^N$, the *join* operator is defined as $x \vee x' = (\min\{x_1, x'_1\}, \min\{x_2, x'_2\}, \ldots, \min\{x_N, x'_N\})$. Similarly the *meet* operator is defined as $x \wedge x' = (\max\{x_1, x'_1\}, \max\{x_2, x'_2\}, \ldots, \max\{x_N, x'_N\})$. Thus, the subset $\mathbb{S} \subseteq \mathbb{X}$ is a sublattice if it is closed under under the *join* and *meet* operations.[6]

Therefore, we assume that $\mathbb{X}$, the set of control variables, is a lattice, while $\mathbb{Z}$, the set of environmental variables, is a partially ordered set. Thus, $\pi(\boldsymbol{x}; \boldsymbol{Z})$ is supermodular in $\mathbb{X}$ if $\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{X}$, and $\forall \boldsymbol{Z} \in \mathbb{Z}$, the following condition holds:

$$\pi(\boldsymbol{x}; \boldsymbol{Z}) + \pi(\boldsymbol{x}'; \boldsymbol{Z}) \leq \pi(\boldsymbol{x} \vee \boldsymbol{x}'; \boldsymbol{Z}) + \pi(\boldsymbol{x} \wedge \boldsymbol{x}'; \boldsymbol{Z}). \tag{4}$$

[6] For an extensive introduction to lattices and set defined functions see Topkis (1998).

Notice that the very same definition of supermodularity of the profit function embodies the idea of complementarity among the decision variables, $\boldsymbol{x}$. Increasing all decision variables separately does not increase profits in the same magnitude than increasing all of them simultaneously. This can easily be proven by rewriting condition (4) as

$$[\pi(\boldsymbol{x};\boldsymbol{Z}) - \pi(\boldsymbol{x} \vee \boldsymbol{x}';\boldsymbol{Z})] + [\pi(\boldsymbol{x}';\boldsymbol{Z}) - \pi(\boldsymbol{x} \vee \boldsymbol{x}';\boldsymbol{Z})] \leq \pi(\boldsymbol{x} \wedge \boldsymbol{x}';\boldsymbol{Z}) - \pi(\boldsymbol{x} \vee \boldsymbol{x}';\boldsymbol{Z}). \qquad (5)$$

At first sight, it may appear that our model is focused on many non testable hypotheses. However, we are only making weak assumptions that limit the pairwise interactions between production and innovation strategies. For the profit function (1) to be supermodular in production and innovation strategies, it is just needed that product innovation, $x_{d\,i}$, shifts the firm's marginal revenue up, and that process innovation, $x_{c\,i}$, shifts marginal production costs down. Furthermore, complementarity between product and process innovation denotes the existence of some scope economies in the adoption of such strategies. When synergies are present, we should expect that choice variables move all together. Contemporaneous complementarity will therefore induce positive pairwise correlation among strategies in a cross-section sample. In the following section we specify an econometric model that can accommodate these restricted pairwise movements of the decision variables to identify the existence and magnitude of complementarities among the strategies of firms.

3.2 Model Specification

At the empirical level, the dichotomous nature of some of the choice variables requires that we set up a structural discrete choice model that predicts the proportions in which different combinations of these discrete strategies appear in the sample. In order to do so we assume a second order approximation to the component functions of firms' profits. This approximation is effectively quadratic in output but only includes innovation dummies and their products. Thus, the revenue, production cost and adoption cost functions of firm i can be written as:

$$\begin{aligned} R(x_{d\,i}, x_{y\,i}; \boldsymbol{Z}_{r\,i}) = {} & \alpha_d x_{d\,i} + \alpha_y x_{y\,i} + \delta_{dy} x_{d\,i} x_{y\,i} \\ & + \boldsymbol{\theta}'_{dr} \boldsymbol{z}_{r\,i} x_{d\,i} + \boldsymbol{\theta}'_{yr} \boldsymbol{z}_{r\,i} x_{y\,i} + \boldsymbol{\psi}'_{dr} \boldsymbol{\zeta}_{r\,i} x_{d\,i} + \boldsymbol{\psi}'_{yr} \boldsymbol{\zeta}_{r\,i} x_{y\,i} - (\gamma_r/2) x^2_{y\,i}, \end{aligned} \qquad (6a)$$

$$\begin{aligned} C(x_{c\,i}, x_{y\,i}; \boldsymbol{Z}_{c\,i}) = {} & \beta_c x_{c\,i} + \beta_y x_{y\,i} - \delta_{cy} x_{c\,i} x_{y\,i} \\ & - \boldsymbol{\theta}'_{cc} \boldsymbol{z}_{c\,i} x_{c\,i} - \boldsymbol{\theta}'_{yc} \boldsymbol{z}_{c\,i} x_{y\,i} - \boldsymbol{\psi}'_{cc} \boldsymbol{\zeta}_{c\,i} x_{c\,i} - \boldsymbol{\psi}'_{yc} \boldsymbol{\zeta}_{c\,i} x_{y\,i} - (\gamma_c/2) x^2_{y\,i}, \end{aligned} \qquad (6b)$$

$$\begin{aligned} K(x_{d\,i}, x_{c\,i}; \boldsymbol{Z}_{k\,i}) = {} & \eta_d x_{d\,i} + \eta_c x_{c\,i} - \delta_{dc} x_{d\,i} x_{c\,i} \\ & - \boldsymbol{\theta}'_{dk} \boldsymbol{z}_{k\,i} x_{d\,i} - \boldsymbol{\theta}'_{ck} \boldsymbol{z}_{k\,i} x_{c\,i} - \boldsymbol{\psi}'_{dk} \boldsymbol{\zeta}_{k\,i} x_{d\,i} - \boldsymbol{\psi}'_{ck} \boldsymbol{\zeta}_{k\,i} x_{c\,i}, \end{aligned} \qquad (6c)$$

where vectors $\boldsymbol{Z}_{r\,i} = (\boldsymbol{z}'_{r\,i}, \boldsymbol{\zeta}'_{r\,i})'$, $\boldsymbol{Z}_{c\,i} = (\boldsymbol{z}'_{c\,i}, \boldsymbol{\zeta}'_{c\,i})'$, and $\boldsymbol{Z}_{k\,i} = (\boldsymbol{z}'_{k\,i}, \boldsymbol{\zeta}'_{k\,i})'$ comprise all environmental variables of firms. Among those variables, $\boldsymbol{z}_{r\,i}$, $\boldsymbol{z}_{c\,i}$, and $\boldsymbol{z}_{k\,i}$ are observed by the econometrician but $\boldsymbol{\zeta}_{r\,i}$, $\boldsymbol{\zeta}_{c\,i}$, and $\boldsymbol{\zeta}_{k\,i}$ are not. This latter set of variables represents the unobserved heterogeneity of firms. Notice that both, observed and unobserved heterogeneity affect the marginal return of the different choices. Therefore, a conditional association analysis similar to that of Table 3 but controlling for the effect of observable characteristics does not suffice to conclude whether firms' strategies are truly complements, or on the contrary the observed association is only induced by the unobserved heterogeneity of firms.[7] After substituting (6a)–(6c) into (1), the profit function

[7]This is the main point of the work by Athey and Stern (1998).

becomes:

$$\begin{aligned}\pi(x_{d\,i}, x_{c\,i}, x_{y\,i}) = {} & (\theta_{d0} + \boldsymbol{\theta}_{dr}' \boldsymbol{z}_{r\,i} + \boldsymbol{\theta}_{dk}' \boldsymbol{z}_{k\,i} + \boldsymbol{\psi}_{dr}' \boldsymbol{\zeta}_{r\,i} + \boldsymbol{\psi}_{dk}' \boldsymbol{\zeta}_{k\,i}) x_{d\,i} \\ & + (\theta_{c0} + \boldsymbol{\theta}_{cc}' \boldsymbol{z}_{c\,i} + \boldsymbol{\theta}_{ck}' \boldsymbol{z}_{k\,i} + \boldsymbol{\psi}_{cc}' \boldsymbol{\zeta}_{c\,i} + \boldsymbol{\psi}_{ck}' \boldsymbol{\zeta}_{k\,i}) x_{c\,i} \\ & + (\theta_{y0} + \boldsymbol{\theta}_{yr}' \boldsymbol{z}_{r\,i} + \boldsymbol{\theta}_{yc}' \boldsymbol{z}_{c\,i} + \boldsymbol{\psi}_{yr}' \boldsymbol{\zeta}_{r\,i} + \boldsymbol{\psi}_{yc}' \boldsymbol{\zeta}_{c\,i}) x_{y\,i} \\ & + \delta_{dy} x_{d\,i} x_{y\,i} + \delta_{dc} x_{d\,i} x_{c\,i} + \delta_{cy} x_{c\,i} x_{y\,i} - (\gamma/2) x_{y\,i}^2,\end{aligned} \tag{7}$$

where $\theta_{d0} = \alpha_d - \eta_d$, $\theta_{c0} = -\beta_c - \eta_k$, $\theta_{y0} = \alpha_y - \eta_y$, and $\gamma = \gamma_r - \gamma_c$. The profit function is concave in the scale of production whenever $\gamma > 0$. This parameter is however not identifiable with the current data because we do not observe the level of profits associated to each scale of production and innovation profile of firms. Our estimation procedure, described below, is based only on the information revealed by the optimal decisions of the firm. So, we have no means to identify γ as a consequence of the invariance of the set of maximizers under monotone transformations of the profit function. We therefore normalize $\gamma = 1$ and implicitly assume that the profit function is well behaved.

Next, distinguishing (and grouping) the elements of observed and unobserved heterogeneity of firms environment we can rewrite profit function (7) as:

$$\begin{aligned}\pi(x_{d\,i}, x_{c\,i}, x_{y\,i}) = {} & (\theta_{d\,i} + \epsilon_{d\,i}) x_{d\,i} + (\theta_{c\,i} + \epsilon_{c\,i}) x_{c\,i} + (\theta_{y\,i} + \epsilon_{y\,i}) x_{y\,i} \\ & + \delta_{dc} x_{d\,i} x_{c\,i} + \delta_{dy} x_{d\,i} x_{y\,i} + \delta_{cy} x_{c\,i} x_{y\,i} - x_{y\,i}^2/2.\end{aligned} \tag{8}$$

Comparing equations (7) and (8) and equating coefficients, we realize that functions $\theta's$ and $\epsilon's$ summarize the effect of observable and unobservable firm heterogeneity as represented by the following linear transformations:

$$\theta_{d\,i} = \theta_d(\boldsymbol{z}_{r\,i}, \boldsymbol{z}_{k\,i}) = \theta_{d0} + \boldsymbol{\theta}_{dr}' \boldsymbol{z}_{r\,i} + \boldsymbol{\theta}_{dk}' \boldsymbol{z}_{k\,i}, \tag{9a}$$

$$\theta_{c\,i} = \theta_c(\boldsymbol{z}_{c\,i}, \boldsymbol{z}_{k\,i}) = \theta_{c0} + \boldsymbol{\theta}_{cc}' \boldsymbol{z}_{c\,i} + \boldsymbol{\theta}_{ck}' \boldsymbol{z}_{k\,i}, \tag{9b}$$

$$\theta_{y\,i} = \theta_y(\boldsymbol{z}_{r\,i}, \boldsymbol{z}_{c\,i}) = \theta_{y0} + \boldsymbol{\theta}_{yr}' \boldsymbol{z}_{r\,i} + \boldsymbol{\theta}_{yc}' \boldsymbol{z}_{c\,i}, \tag{9c}$$

$$\epsilon_{d\,i} = \boldsymbol{\psi}_{dr}' \boldsymbol{\zeta}_{r\,i} + \boldsymbol{\psi}_{dk}' \boldsymbol{\zeta}_{k\,i}, \tag{9d}$$

$$\epsilon_{c\,i} = \boldsymbol{\psi}_{cc}' \boldsymbol{\zeta}_{c\,i} + \boldsymbol{\psi}_{ck}' \boldsymbol{\zeta}_{k\,i}, \tag{9e}$$

$$\epsilon_{y\,i} = \boldsymbol{\psi}_{yr}' \boldsymbol{\zeta}_{r\,i} + \boldsymbol{\psi}_{yc}' \boldsymbol{\zeta}_{c\,i}. \tag{9f}$$

In order to estimate the parameters of (8) by the maximum-likelihood method, we need to specify a known family of distributions from which particular realizations of unobservables $\boldsymbol{\epsilon}_i = (\epsilon_{d\,i}, \epsilon_{c\,i}, \epsilon_{y\,i})$ are drawn. We assume that $\boldsymbol{\epsilon}_i$ follows a trivariate normal distribution with zero mean, standard deviations denoted as $(\sigma_d, \sigma_c, \sigma_y)$, and correlation matrix given by:

$$\boldsymbol{R} = \begin{bmatrix} 1 & \rho_{dc} & \rho_{dy} \\ \rho_{dc} & 1 & \rho_{cy} \\ \rho_{dy} & \rho_{cy} & 1 \end{bmatrix}. \tag{10}$$

Several comments are worth being pointed here. First, equation (8) shows the different strategy-related sources that contribute to the profits of firm i. The first term in the right hand side of equation (8) captures the direct profitability of adopting the demand-enhancing innovation. This direct return is divided into two components: $\theta_{d\,i}$, which is related to the observable characteristics of the firm $\boldsymbol{z}_{r\,i}$ and $\boldsymbol{z}_{k\,i}$, as shown in equation (9a), and $\epsilon_{d\,i}$, which comprises organizational, managerial, or simply non-observed environmental factors that also affect the profitability of product innovation. In a similar manner, the following two terms represent the direct return of process inovation, $(\theta_{c\,i} + \epsilon_{c\,i})$, and, apart of a second order term, the marginal profitability of the scale of production, $(\theta_{y\,i} + \epsilon_{y\,i})$. In addition to these direct returns, the magnitude of intrinsic complementarities among strategies, identified by parameters δ_{dc}, δ_{dy}, and δ_{cy}, also affect profits. The last term of equation (8) captures the curvature of the profit function with respect to $x_{y\,i}$.

Second, supermodularity in the decision variables of profit function (8) depends solely on the signs of parameters δ_{dc}, δ_{dy}, and δ_{cy}. For example, (8) is supermodular in $(x_{y\,i}, x_{d\,i}, x_{c\,i})$ as long as $\delta_{dy} > 0$, $\delta_{dc} > 0$, and $\delta_{cy} > 0$. Thus, as discussed above, $\delta_{dy} > 0$ implies that the adoption of a product innovation increases marginal returns and, therefore, favors the simultaneous expansion of the scale of production. Alternatively, $\delta_{dy} > 0$ also means that higher levels of production increase the total profitability of adopting demand-enhancing innovations. Similarly, $\delta_{cy} > 0$ indicates that process innovation shifts the marginal costs of production down, so carrying out this kind of innovation leads to higher levels of production. Finally, $\delta_{dc} > 0$ indicates that the adoption of one of the innovation practices reduces the cost of adopting the other innovation strategy.

Third, equation (8) also points out other sources of association among decision variables besides intrinsic complementarities. For example, a positive association between product innovation and scale of production could be due to a variation in one of the elements of $\boldsymbol{z}_{r\,i}$ that simultaneously increases the direct returns to $x_{d\,i}$, through the $\theta_{d\,i}$ function of equation (9a), and the marginal returns to $x_{y\,i}$, through $\theta_{y\,i}$, see (9c). In a similar fashion, the unobserved heterogeneity captured by the variables in $\boldsymbol{\epsilon}_i$ could also lead to co-movements of the elements of $\boldsymbol{x}_i$. If, for example, there is a positive correlation between $\epsilon_{d\,i}$ y $\epsilon_{c\,i}$, *i.e.*, $\rho_{dc} > 0$, then a simultaneous increase in both $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ rises the returns to both innovation activities. Therefore, the effect of association of strategies due to firms' unobserved heterogeneity is captured by parameters ρ_{dy}, ρ_{cy}, and ρ_{dc}. Notice that, as could be seen in equations (9d)–(9f), the unobservables $\epsilon_{d\,i}$, $\epsilon_{c\,i}$, and $\epsilon_{c\,i}$ share common determinants, so we should not neglect in advance the presence of complementarities *induced* by unobserved organizational and/or managerial factors.

Fourth, observe, for instance, that $\theta_{d\,i}$ depends on revenue shifting and cost of adoption environmental variables, *i.e.*, $\boldsymbol{z}_{r\,i}$ and $\boldsymbol{z}_{k\,i}$, respectively. However, it does not depend on cost of production variables, $\boldsymbol{z}_{c\,i}$. A similar analysis for $\theta_{c\,i}$ and $\theta_{y\,i}$ also reveals that these exclusion restrictions on the elements of $\boldsymbol{z}_i = (\boldsymbol{z}'_{r\,i}, \boldsymbol{z}'_{c\,i}, \boldsymbol{z}'_{k\,i})'$ clearly identify shifts of different elements of the profit function and allow us to identify whether profit movements are originated by variations of the returns to product innovation, process innovation, or scale of production.

The model is therefore very flexible and it may acommodate three sources of association among decision variables: *intrinsic* complementarities, measured by δ_{dc}, δ_{dy}, and δ_{cy}; common impacts on returns to strategies originated by observable features of the production process that we can control, at least partially, through the observable environmental variables $\boldsymbol{z}_i$; and association *induced* by unobserved factors that also affects the returns to different firms' activities.

3.3 Estimation Approach

Ideally, we would like to observe the strategies used by each firm as well as a measure of their combined profitability. Unfortunately, we do not observe the revenue and cost functions of each firm. Thus, many of the parameters of equation (7) cannot be identified. Still, from the observed decisions on scale of production and innovation we can recover enough parameters —those of equation (8)— to consistently test for complementarity and identify whether the association among strategies is due to the existence of intrinsic complementarity or if it is only induced by the unobserved heterogeneity of firms.

In the absence of profit data, we should base our inference in the set of optimality conditions that determine the simultaneous choice of strategies that characterize each innovation profile. However, the usual practice in the analysis of complementarity is to study the determinants of adoption of each innovation strategy in isolation from each other. These reduced form demands for innovations are not formally derived from maximizing behavior, and they do not necessarily correspond to a properly defined profit function. We prove in Appendix A that this misleading approach leads to an incoherent system of equations with continuous plus dichotomous endogenous variables that will only allow us to study the restrictive case where all complementarity is *induced*, *i.e.*, when $\delta_{dc} = \delta_{dy} = \delta_{cy} = 0$. *Intrinsic* complementarity would be ruled out, and thus the estimates could turn out to be inconsistent if such source of complementarity is present.

Instead of considering the choice of strategies in isolation, we deal with the optimal choice of innovative profiles and scale of production. First, we simplify the objective function by substituting the optimal value of the continuous scale variable, through the corresponding optimality condition. From the first order condition of maximizing (8) with respect to $x_{y\,i}$ we obtain the optimal scale choice as a function of the innovation strategies $x_{d\,i}$ and $x_{c\,i}$:

$$x_{y\,i} = \theta_{y\,i} + \delta_{dy} x_{d\,i} + \delta_{cy} x_{c\,i} + \epsilon_{y\,i}. \tag{11}$$

After substituting this expression into the profit function (8), the profits can now be written as an exclusive function of the discrete innovation strategies:

$$\begin{aligned} \pi(x_{d\,i}, x_{c\,i}) = (\theta_{d\,i} + \epsilon_{d\,i}) x_{d\,i} + (\theta_{c\,i} + \epsilon_{c\,i}) x_{c\,i} + \delta_{dc} x_{d\,i} x_{c\,i} \\ + \frac{1}{2}(\theta_{y\,i} + \epsilon_{y\,i} + \delta_{dy} x_{d\,i} + \delta_{cy} x_{c\,i})^2. \end{aligned} \tag{12}$$

Next, it is convenient to define the following magnitudes:

$$\kappa_{y\,i} = \theta_{y\,i} + \epsilon_{y\,i}, \tag{13a}$$

$$\kappa_{d\,i} = \theta_{d\,i} + \delta_{dy}^2/2 + \delta_{dy}\kappa_{y\,i}, \tag{13b}$$

$$\kappa_{c\,i} = \theta_{c\,i} + \delta_{cy}^2/2 + \delta_{cy}\kappa_{y\,i}, \tag{13c}$$

$$\pi_{0\,i} = \kappa_{y\,i}^2/2, \tag{13d}$$

$$\delta = \delta_{dc} + \delta_{dy}\delta_{cy}, \tag{13e}$$

so that we can write the profits function (12) as follows:

$$\pi(x_{d\,i}, x_{c\,i}) = (\kappa_{d\,i} + \epsilon_{d\,i}) x_{d\,i} + (\kappa_{c\,i} + \epsilon_{c\,i}) x_{c\,i} + \delta x_{d\,i} x_{c\,i} + \pi_{0\,i}. \tag{14}$$

Profits from different innovation profiles are divided into $\pi_{0\,i}$, the profits from not innovating at all; the direct returns of product innovation, whether observable, $\kappa_{d\,i}$, or unobservable, $\epsilon_{d\,i}$; the direct returns of process innovation, again distinguishing between observable, $\kappa_{c\,i}$, or unobservable, $\epsilon_{c\,i}$; and δ, which measures the the intrinsic complementarity between innovation strategies. Notice that the the observable returns, $\kappa_{d\,i}$, $\kappa_{c\,i}$, and δ, already include the interaction among the innovation strategies and the optimal scale, as depicted in equations (13a)–(13e).

A profit maximizing firm chooses the combination of innovation strategies that leads to higher profits. For instance, a firm engages in simultaneous product and process innovation if

Table 4: Unobserved Heterogeneity and Choice of Innovation Profile

$S_i(1,1): \begin{cases} \epsilon_{d\,i} > -\kappa_{d\,i} - \delta \\ \epsilon_{c\,i} > -\kappa_{c\,i} - \delta \\ \epsilon_{c\,i} > -\kappa_{c\,i} - \kappa_{d\,i} - \delta - \epsilon_{d\,i}\ ^{\text{a}} \end{cases}$	$S_i(1,0): \begin{cases} \epsilon_{d\,i} > -\kappa_{d\,i} \\ \epsilon_{c\,i} < -\kappa_{c\,i} - \delta \\ \epsilon_{c\,i} < -\kappa_{c\,i} + \kappa_{d\,i} + \epsilon_{d\,i}\ ^{\text{b}} \end{cases}$
$S_i(0,1): \begin{cases} \epsilon_{d\,i} < -\kappa_{d\,i} - \delta \\ \epsilon_{c\,i} > -\kappa_{c\,i} \\ \epsilon_{c\,i} > -\kappa_{c\,i} + \kappa_{d\,i} + \epsilon_{d\,i}\ ^{\text{b}} \end{cases}$	$S_i(0,0): \begin{cases} \epsilon_{d\,i} < -\kappa_{d\,i} \\ \epsilon_{c\,i} < -\kappa_{c\,i} \\ \epsilon_{c\,i} < -\kappa_{c\,i} - \kappa_{d\,i} - \delta - \epsilon_{d\,i}\ ^{\text{a}} \end{cases}$

These are the conditions that simultaneoslly fulfill all points pertaining to $S_i(d,c)$, the set of pairs $(\epsilon_{d\,i}, \epsilon_{c\,i})$ for which the optimal decision on innovative profile is $(x_{d\,i}, x_{c\,i}) = (d,c)$.

[a] This condition is not binding when $\delta \leq 0$.

[b] This condition is not binding when $\delta \geq 0$.

the following three conditions are fulfilled:

$$\pi(1,1) > \pi(0,1) \implies \kappa_{d\,i} + \epsilon_{d\,i} + \kappa_{c\,i} + \epsilon_{c\,i} + \delta + \pi_{0\,i} > \kappa_{c\,i} + \epsilon_{c\,i} + \pi_{0\,i}, \tag{15a}$$

$$\pi(1,1) > \pi(1,0) \implies \kappa_{d\,i} + \epsilon_{d\,i} + \kappa_{c\,i} + \epsilon_{c\,i} + \delta + \pi_{0\,i} > \kappa_{d\,i} + \epsilon_{d\,i} + \pi_{0\,i}, \tag{15b}$$

$$\pi(1,1) > \pi(0,0) \implies \kappa_{d\,i} + \epsilon_{d\,i} + \kappa_{c\,i} + \epsilon_{c\,i} + \delta + \pi_{0\,i} > \pi_{0\,i}. \tag{15c}$$

For convenience, we denote as $S_i(1,1)$ the subset of realization of errors $(\epsilon_{d\,i}, \epsilon_{c\,i})$ leading firm i to jointly adopt innovation strategies $x_{d\,i} = 1$ and $x_{c\,i} = 1$. After simplifying (15a)–(15c), $S_i(1,1)$ comprises all values of $(\epsilon_{d\,i}, \epsilon_{c\,i})$ that simultaneously fulfill the following three conditions:

$$\epsilon_{d\,i} > -\kappa_{d\,i} - \delta \tag{16a}$$

$$\epsilon_{c\,i} > -\kappa_{c\,i} - \delta \tag{16b}$$

$$\epsilon_{c\,i} + \epsilon_{d\,i} > -\kappa_{d\,i} - \kappa_{c\,i} - \delta \tag{16c}$$

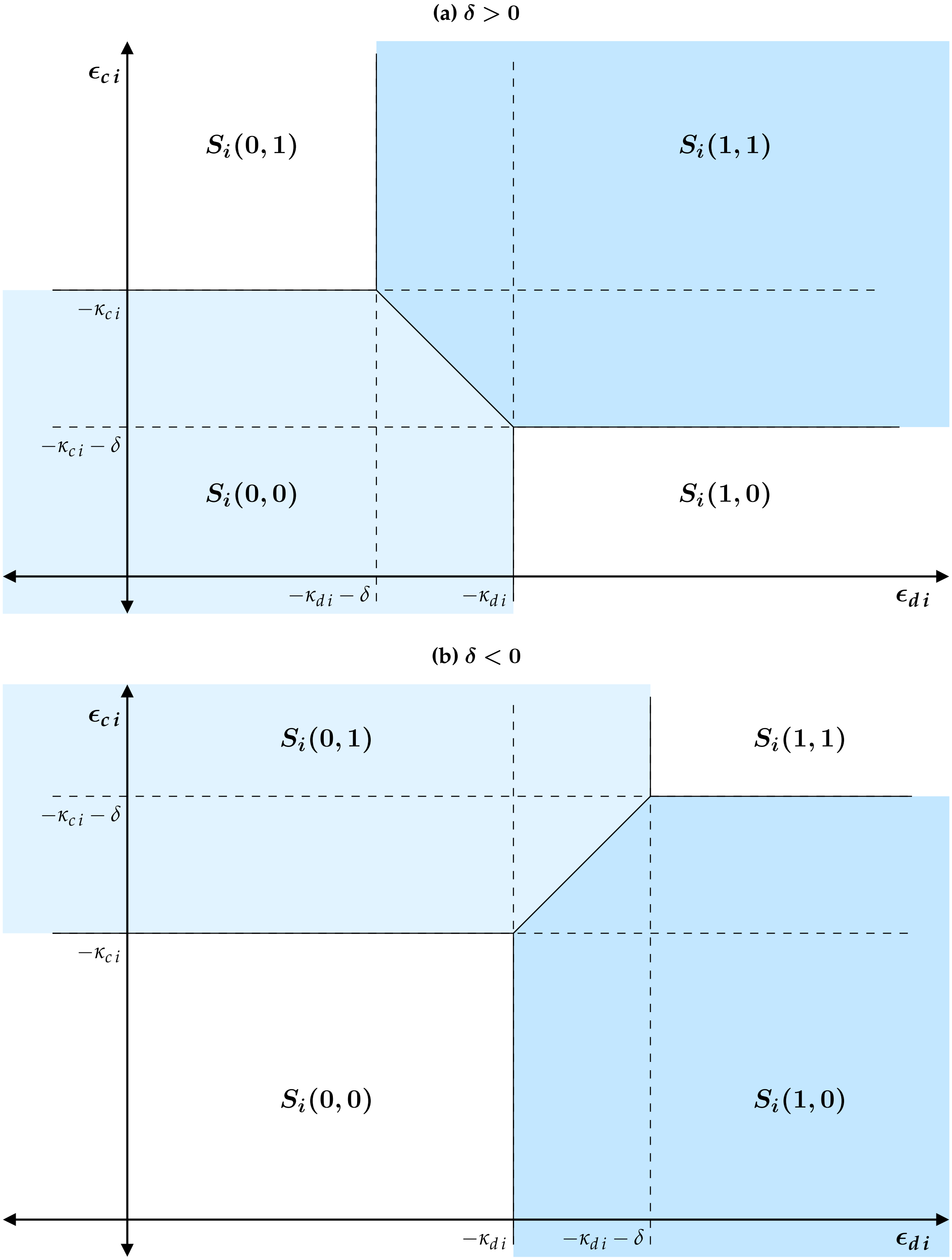
It is straightforward to show that the third condition is not binding when $\delta \leq 0$. Similar conditions for every strategy profile are presented in Table 4, which implicitly define $S_i(x_{d\,i}, x_{c\,i})$ for all possible values of $x_{d\,i}$ and $x_{c\,i}$.

Figure 1a and Figure 1b show the support of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ for a given realization of $\epsilon_{y\,i}$, and, by means of the sets $S_i(x_{d\,i}, x_{c\,i})$ of Table 4, the optimal choices of innovative profile induced by every combination of unobservables $(\epsilon_{d\,i}, \epsilon_{c\,i})$. Two comments are worth being pointed out here. First, there is no overlap among the different sets $S_i(x_{d\,i}, x_{c\,i})$, and thus, our model does not suffers from incoherence problems as those discussed in Appendix A. Second, the $S_i(x_{d\,i}, x_{c\,i})$ sets are not in general (displaced) quadrants and the sign of δ determines the shapes of these sets. Only when $\delta = 0$ all four sets $S_i(x_{d\,i}, x_{c\,i})$ have the simple form of quadrants. We have therefore to control for this latter feature of the problem when computing the likelihood function of each observation.

Therefore, our model is characterized by a linear equation for the continuous variable $x_{y\,i}$ and a set of inequalities from which we can infer bounds for $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ from any given observed choice on $x_{d\,i}$ and $x_{c\,i}$. With this information in hand, we can write the contribution to the likelihood funcion of observation i as:

$$L_i(x_{d\,i}, x_{c\,i}, x_{y\,i}) = f(\epsilon_{y\,i}) \Pr(x_{d\,i}, x_{c\,i} | \epsilon_{y\,i}). \tag{17}$$

Figure 1: Innovation Profile Defining Regions



where $f(\cdot)$ is the probability density function of $\epsilon_{y\,i}$. Thus, analogously to other econometric models like the family of generalized tobit models — see for instance Amemiya (1985)—, the probabilistic structure of our model mixes a continuous density with a discrete probability. This latter probability is evaluated integrating $g(\cdot)$, the joint density of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ conditional on $\epsilon_{y\,i}$, over the corresponding $S_i(x_{d\,i}, x_{c\,i})$ region. Thus:

$$L_i(x_{d\,i}, x_{c\,i}, x_{y\,i}) = f(\epsilon_{y\,i}) \iint_{S_i(x_{d\,i}, x_{c\,i})} g(\epsilon_{d\,i}, \epsilon_{c\,i} | \epsilon_{y\,i}) d\epsilon_{c\,i} d\epsilon_{d\,i}. \tag{18}$$

The results of this section together with a convenient distributional assumption on the elements of $\boldsymbol{\epsilon}_i$ enables us to perform the estimation of parameters of interest based on the likelihood function sketched in (18). Appendix B presents a detailed derivation of the likelihood function under the assumption of normally distributed $\boldsymbol{\epsilon}_i$.

4 The Nature of Complementarities in the Ceramic Tiles Industry

The parameters that capture the *intrinsic* complementarity, δ_{dc}, δ_{dy}, and δ_{cy} are identified through the cross products of the decision variables on the profit function (7). Regarding the determinants of the *induced* complementarity, they are the unobserved elements of the environments of the firms as defined in (9d)–(9f). Evidently, as they capture organizational features, managerial ability, and/or other issues not known to the econometrician, we can only determine the most likely joint distribution of such effects. So, in addition to the parameters of profits function, we have to estimate the parameters of a general normal distribution of $\epsilon = (\epsilon_{d\,i}, \epsilon_{c\,i}, \epsilon_{y\,i})'$. These parameters includes the correlation coefficients ρ_{dc}, ρ_{dy}, and ρ_{cy} that measure the *induced* complementarities, and the standard deviation of $\epsilon_{y\,i}$, denoted as σ_y.[8]

The remaining basic elements of our model are given by the observable environmental variables of equations (9a)–(9c). Functions $\theta_d(\boldsymbol{z}_{r\,i}, \boldsymbol{z}_{k\,i})$, $\theta_c(\boldsymbol{z}_{c\,i}, \boldsymbol{z}_{k\,i})$, and $\theta_y(\boldsymbol{z}_{r\,i}, \boldsymbol{z}_{c\,i})$ include the combined effects of the environmental variables that affect revenues, $\boldsymbol{z}_{r\,i}$; cost of production, $\boldsymbol{z}_{c\,i}$; and cost of adoption $\boldsymbol{z}_{k\,i}$. Which variables are included in each one of these categories is defined by our behavioral model of production and innovation. Our selection of variables is not unlimited and the following paragraphs describe what we think is a reasonable behavioral model for the Spanish ceramic tiles industry.[9]

[8]In practice we normalize the standard deviations of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ as $\sigma_d = \sigma_c = 1$. The reason why we have to assume that the marginal distribution of those errors have the same dispersion has to do with local identification problems. Whenever $\delta = 0$ our model reduces to the estimation of a linear equation plus a bivariate probit. In such a case neither of the two standard deviations of the errors are identified. In principle, whenever $\delta \neq 0$, one of the two standard deviations could be identified. However, we cannot rule out *ex ante* the possibility that $\delta = 0$ and we encountered severe difficulties to achieve convergence as soon as our iterations took us in the surrounding of $\delta = 0$.

[9]Actually, many of our regressors could be considered endogenous. Whether a firm exits the market, exports to a particular region, owns one or more brands, and produces one or several products, all are in the end decisions of the firms. There is little else that we can do since there are no additional instruments available to us. In this paper we want mostly to stress the validity of our estimation method and its applicability to better data sets. To justify our approach, we argue that exit is not a very frequent event in this sample. Similarly, all the other potentially endogenous variables can easily be considered predetermined, at least in the short run. We will assume that they are at least weakly exogenous as marketing new brands, increasing the number of production lines, and gaining access to foreign markets requires time, resources, and managerial effort. In addition to these arguments, it is possible to derive orthogonality restrictions from the likelihood function described in detail in Appendix B. Then, if additional valid exogenous instruments were available, we could implement a generalized method of moments estimator robust to the possible endogeneity of the elements in $\boldsymbol{Z}_i$.

There is first a set of common variables affecting revenues, costs of production, and costs of adoption. Besides a CONSTANT capturing the average level of revenues, production costs, and adoption costs, we allow for a TIME effect over the three environmental functions. The time dummy may capture dynamic effects on demand, as new products gain access to new distribution channels, firms build up their reputation, or consumers learn about the new varieties of ceramic tiles and their improved quality. Similarly, as time goes by, firms may be able to reduce their unit costs as they gain experience in using the new technology, or because of a downwards trend in the costs of inputs. It is expected that the cost of adoption also becomes less important with time due to reduction in the actual cost of the single firing furnace, as well as for the improved local knowledge of technicians and engineers (the ceramic tiles industry is clustered in a small region on the east coast of Spain).

We also include among the common variables two dummies to control for ENTRY into or EXIT of firms from the sample. These variables may capture the differentiated environment and behavior of recent startups and of declining firms. Overall, the Spanish ceramic tiles is not a declining industry, which helps us avoid dealing with endogenous exit of firms.

Among the environmental variables affecting exclusively to revenues we include EX, the percent of production that is exported; whether the largest foreign market was the European Union, EU; and wether the firm owns at least one registered trademark, TM.[10] Most ceramic tiles firms concentrate their sales in the domestic market. The European market provides with higher revenues, but also demands, in general, higher quality products. Those Spanish ceramic tiles firms who are present in Europe are also among the largest, and more innovative, and we expect that all these effects affect positively to their revenues. Trademarks are the common way to identify a producer with the quality of their products. We interpret this indicator as the valuation of goodwill and reputation, and again, we expect it to be associated to positive revenue shifts.

The only variable associated exclusively to the cost of production is AGE. This variable captures the potential learning by doing effects of experience. The number of years that a firm has been active in the industry works as a signal related to the accumulated output that will eventually be responsible of potential unit cost reductions.

Finally, the number of varieties produced by the firms may affect the costs of adopting innovations. Its effect is however ambiguous. The dummy MPROD identifies those firms that produce more than one product (about 39% of the sample).[11] Costs of adopting a new innovation might presumably be higher if they have to be integrated with the joint production of several products. Coordination and organizational problems may then arise. On the contrary, by simplifying matters, or by reducing the fixed costs common to the different production lines, multiproduct firms may enjoy a significant cost savings if adopting both product and process innovations.

Table 5 reports the maximum likelihood estimates of our model. The revenue, production costs, and adoption cost specific argument enter the return functions $\theta_d(z_{r\,i}, z_{k\,i})$, $\theta_c(z_{c\,i}, z_{k\,i})$, and $\theta_y(z_{r\,i}, z_{c\,i})$ as indicated in equations (9a)–(9c). Therefore, and according to (7) the estimates

[10] We have encountered nonlinearities among these regressors. After trying several combinations and alternative definitions for some of these dummies, we included EX · EU and TMHI, an indicator that firms produces at least two products, among the observable environmental variables affecting revenues.

[11] Here we have also encountered significant nonlinearities, and again after several attempts, we decided to include MPRODHI to identify those firms that produce at least three products (about 4% of the sample).

Table 5: Maximum Likelihood Estimates

		Model [I]	Model [II]	Model [III]	Model [IV]
θ_d	CONSTANT	−0.64 (0.16)***	−0.63 (0.16)***	−0.51 (0.21)**	0.90 (0.67)
	EX	−0.21 (0.35)	−0.29 (0.32)	−0.17 (0.35)	0.34 (0.36)
	EX · EU	0.88 (0.36)**	0.97 (0.33)***	0.99 (0.36)***	0.43 (0.53)
	TM	0.35 (0.15)**	0.36 (0.14)***	0.43 (0.15)***	0.39 (0.19)**
	TMHI	−0.03 (0.16)	−0.05 (0.15)	0.04 (0.17)	0.37 (0.18)**
	MPROD	0.07 (0.14)	0.07 (0.13)	0.08 (0.14)	0.03 (0.12)
	MPRODHI	0.38 (0.34)	0.43 (0.34)	0.20 (0.34)	0.69 (0.40)*
	TIME	−0.16 (0.14)	−0.17 (0.14)	−0.13 (0.14)	−0.08 (0.13)
	EXIT	−0.36 (0.27)	−0.35 (0.27)	−0.41 (0.28)	−0.62 (0.26)**
	ENTRY	0.33 (0.30)	0.33 (0.30)	0.24 (0.31)	0.01 (0.32)
θ_c	CONSTANT	−0.24 (0.28)	−0.35 (0.26)	−0.49 (0.30)	−0.56 (0.41)
	AGE	−0.03 (0.10)	0.01 (0.09)	−0.03 (0.10)	−0.20 (0.11)*
	MPROD	−0.02 (0.13)	−0.02 (0.13)	−0.03 (0.13)	−0.02 (0.11)
	MPRODHI	1.54 (0.41)***	1.59 (0.42)***	1.58 (0.43)***	1.30 (0.47)***
	TIME	−0.11 (0.14)	−0.12 (0.14)	−0.08 (0.14)	−0.13 (0.12)
	EXIT	−0.30 (0.27)	−0.29 (0.27)	−0.25 (0.27)	−0.12 (0.26)
	ENTRY	0.09 (0.32)	0.15 (0.32)	0.05 (0.33)	0.15 (0.29)
θ_y	CONSTANT	3.26 (0.38)***	3.33 (0.38)***	3.30 (0.38)***	3.29 (0.39)***
	EX	1.02 (0.44)**	1.04 (0.45)**	1.01 (0.44)**	1.05 (0.42)**
	EX · EU	−0.10 (0.48)	−0.11 (0.48)	−0.07 (0.48)	0.04 (0.48)
	TM	0.44 (0.19)**	0.44 (0.19)**	0.45 (0.19)**	0.48 (0.19)**
	TMHI	0.97 (0.22)***	0.97 (0.22)***	0.97 (0.22)***	0.90 (0.23)***
	AGE	0.52 (0.13)***	0.49 (0.13)***	0.52 (0.13)***	0.51 (0.13)***
	TIME	0.13 (0.19)	0.14 (0.19)	0.13 (0.19)	0.12 (0.19)
	EXIT	−1.00 (0.35)***	−1.01 (0.35)***	−1.01 (0.35)***	−1.02 (0.35)***
	ENTRY	−0.40 (0.44)	−0.44 (0.44)	−0.40 (0.44)	−0.39 (0.44)
δ_{dc}				0.52 (0.08)***	−0.50 (0.36)
δ_{dy}				−0.08 (0.03)***	−0.27 (0.14)**
δ_{cy}				0.01 (0.03)	0.19 (0.10)*
ρ_{dc}			0.55 (0.06)***		0.64 (0.26)**
ρ_{dy}			−0.15 (0.06)**		0.40 (0.29)
ρ_{cy}			−0.04 (0.06)		−0.40 (0.20)**
σ_y		1.75 (0.06)***	1.75 (0.06)***	1.75 (0.06)***	1.74 (0.06)***
$\ln L$		−1396.3	−1367.8	−1367.9	−1364.5
χ^2		132.7***	137.2***	136.3***	121.2***

Maximum likelihood estimates and their standard errors in parentheses. Estimates found different from zero at significance levels 10%, 5% and 1% are marked with ***, **, and *, respectively. $\ln L$ is the value of the likelihood function at the maximum. Last row shows Wald tests for the joint significance of the slopes of $\theta_{d\,i}$, $\theta_{c\,i}$, and $\theta_{y\,i}$. These tests are asymptotically distributed as a χ^2 with 23 degrees of freedom.

Table 6: Model Selection Tests

H_0	H_1	d. f.	Test	p-value
Model [I]	Model [II]	3	57.02	0.000
Model [I]	Model [III]	3	56.94	0.000
Model [I]	Model [IV]	6	63.59	0.000
Model [II]	Model [IV]	3	6.57	0.087
Model [III]	Model [IV]	3	6.65	0.084
Model [II]	Model [III]		0.04	0.972

All rows, except the last one, reports likelihood ratio tests with null and alternative hypotheses as indicated in columns 'H_0' and 'H_1'. The asymptotic distribution of these tests is a χ^2 with 'd. f.' degrees of freedom. The last row reports a Vuong test of non-nested hypotheses to compare Models [II] and [III] which is distributed as standard normal under the null hypothesis of equivalence of both models. Significant positive values of this test favor Model [II] while significant negative values favor Model [III].

of the observable environmental variables have to be interpreted as increasing o reducing the direct returns to engaging in product innovation, process innovation, and setting the scale of production.

Table 5 presents four different specifications of the model of production decision and adoption of innovations. Model [I], which is primarily intended for testing purpouses, assumes that there is no complementarity at all, neither *intrinsic* or *induced*. Model [II] considers that all complementarity among strategies, if any, has its origin on organizational features or other unobserved characteristics of firms. This is the type of model commonly estimated —*e.g.*, Arora and Gambardella (1990), Mohnen and Röller (2003), Kaiser (2003), Cassiman and Veugelers (2002)— where the choice of innovation strategies is studied in isolation of each other but complementarity is studied by analyzing the correlation across error terms of each choice equation. As we discussed before, such approach lead to inconsintent estimates in the presence of *intrinsic* complementarity. Model [III], on the contrary, assumes that there is no unobserved firm heterogeneity, and consequently, all existing complementarity will be of *intrinsic* nature only, and thus complementarity will not have its origin on unobserved organizational features of the firm or ability and experience of management. Finally, model [IV] is the general model discussed above and it allows for both *intrinsic* and *induced* complementarity.

At this time, it is worth turning our attention to the likelihood ratio tests presented in Table 6. Regardless of the alternative hypothesis considered, Model [I] is always rejected in favor of any other model that allows for the possibility of complementarity of any kind. So, complementarity among strategies is a relevant issue in our data set. Last line of Table 6 reports a Vuong (1989) test of non-nested hypotheses to compare models [II] and [III].[12] This test reveals that these two models give essentially equivalent explanations of firms behaviour in our sample, so we cannot distinguish between models with a unique source of complementarity. In addition, model [IV], which includes different types of complementarities, is always preferred (at a 10% significance level) to models [II] or [III] that only contemplate one single source to explain the

[12]Models [II] and [III] are an example of the kind of overlapping models discussed in Vuong (1989, §6). The asymptotic distribution of tests comparing overlapping models depends on a variance term being zero or not. An equivalent procedure to the variance tests of Vuong (1989) is the standard likelihood ratio test comparing Model [I] versus Model [IV]. As this test strongly rejects the null hypothesis, we can rely on the normal asymptotic distribution of the test in last row of Table 6.

complementarity among firms' strategies. We will therefore focus our comments in the most general specification of our production and innovation model.

It is worth making a couple of general remarks at this point. First, in addition for *intrinsic* or *induced* complementarity, we condition the analysis of the different innovation profiles by a set of observable characteristics of firms. The last row of Table 5 reports a Wald test of joint significance of all these characteristics of firms. The alternative model would include constant θ_d, θ_c, and θ_y. Such a model is always rejected, regardless of the assumed working hypothesis about the nature of complementarities across strategies. Returns of innovations and production decisions vary across firms, and the estimation improves substantially if we control for observable characteristics of firms.

The second remark is referred to the precision of the estimates. Model [IV] is less precise than any of the other models and this is due to the difficulty to distinguish the source of complementarity. Some parameters are only significant under restrictive complementarity relationships. However, in the presence of both *intrinsic* and *induced* complementarity, the estimates of environmental variables are no longer consistent in models [I]–[III].

We first turn our attention to the determinants of the returns of product innovation. Trademarks have a quite significant effect on the return of product innovation, θ_d. This result is intuitive as trademarks ease the way for firms to appropriate the profits of their innovations. We also document that the effect of trademarks on the return of product innovation increases more than proportionally with the number of registered trademarks. Similarly, returns to demand innovation are higher when firms offer many products. This can be the result of reputation or consumer learning spillovers across different products of a firm. An increase in the quality of one of the products of a firm may lead consumers to purchase some of the other varieties offered by the firm. Only firms that exit the market have a significantly lower return to engage in demand enhancing innovations.

Only two variables among our regressors appear to have significant effects on the returns to adopt process innovations. The older firms get, the less likely they innovate in cost reducing innovations. The single firing furnace of the ceramic tile industry represents a major innovation that requires the production plant to be completely redesigned. It is in many cases more efficient to build a brand new plant than have an old one remodelled. Thus, firms that entered during the early 1980s were more likely to have adopted such major innovation by the end of the decade. Process innovation is also more profitable for multiproduct firms. This together with the high return of product innovation for multiproduct firms is consistent with the existence of economies of scope in the ceramic tile industry.

Firms with access to foreign markets increase significantly the returns to a larger scale of production. This is also the case of firms with several registered trademarks. As firms can appropriate the benefits of their innovations, they take advantage by expanding production. The profitability of large production also increases with time, as firms get established, and only decline before firms leave the market.

Perhaps the most interesting results are those involving the complementarity of strategies. We should notice that when we restrict the model to include exclusively either *intrinsic* or *induced* complementarities, the estimates wrongly pick the effect of the excluded source of complementarity. This can be seen comparing the estimates of δ_{dc}, δ_{dy}, δ_{cy}, and those of ρ_{dc}, ρ_{dy}, ρ_{cy} across models [II], [III], and [IV]. This latter specification reveals, for instance, that the product and process innovation complementarity already documented in Table 3 has its origin on unobserved firm heterogeneity. This is consistent with simultaneous innovation in product and process being the result of organizational features of the firm difficult to account for, such as the experience, background, and ability of managers that may realize of the high profitability of developing a combined set of strategies simultaneously.

Smaller firms obtain a larger return of adopting product innovations. This is mostly a technological relationship. The single firing furnace reduced effectively the minimum efficient scale of ceramic tiles firms, thus allowing that smaller firms engaged in design of new products. When comparing models [II] and [IV] we observe that the inconsistent estimate of a negative *induced* complementarity in model [II] is only due to the fact that we are excluding the possibility of *intrinsic* complementarities.

In models [II] and [III], scale and process do not appear to be related at all. This is again the result of a misspecified model based on a restricted structure of complementarity relationship. In model [IV], scale and process innovation show significant complementarity relationships of opposite sign depending of their nature. Returns to innovation are higher for larger firms based on technological features of the production process (*intrinsic* complementarity). Larger firms may benefit the most from employing the new single firing furnace and using all the increased capacity of production that such innovation brings to the firm. The negative effect of the *induced* complementarity relationship could perhaps be explained by the lack of technical personnel, access to markets, and manager backgrounds of small firms, which deters them (actually most of the firms in the sample) to adopt the single firing furnace even though it is designed to achieve its minimum efficient scale at a low level of production.

To conclude, the estimation shows a rich pattern of complementarity among strategies distinguishing by their nature and origin. These estimates conform the well established facts of the Spanish ceramic tile industry and point out the potential importance of organizational issues in the delay of adopting innovations. Among the policy recommendations we can think of management training, fostering of mergers and consolidation of the industry, and/or the development of specialized technological institutes has potential ways to help spreading knowledge about the new technologies and thus prompt more firms to adopt this major innovation so that they can compete successfully in international markets where the return of the investment appears to be higher. Technological institutes and research centers sponsored by ceramic tiles firms or local governments were actually developed during the 1990s to pool resources across firms in the development of new materials.

5 Concluding Remarks

This paper has introduced and estimated an econometric model that can distinguish the nature of the complementarity relationship among the different strategies of a firm. Our estimates of the Spanish ceramic tiles industry show that an econometric model that allows for complementarities among production, product, and process innovation is always preferred to one where firms strategies are independent and all heterogeneity is known to the econometrician.

Our results show that there is significant complementarity between product and process innovation, and that this is mostly due to unobserved heterogeneity. This opens the door to interpretations where the organizational form of firms and/or the experience and ability of managers become the key element to coordinate and take advantage of the innovation possibilities offered by technology. Smaller firms appear to be more inclined to innovate. This is the result of technology in the case of demand innovations (*intrinsic* complementarity) but of organizational matters (*induced* complementarity) are more important in the case of process innovation.

Our model is a first step in the direction of evaluating the importance and origin of innovation complementarities. We envision at least two ways in which the present framework could be extended. First, we could add a temporal dimension to the analysis of complementarities. A panel data of this model will be able to identify the existence of dynamic complementarities through the identification of state dependence of the sequence of realized types of innovations. Second, we could consider contemplating several additional strategies. If these additional strategies were continuous, the extension of the model would be straightforward. If, alternatively, we considered dichotomous strategies, the definition of $S_i(x_{d\,i}, x_{c\,i})$, the regions of realizations of shocks associated to each innovation profile becomes much harder to delimit and we will have to resort to simulator estimators such as those of McFadden (1989) and Pakes and Pollard (1989).

References

ABERNATHY, W. J. AND J. M. UTTERBACK (1978): "Patterns of Industrial Innovation." *Technology Review*, 80, 41–47.

AMEMIYA, T. (1985): *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

ARORA, A. AND A. GAMBARDELLA (1990): "Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology." *The Journal of Industrial Economics*, 38, 361–379.

ATHEY, S. AND A. SCHMUTZLER (1995): "Product and Process Flexibility in an Innovative Environment." *Rand Journal of Economics*, 26, 557–574.

ATHEY, S. AND S. STERN (1998): "An Empirical Framework for Testing Theories About Complementarity in Organizational Design." Working Paper 6600, NBER.

CASSIMAN, B. AND R. VEUGELERS (2002): "Complementarity in the Innovation Startegy: Internal R & D, External Technology Acquisition, and Cooperation in R & D." Mimeo, DTEW, Katholieke Universiteit Leuven.

DOORNIK, J. A. (2002): *Object-Oriented Matrix Programming using Ox*, 3rd edn. London: Timberlake Consultants Press.

HÖLMSTROM, B. AND P. MILGROM (1994): "The Firm as an Incentive System." *American Economic Review*, 84, 972–991.

ICHNIOWSKI, C., K. SHAW, AND G. PRENNUSHI (1997): "The Effects of Human Resource Management Practices on Productivity." *American Economic Review*, 87, 291–313.

KAISER, U. (2003): "Startegic Complementarities Between Different Types of ICT-Expenditures." Discussion Paper 2003-05, CEBR.

KLEEPER, S. (1996): "Entry, Exit, Growth, and Innovation over the Product Life Cycle." *American Economic Review*, 86, 562–583.

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica*, 57, 995–1026.

MILGROM, P. AND J. ROBERTS (1990): "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economic Review*, 80, 511–528.

MILGROM, P. AND J. ROBERTS (1995): "Complementarities and Fit: Strategy, Structure and Organizational Change in Manufacturing." *Journal of Accounting and Economics*, 19, 179–208.

MOHNEN, P. AND L.-H. RÖLLER (2003): "Complementarities in Innovation Policy." Discussion Paper 2003s-60, CIRANO.

PAKES, A. AND D. POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators." *Econometrica*, 57, 1027–1058.

PRESS, W., B. FLANNERY, S. TEULOSKY, AND W. VETTERLING (1986): *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. Cambridge, MA: Cambridge University Press.

SCHMIDT, P. (1981): "Constraints on the Parameters in Simultaneous Tobit and Probit Models." In C. F. Manski and D. L. McFadden (eds.): *Structural Analysis of Discrete Data and Econometric Applications*. Cambridge, MA: MIT Press.

STROUD, A. H. AND D. SECREST (1966): *Gaussian Quadrature Formulas*. Englewood Cliffs, NJ: Prentice-Hall.

TOPKIS, D. M. (1998): *Supermodularity and Complementarity*. Princeton, NJ: Princeton University Press.

VUONG, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica*, 57, 307–333.

Appendices

A Incoherence of Estimating the Determinants of Innovation Strategies in Isolation

Our estimation approach identify the parameters of the model by assuming that firms maximize profits when they choose a particular combination of innovation strategies. Given the continuous nature of the scale of production, $x_{y\,i}$, we can make use of the standard first order condition that determines the optimal choice of production level:

$$\frac{\partial \pi}{\partial x_{y\,i}} = \theta_{y\,i} + \delta_{dy} x_{d\,i} + \delta_{cy} x_{c\,i} + \epsilon_{y\,i} - x_{y\,i} = 0. \tag{A.1}$$

From here we can obtain the optimal scale of production conditional on the innovation profile of the firm:

$$x_{y\,i} = \theta_{y\,i} + \delta_{dy}x_{d\,i} + \delta_{cy}x_{c\,i} + \epsilon_{y\,i}. \tag{A.2}$$

In addition, firms will adopt a particular innovation if such decision leads to an increase in profits. Let $x^{\star}_{d\,i}$ denote the unobservable increase in profits related to the adoption of a product innovation:

$$x^{\star}_{d\,i} = \pi(1, x_{c\,i}, x_{y\,i}) - \pi(0, x_{c\,i}, x_{y\,i}). \tag{A.3}$$

Similarly $x^{\star}_{c\,i}$ represents the unobservable return of adopting a process innovation:

$$x^{\star}_{c\,i} = \pi(x_{d\,i}, 1, x_{y\,i}) - \pi(x_{d\,i}, 0, x_{y\,i}). \tag{A.4}$$

From the profit function (8), we get:

$$x^{\star}_{d\,i} = \theta_{d\,i} + \delta_{dc}x_{c\,i} + \delta_{dy}x_{y\,i} + \epsilon_{d\,i}, \tag{A.5}$$

and similarly:

$$x^{\star}_{c\,i} = \theta_{c\,i} + \delta_{dc}x_{d\,i} + \delta_{cy}x_{y\,i} + \epsilon_{c\,i}. \tag{A.6}$$

Next, we define the innovation adoption indicators as a function of whether firms realize positive profits if they engage in each innovation strategy:

$$x_{j\,i} = \begin{cases} 1 & \text{si } x^{\star}_{j\,i} > 0, \\ 0 & \text{si } x^{\star}_{j\,i} \leq 0, \end{cases} \qquad (j = d, c). \tag{A.7}$$

In these expressions, optimal firm choices are indeed conditional on values of $\boldsymbol{z}_{r\,i}$, $\boldsymbol{z}_{c\,i}$, and $\boldsymbol{z}_{k\,i}$ as defined by $\theta_{d\,i}$, $\theta_{c\,i}$, and $\theta_{y\,i}$ in equations (9a)–(9c). These relations can be expressed in matrix form as:

$$\boldsymbol{\theta}_i = \boldsymbol{\Theta}\boldsymbol{z}_i, \tag{A.8}$$

where $\boldsymbol{\theta}_i = (\theta_{d\,i}, \theta_{c\,i}, \theta_{y\,i})'$, $\boldsymbol{z}_i = (1, \boldsymbol{z}'_{r\,i}, \boldsymbol{z}'_{c\,i}, \boldsymbol{z}'_{k\,i})'$ and:

$$\boldsymbol{\Theta} = \begin{bmatrix} \theta_{d0} & \boldsymbol{\theta}'_{dr} & \boldsymbol{0} & \boldsymbol{\theta}'_{dk} \\ \theta_{c0} & \boldsymbol{0} & \boldsymbol{\theta}'_{cc} & \boldsymbol{\theta}'_{ck} \\ \theta_{y0} & \boldsymbol{\theta}'_{yr} & \boldsymbol{\theta}'_{yc} & \boldsymbol{0} \end{bmatrix}. \tag{A.9}$$

Thus, the optimal innovation rules of firms, (A.2), (A.5), and (A.6) can be written in matrix form as:

$$\boldsymbol{x}^{*}_{i} = \boldsymbol{\Theta}\boldsymbol{z}_i + \boldsymbol{\Gamma}\boldsymbol{x}_i + \boldsymbol{\epsilon}_i, \tag{A.10}$$

where $\boldsymbol{x}^{*}_{i} = (x^{\star}_{d\,i}, x^{\star}_{c\,i}, x_{y\,i})'$, and:

$$\boldsymbol{\Gamma} = \begin{bmatrix} 0 & \delta_{dc} & \delta_{dy} \\ \delta_{dc} & 0 & \delta_{cy} \\ \delta_{dy} & \delta_{cy} & 0 \end{bmatrix}. \tag{A.11}$$

Equation (A.10) together with the observation rules (A.7), conform a model of simultaneous equations where two out of three endogenous variables, $x^{\star}_{d\,i}$ and $x^{\star}_{c\,i}$, are only partially observable. Notice that the right hand side of (A.10) only includes the observable indicators, $x_{d\,i}$ and $x_{c\,i}$.

Schmidt (1981) discusses specific restrictions that the parameters of this class of models need to fulfil. In particular, it is required some minimum degree of recursion so that for any vector $\boldsymbol{z}_i$, the realization of errors $\boldsymbol{\epsilon}_i$ uniquely determine the choice of endogenous variables. Let partition matrix $\boldsymbol{\Gamma}$ as follows:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{bmatrix} = \left[\begin{array}{cc|c} 0 & \delta_{dc} & \delta_{dy} \\ \delta_{dc} & 0 & \delta_{cy} \\ \hline \delta_{dy} & \delta_{cy} & 0 \end{array}\right]. \tag{A.12}$$

The model coherence requires that the principal minors of matrix $\boldsymbol{A} = \boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{12}(\mathbf{I} - \boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}$ be equal to zero (Schmidt, 1981, §8.2). In our case matrix $\boldsymbol{A}$ is:

$$\boldsymbol{A} = \begin{bmatrix} \delta_{dy}^2 & \delta_{dc} + \delta_{dy}\delta_{cy} \\ \delta_{dc} + \delta_{dy}\delta_{cy} & \delta_{cy}^2 \end{bmatrix}. \tag{A.13}$$

It is straightforward to show that the principal minors of this matrix can only be zero if parameters δ_{dc}, δ_{dy}, and δ_{cy} are also zero. Using the approach summarized by equations (A.10) and (A.7), where innovations are considered in isolation, it is not possible to address the case of *intrinsic* complementarity.

We can illustrate the meaning of the coherence problem by substituting the definitions of $x^\star_{d\,i}$ y $x^\star_{c\,i}$ given by (A.5) and (A.6), into the observability rules (A.7), so that:

$$x_{d\,i} = \begin{cases} 1 & \text{si } \epsilon_{d\,i} > -\theta_{d\,i} - \delta_{dc}x_{c\,i} - \delta_{dy}x_{y\,i}, \\ 0 & \text{si } \epsilon_{d\,i} \leq -\theta_{d\,i} - \delta_{dc}x_{c\,i} - \delta_{dy}x_{y\,i}, \end{cases} \tag{A.14}$$

and:

$$x_{c\,i} = \begin{cases} 1 & \text{si } \epsilon_{c\,i} > -\theta_{d\,i} - \delta_{dc}x_{d\,i} - \delta_{cy}x_{y\,i}, \\ 0 & \text{si } \epsilon_{c\,i} \leq -\theta_{d\,i} - \delta_{dc}x_{d\,i} - \delta_{cy}x_{y\,i}. \end{cases} \tag{A.15}$$

Next, let define:

$$S_i(1,1) = \big\{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \arg\max \pi_i(x_{d\,i}, x_{c\,i}) = (1,1)\big\} \tag{A.16}$$

as the set of values of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ that induce firm i to optimally adopt both innovations simultaneously. The combinations of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ leading to the remaining innovation profiles, $S_i(1,0)$, $S_i(0,1)$, and $S_i(0,0)$ can be defined in a similar manner to equation (A.16). Making use of (A.14) and (A.15), we obtain:

$$S_i(1,1) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} > -\theta_{d\,i} - \delta_{dc} - \delta_{dy}x_{y\,i},\ \epsilon_{c\,i} > -\theta_{c\,i} - \delta_{dc} - \delta_{cy}x_{y\,i}\}, \tag{A.17}$$

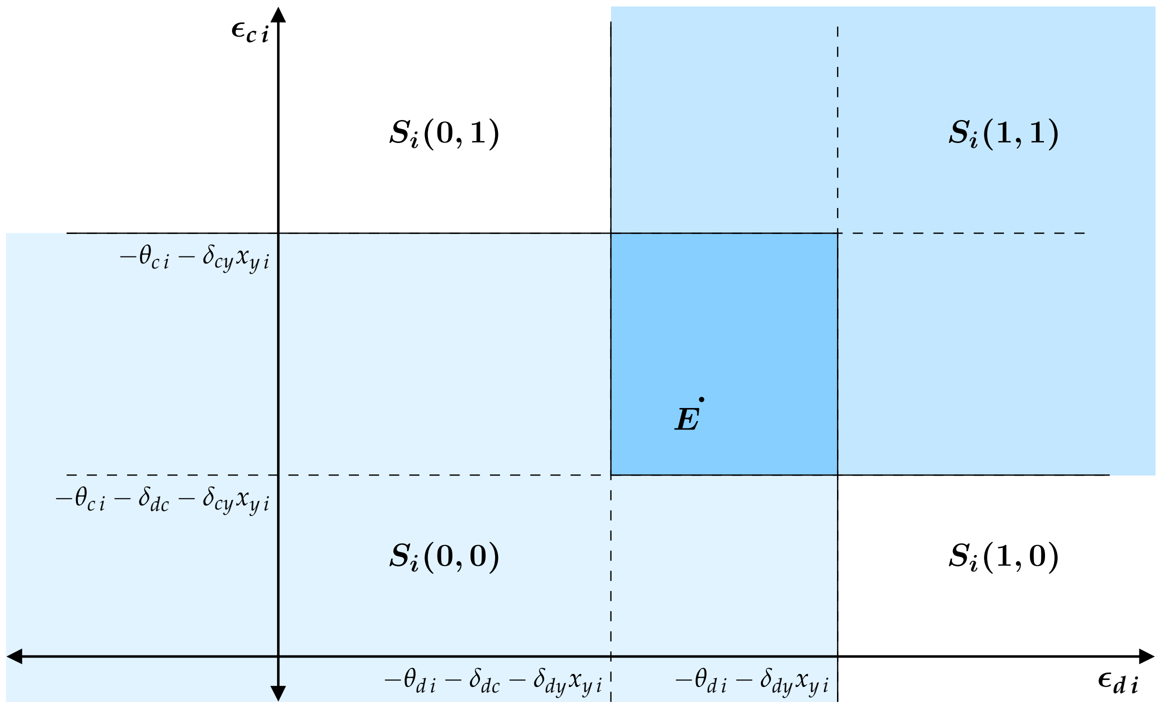
$$S_i(1,0) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} > -\theta_{d\,i} - \delta_{dy}x_{y\,i},\ \epsilon_{c\,i} < -\theta_{c\,i} - \delta_{dc} - \delta_{cy}x_{y\,i}\}, \tag{A.18}$$

$$S_i(0,1) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} < -\theta_{d\,i} - \delta_{dc} - \delta_{dy}x_{y\,i},\ \epsilon_{c\,i} > -\theta_{c\,i} - \delta_{cy}x_{y\,i}\}, \tag{A.19}$$

$$S_i(0,0) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} < -\theta_{d\,i} - \delta_{dy}x_{y\,i},\ \epsilon_{c\,i} < -\theta_{c\,i} - \delta_{cy}x_{y\,i}\}. \tag{A.20}$$

These four regions in the $\epsilon_{d\,i}$–$\epsilon_{c\,i}$ space are represented in Figure 2 for the case where parameters δ_{dc}, δ_{dy}, and δ_{cy} are positive. This figure shows that, according to (A.14) and (A.15), subsets $S_i(1,1)$ and $S_i(0,0)$ overlap each other. This means that the combinations of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ leading to the optimal choice of both innovations and the optimal choice of not innovating at all, respectively are not disjoint sets —their intersection is the dark shaded square at the center of the figure—, and thus, some realizations of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$, such as that represented by point $\boldsymbol{E}$ in Figure 2, may actually lead to the optimal choice of completely different observable innovation profiles. This is the exact meaning of incoherence of the system defined by (A.10) and (A.7).

Figure 2: Incoherence of a System of Equations



To conclude we now address the special case where *intrinsic* complementarities are absent, *i.e.*, when $\delta_{dc} = \delta_{dy} = \delta_{cy} = 0$. In this case matrix $\mathbf{\Gamma}$ is a null matrix and equation (A.10) becomes:

$$\boldsymbol{x}_i^* = \mathbf{\Theta}\boldsymbol{z}_i + \boldsymbol{\epsilon}_i. \tag{A.21}$$

Without *intrinsic* complementarities, the model stops being simultaneous. The optimal choice strategies is independent of each other. Similarly, the rules defining our innovation indicators simplify to:

$$x_{j\,i} = \begin{cases} 1 & \text{si } x^\star_{j\,i} = \epsilon_{j\,i} + \theta_{j\,i} > 0, \\ 0 & \text{si } x^\star_{j\,i} = \epsilon_{j\,i} + \theta_{j\,i} \leq 0. \end{cases} \qquad (j = d, c). \tag{A.22}$$

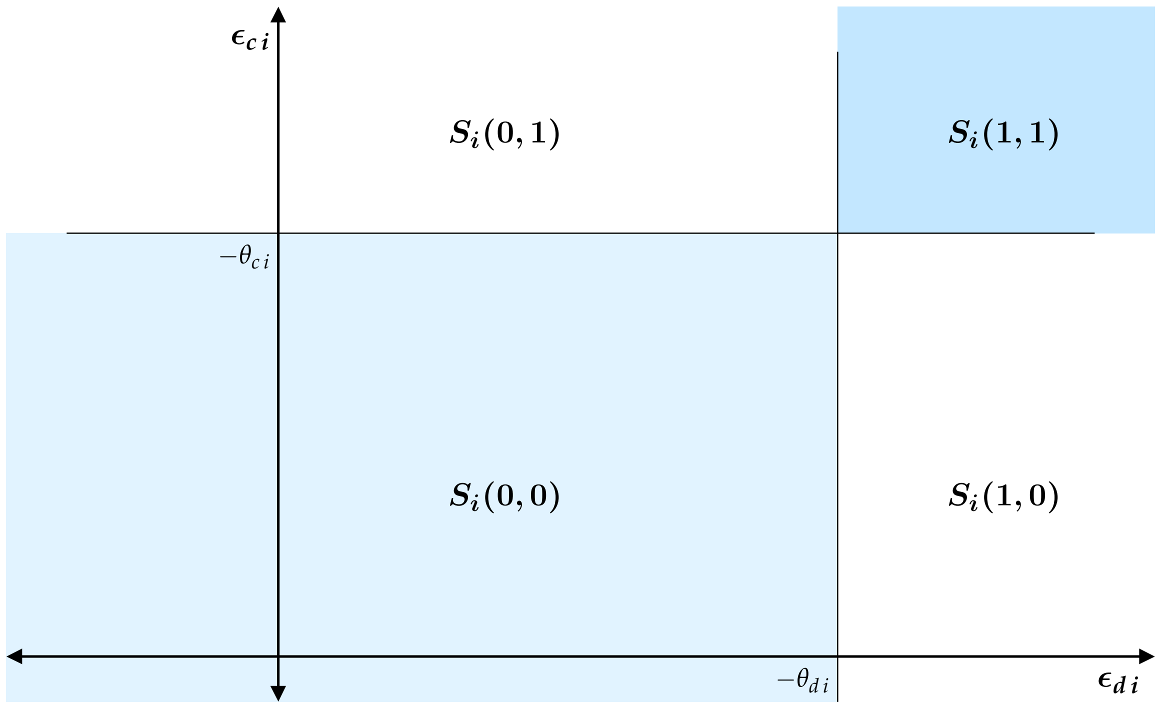
Values of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ leading to each of the four possible innovation profiles are now:

$$S_i(1,1) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} > -\theta_{d\,i},\ \epsilon_{c\,i} > -\theta_{c\,i}\}, \tag{A.23}$$
$$S_i(1,0) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} > -\theta_{d\,i},\ \epsilon_{c\,i} < -\theta_{c\,i}\}, \tag{A.24}$$
$$S_i(0,1) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} < -\theta_{d\,i},\ \epsilon_{c\,i} > -\theta_{c\,i}\}, \tag{A.25}$$
$$S_i(0,0) = \{(\epsilon_{d\,i}, \epsilon_{c\,i}) : \epsilon_{d\,i} < -\theta_{d\,i},\ \epsilon_{c\,i} < -\theta_{c\,i}\}. \tag{A.26}$$

Figure 3 represents the combinations of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ leading to the choice of $x_{d\,i}$ and $x_{c\,i}$. The characterization of the regions associated to each innovation profile becomes much simpler in the case where the interdependence among the choice variables in the profit function is absent. For instance, we now observe $x_{d\,i} = 1$ whenever $\epsilon_{d\,i} > -\theta_{d\,i}$, regardless of any other optimality condition. Figure 3 also shows that set (A.23)–(A.26) are disjoint and they divide the whole $\epsilon_{d\,i}$–$\epsilon_{c\,i}$ space. Therefore, without *intrinsic* complementarity, any combination of values of $\epsilon_{d\,i}$ and $\epsilon_{c\,i}$ uniquely determine the innovation choices.

The model defined by equations (A.21) and (A.22) is coherent and its parameters can be estimated using common methods (equation by equation estimation). Given our assumption on

Figure 3: Adoption of Innovations without *Intrinsic* Complementarity



the distribution of ϵ_i, parameters of the $x_{y\,i}$ equation can be estimated by ordinary least squares, while the determinants of the innovation decisions, $x^\star_{d\,i}$ and $x^\star_{c\,i}$, can be estimated by means of a standard probit model. It suffices to estimate each equation independently of each other in order to estimate consistent estimates of $\mathbf{\Theta}$. A more efficient approach is to estimate the scale and innovation equations jointly in order to account for potential correlation across elements of ϵ_i. Actually our model provides with estimates for correlations among $\epsilon_{d\,i}$, $\epsilon_{c\,i}$, and $\epsilon_{y\,i}$. Observe however that according to our model, the existence of correlation across error terms do not respond to the existence of *intrinsic* complementarity, but rather to complementary *induced* by firms' unobserved heterogeneity.

B Likelihood Function

To distinguish the source and nature of complementarity relations, we estimate the model by maximum likelihood. The following pages describe the derivation of the likelihood function in detail. We assume that the vector of disturbances $\epsilon_i = (\epsilon_{d\,i}, \epsilon_{c\,i}, \epsilon_{y\,i})'$ is normally distributed with zero mean, variances denoted by $(\sigma_d^2, \sigma_c^2, \sigma_y^2)'$ and correlation matrix as given in equation (10). In addition, we assume that environmental variables $\boldsymbol{z}_{r\,i}$, $\boldsymbol{z}_{c\,i}$, and $\boldsymbol{z}_{k\,i}$ are orthogonal to the different components of ϵ_i. Then, the contribution of each observation $(x_{d\,i}, x_{c\,i}, x_{y\,i})$ to the likelihood function (18) can be written as:

$$L_i(x_{d\,i}, x_{c\,i}, x_{y\,i}) = \iint_{S_i(x_{d\,i}, x_{c\,i})} \sigma_y^{-1} \phi_3(\mu_{d\,i}, \mu_{c\,i}, \mu_{y\,i}; \boldsymbol{R}) d\epsilon_{c\,i} d\epsilon_{d\,i}, \tag{B.1}$$

where $\phi_3(\cdot; \boldsymbol{R})$ is a standard trivariate normal density function with correlation matrix $\boldsymbol{R}$, and:

$$\mu_{d\,i} = \epsilon_{d\,i}/\sigma_d, \tag{B.2}$$
$$\mu_{c\,i} = \epsilon_{c\,i}/\sigma_c, \tag{B.3}$$

$$\mu_{y\,i} = \epsilon_{y\,i}/\sigma_y = (x_{y\,i} - \theta_{y\,i} - \delta_{dy}x_{d\,i} - \delta_{cy}x_{c\,i})/\sigma_y, \tag{B.4}$$

are the standardized error terms of the model.

We first define a couple of variables to ease rewriting the relationships described in Table 4 in a more convenient manner for our analysis:

$$q_{j\,i} = 2(x_{j\,i} - 1), \qquad (j = d, c), \tag{B.5}$$

that takes value 1 when $x_{j\,i} = 1$ and -1 whenever innovation strategy j is not adopted. We then define the following indicators:

$$s_i = q_{d\,i}q_{c\,i}, \tag{B.6}$$
$$m_i = (s_i + 1)/2. \tag{B.7}$$

Therefore, s_i (m_i) takes value -1 (1) when only one of the innovation strategies is adopted and value 1 (0) otherwise, *i.e.*, when either both innovations are adopted, or when both innovations are not adopted simultaneously. We can then define $S_i(x_{d\,i}, x_{c\,i})$, the set of realizations $(\epsilon_{d\,i}, \epsilon_{c\,i})$ leading to the observed choices of $x_{d\,i}$ and $x_{c\,i}$, from the following inequalities:

$$q_{d\,i}\epsilon_{d\,i} > -q_{d\,i}(\kappa_{d\,i} + \delta x_{c\,i}), \tag{B.8}$$
$$q_{c\,i}\epsilon_{c\,i} > -q_{c\,i}(\kappa_{c\,i} + \delta x_{d\,i}), \tag{B.9}$$

and if: $s_i\delta > 0$,

$$q_{c\,i}\epsilon_{c\,i} > -q_{c\,i}(\kappa_{c\,i} + m_i\delta + s_i\kappa_{d\,i} + s_i\epsilon_{d\,i}). \tag{B.10}$$

Computation of the likelihood function gets complicated by the fact that when $s_i\delta > 0$, the integration areas are not rectangular any more. To illustrate this issue, let analyze the contribution to the likelihood of an observation where a firm innovate both in product and process: $x_{d\,i} = 1, x_{c\,i} = 1$. According to (B.6), $s_i = 1$ in this case, so the sign of $s_i\delta$ is the same as the sign of δ. If $\delta < 0$ the integration is defined on the rectangular region $S_i(1,1) = [-\kappa_{d\,i} - \delta, \infty) \times [-\kappa_{c\,i} - \delta, \infty)$, shown in the upper right area of Figure 1b. On the contrary, if $\delta > 0$ the integration region for joint adoption of innovations is not rectangular. Figure 1a shows that $S_i(1,1)$ is a subset of $[-\kappa_{d\,i} - \delta, \infty) \times [-\kappa_{c\,i} - \delta, \infty)$ when $\delta s_i > 0$. Therefore, making use of the rules that define the sets $S_i(x_{d\,i}, x_{c\,i})$, (B.8)–(B.10), we can write (B.11) as:

$$L_i(x_{d\,i}, x_{c\,i}, x_{y\,i}) = \int_{k_{d\,i}}^{\infty}\int_{k_{c\,i}}^{\infty} \sigma_y^{-1}\phi_3(\mu_{d\,i}, \mu_{c\,i}, \mu_{y\,i}; \boldsymbol{R}^*)d\mu_{c\,i}d\mu_{d\,i}$$
$$- \boldsymbol{1}(s_i\delta > 0)q_{d\,i}\int_{a_{d\,i}}^{b_{d\,i}}\int_{a_{c\,i}}^{b_{c\,i}} \sigma_y^{-1}\phi_3(\mu_{d\,i}, \mu_{c\,i}, \mu_{y\,i}; \boldsymbol{R})d\mu_{c\,i}d\mu_{d\,i}, \tag{B.11}$$

where $\boldsymbol{R}$ is the correlation matrix of $\boldsymbol{\epsilon}_i$, $\boldsymbol{R}^*$ is the correlation matrix of $(q_{d\,i}\epsilon_{d\,i}, q_{c\,i}\epsilon_{c\,i}, \epsilon_{y\,i})$, that is:

$$\boldsymbol{R}^* = \begin{bmatrix} 1 & s_i\rho_{dc} & q_{d\,i}\rho_{dy} \\ s_i\rho_{dc} & 1 & q_{c\,i}\rho_{cy} \\ q_{d\,i}\rho_{dy} & q_{c\,i}\rho_{cy}\rho_{dc} & 1 \end{bmatrix} \tag{B.12}$$

and $\boldsymbol{1}(\cdot)$ is the indicator function that takes value 1 when its argument is true and 0 otherwise. Finally, the limits of integration of (B.11) are:

$$k_{d\,i} = -q_{d\,i}(\kappa_{d\,i} + \delta x_{c\,i})/\sigma_d, \qquad k_{c\,i} = -q_{c\,i}(\kappa_{c\,i} + \delta x_{d\,i})/\sigma_c, \tag{B.13}$$
$$a_{d\,i} = -(\kappa_{d\,i} + \delta)/\sigma_d, \qquad b_{d\,i} = -\kappa_{d\,i}/\sigma_d, \tag{B.14}$$
$$a_{c\,i} = -(\kappa_{c\,i} + \delta x_{d\,i})/\sigma_c, \qquad b_{c\,i} = -(\kappa_{c\,i} + m_i\delta + s_i\kappa_{d\,i} + s_i\sigma_d\mu_{d\,i})/\sigma_c. \tag{B.15}$$

The first integral (B.11) contains the mass of probability associated to the rectangular regions of the error space defined by equations (B.8) and (B.9). When $s_i\delta > 0$ and $S_i(x_{d\,i}, x_{c\,i})$ are not rectangular, this integral overestimates L_i. The second integral of (B.11) corrects this bias.

The first integral of (B.11) can be written, after conditioning on $\mu_{y\,i}$, as a function of a single dimensional normal probability density function, $\phi(\cdot)$, and a bivariate normal probability distribution function, $\Phi_2(\cdot;\rho)$, as follows:

$$\int_{k_{d\,i}}^{\infty}\int_{k_{c\,i}}^{\infty}\sigma_y^{-1}\phi_3(\mu_{d\,i},\mu_{c\,i},\mu_{y\,i};\boldsymbol{R}^*)d\mu_{c\,i}d\mu_{d\,i} = \sigma_y^{-1}\phi(\mu_{y\,i})\Phi_2(-k_{d.y\,i},-k_{c.y\,i};s_i\rho_{dc.y}), \tag{B.16}$$

where:

$$k_{j.y\,i} = \frac{k_{j\,i} - q_{j\,i}\rho_{jy}\mu_{y\,i}}{(1-\rho_{jy}^2)^{1/2}}, \qquad (j = d, c), \tag{B.17}$$

and

$$\rho_{dc.y} = \frac{\rho_{dc} - \rho_{dy}\rho_{cy}}{\left[(1-\rho_{dy}^2)(1-\rho_{cy}^2)\right]^{1/2}}. \tag{B.18}$$

We proceed similarly to evaluate the second integral of (B.11). Conditioning on $\mu_{y\,i}$, we get:

$$\int_{a_{d\,i}}^{b_{d\,i}}\int_{a_{c\,i}}^{b_{c\,i}}\sigma_y^{-1}\phi_3(\mu_{d\,i},\mu_{c\,i},\mu_{y\,i};\boldsymbol{R})d\mu_{c\,i}d\mu_{d\,i} = \sigma_y^{-1}\phi(\mu_{y\,i})\int_{a_{d.y\,i}}^{b_{d.y\,i}}\int_{a_{c.y\,i}}^{b_{c.y\,i}}\phi_2\big(\mu_{d.y\,i},\mu_{c.y\,i};\rho_{dc.y}\big)d\mu_{c.y\,i}d\mu_{d.y\,i}, \tag{B.19}$$

where:

$$a_{j.y\,i} = \frac{a_{j\,i} - \rho_{jy}\mu_{y\,i}}{(1-\rho_{jy}^2)^{1/2}}, \qquad b_{j.y\,i} = \frac{b_{j\,i} - \rho_{jy}\mu_{y\,i}}{(1-\rho_{jy}^2)^{1/2}}, \qquad (j = d, c). \tag{B.20}$$

Next, conditioning $\mu_{c.y\,i}$ on $\mu_{d.y\,i}$, we can write (B.19) as:

$$\int_{a_{d\,i}}^{b_{d\,i}}\int_{a_{c\,i}}^{b_{c\,i}}\sigma_y^{-1}\phi_3(\mu_{d\,i},\mu_{c\,i},\mu_{y\,i};\boldsymbol{R})d\mu_{c\,i}d\mu_{d\,i} = \sigma_y^{-1}\phi(\mu_{y\,i})\int_{a_{d.y\,i}}^{b_{d.y\,i}}\phi\big(\mu_{d.y\,i}\big)\int_{a_{c.dy\,i}}^{b_{c.dy\,i}}\phi\big(\mu_{c.dy\,i}\big)d\mu_{c.dy\,i}d\mu_{d.y\,i}, \tag{B.21}$$

where:

$$a_{c.dy\,i} = \frac{a_{c.y\,i} - \rho_{dc.y}\mu_{d.y\,i}}{(1-\rho_{dc.y}^2)^{1/2}}, \qquad b_{c.dy\,i} = \frac{b_{c.y\,i} - \rho_{dc.y}\mu_{d.y\,i}}{(1-\rho_{dc.y}^2)^{1/2}}. \tag{B.22}$$

Integrating now (B.21) with respect to $\mu_{c.dy\,i}$:

$$\int_{a_{d\,i}}^{b_{d\,i}}\int_{a_{c\,i}}^{b_{c\,i}}\sigma_y^{-1}\phi_3(\mu_{d\,i},\mu_{c\,i},\mu_{y\,i};\boldsymbol{R})d\mu_{c\,i}d\mu_{d\,i} = \sigma_y^{-1}\phi(\mu_{y\,i})\int_{a_{d.y\,i}}^{b_{d.y\,i}}\phi\big(\mu_{d.y\,i}\big)\Big[\Phi(b_{c.dy\,i}) - \Phi(a_{c.dy\,i})\Big]d\mu_{d.y\,i}. \tag{B.23}$$

Finally, substituting (B.16) and (B.23) in(B.11), we get:

$$L(x_{d\,i}, x_{c\,i}, x_{y\,i}) = \sigma_y^{-1}\phi(\mu_{y\,i})\Bigg\{\Phi_2(-k_{d.y\,i}, -k_{c.y\,i}; s_i\rho_{dc.y})$$
$$- \mathbf{1}(s_i\delta > 0)q_{d\,i}\int_{a_{d.y\,i}}^{b_{d.y\,i}} \phi\big(\mu_{d.y\,i}\big)\Big[\Phi(b_{c.dy\,i}) - \Phi(a_{c.dy\,i})\Big]d\mu_{d.y\,i}\Bigg\}. \quad \text{(B.24)}$$

The only remaining difficulty in the evaluation of the likelihood function is the computation of the integral shown in the second line of (B.24). Changing variables so that $\tau_i = 2(\mu_{d.y\,i} - a_{d.y\,i})/(b_{d.y\,i} - a_{d.y\,i}) - 1$, this integral can fortunately be easily evaluated by means of a Gauss-Legendre quadrature (see Stroud and Secrest, 1966, for instance). The results of this paper were obtained using a 40 points rule to evaluate the likelihood functions. All computations were carried out with Ox 3.30 (Doornik, 2002).

A special case of our model occurs when δ_{dc}, δ_{dy} and δ_{cy} equal zero. This case was discussed in Appendix A. Without *intrinsic* complementarity, our model simplifies to a single linear equation and a couple of probits. However, in this case, standard deviations σ_d and σ_c are not identified. To avoid problems of local identification we normalize $\sigma_d = 1$ and $\sigma_c = 1$, as it is commonly done in probit models.