Why do industrial firms in developing countries grow after they begin exporting? Evidence from Brazil[1]

Abstract

In this article, we verify the existence of *ex-post* gains of productivity and performance in the Brazilian companies that entered the export market during the first years of the current decade. It is a comparative analysis of the performance of companies debuting in the international market vis-a-vis those that never exported during the period of our analysis. Several sample drawings were used in this sense, including taking the innovative nature of these companies into consideration. However, our major contribution resides in characterizing the technological standard of *learning by exporting.* While innovative activities, generally speaking, (especially, intramural R&D) showed to be important for all companies, for those that began exporting the array of relevant sources of information was bigger, including other internal sources of the company, licenses, patents, consulting companies, competitors, and machinery and equipment suppliers. Despite the limitations of the exercise herein conducted, the hypothesis of *learning by exporting* resulting from better access to input and imported equipment, bigger exposure to competition and to the possibility of technological cooperation with other companies in the productive chain seems to be true in Brazil.

Key-words: *learning by exporting*, Technological standards, Productivity gains, Principal components analysis, *Propensity score matching*.

[1] The authors are deeply thankful for Mr. João De Negri's comments and for Eric Jardim's and Geovane Lopes' statistical assistance, without, naturally, implicating them in any way for possible errors and omissions present in this version of the article.

Why do industrial firms in developing countries grow after they begin exporting? Evidence from Brazil

Introduction

The relation between productivity and exports is one of the most traditional topics of study in economics. Initially, Ricardo's paradigm of comparative advantages established as the engine of international commerce the existence of advantages of relative costs of production, starting from differences in the relative labor productivity among countries. Later on, Heckscher-Ohlin's theorem and even Krugman-Helpman's "new theories" connected themselves to this paradigm when relating the differences in productivity, respectively, to the relative factors endowments and to the possibility of scale gains resulting from international trade.

Recent availability of data, computational resources and specific econometric techniques had led researchers to evaluate, in an alternative way, the theories of international trade. As a stylized fact from this literature, companies that export present competitiveness indicators which are more favorable than those of companies that do not export (Ellery and Gomes, 2005 and Tybout, 2003). However, one can observe that these competitiveness gains occur *before* the company starts to export, which is consistent with the evidence of fixed entry costs and bigger competition in the international market, which makes companies that are more productive self-select.

Nevertheless, appropriate literature points out that there may be future gains for exporting companies after their entrance in this market, especially in developing countries. The causes for the existence or not of learning effects are related to the possibilities of exporters having better access to input and imported equipment, adapting themselves to higher technological standards and quality – due both to bigger exposure to competition as well as to the possibility of technological cooperation with other companies, in the productive chain, worldwide – and also to scale economies associated with exports, effects which tend to be more intense in developing countries (Aw and Hwang, 1995 and Clerides, Lauch and Tybout, 1998).

In fact, the hypothesis of *learning-by-exporting* has been confirmed in several developing countries.[2] However, in most empirical work in which *ex-post* gains from exports are confirmed, the growth process of the companies that begin exporting activities is not characterized at all: for instance, it is not possible to know if exporting firms innovate, if they import machinery and equipment, or if the suppliers or clients abroad constitute their engine of technological convergence. With the exception of anecdotal and casual evidence, everything that is behind these learning effects is a "black-box" from a technological point of view.

Thus, this paper makes three contributions:

I. The first one is to present the productivity gains of *rookie* exporting companies in comparison to those that never exported during the period of our analysis. These gains are computed the year the company entered the export market as well as the previous and following years of when this debut began. This analysis is made for three kinds of firms: exporting companies that remained exporting after their debut (herein called permanent exporters); those which started to export in the debut year, but did not restart exporting in the subsequent years (herein called occasional exporters); and those that entered the market, but only exported occasionally in the following period (herein called unstable).

[2] For a brief review of the literature on this topic, please read Araújo (2006).

II. The second one is to show that, starting from a sample of comparable innovative companies, Brazilian industrial companies that started operating in the international market between 2001 and 2002 presented higher turnover, employment and productivity gains in comparison to those that did not export in the period 2003-2005.[3]

III. The third contribution, that of a higher level of uniqueness, consists in characterization of the technological learning process of these new-exporters by checking which technological learning patterns are more related to the learning effects resulting from foreign trade.

This contribution is possible because data availability in Brazil allows the follow-up of a panel between 1999 and 2005, in which several databases are connected, especially the Annual Industrial Survey (PIA) and the Brazilian Innovation Survey (Pintec), both from IBGE, information from the Foreign Trade Secretariat (Secex) on exports and imports, in addition to a profile of the workforce provided by the Annual List of Social Information (Rais) of the Ministry of Labor and Employment. This concatenation was carried out by the Institute of Applied Economics (Ipea), which, nevertheless, does not its own physical copy of this data; and it respects all the confidentiality rules involved.

The rest of this article is structured in the following way: in the second section, data used in our analysis are presented. In the third section, we discuss the treatment given the problem of self-selection of the companies that start operating in the international market. In the fourth section, the principal component analysis is carried out in order to characterize the technological standards of the companies in our sample is presented. The fifth section brings the results technological standards have over the learning effects resulting from exports. Finally, in the sixth and last section one can find the conclusions of the article.

Data and description of the variables

The data used to evaluate the relation between the learning effects and the technological standards result from the concatenation of a few databases Annual Industrial Survey (PIA) and the Brazilian Innovation Survey (Pintec), both from IBGE, information from the Foreign Trade Secretariat (Secex), from the Ministry of Development, Industry and Foreign Trade and from the List of Social Information (Rais) of the Ministry of Labor and Employment. The period of the analysis ranged 1999 to 2005. Companies were considered to be *rookies* (or new exporters) if they started exporting either in 2001 or in 2002 and, not having exported at all two years before, began exporting for at least two consecutive years. Additionally, they exported at least in one of the years in the period 2004-2005.

Our analysis restricts itself to innovative companies during the period 2001-2003 and 2003-2005 (the two last rounds of the Pintec), whether they are new exporters or companies that never exported in the period we analyzed. This happened because only innovative companies answered the part of Pintec's questionnaire in which we are directly interested, that is, questions regarding the importance of innovative activities and sources of information for innovation. Nonetheless, we know that this procedure hinders any analysis of the effects of *learning by exporting* over companies that were not innovative and started innovating, and it also makes impossible the extrapolation of the results for the Brazilian industry as a whole. We intend to overcome the external validity problems of our study in a future version of this paper. After having made these considerations, our analysis refers to 1,044 companies, 20% of which are *rookies*.

[3] This result aligns itself with Araújo's paper (2006), which found turnover, size (measured by the number of employees) and productivity gains for new exporting companies in comparison with companies that did not begin to export in the period 1998-2003.

The positive effects of entering the international market are evaluated based upon productivity, employment, turnover and the *market share* of industrial companies. The measure of employment (PO) is an average of total employees. As a measure of productivity, we used the ratio between the value added in industrial transformation and the PO. We measured the company's turnover as if it were the Net Revenue from Sales, and we used this variable in the construction of the variable *market share* of the company at the CNAE sector (National Classification of Economic Activities, comparable to SIC, Standard Industrial Classification) at three-digit level. All these variables were considered in three periods: "before" the company's debut in the international market (average of the period 1999-2000), "during" its debut (average 2001-2002) and "after" its debut (average 2003-2005). The nominal variables of the PIA were deflated by the Wholesale Price Index – Overall Supply (IPA-OG) of Getulio Vargas Foundation (FGV). The advantage of using the IPA-OG is the availability of specific inflation indexes for the CNAE sectors at three-digit level.

Besides these impact variables, the variable of the company's age, estimated based on the RAIS using the methodology created in Costa *et al* (2006) and a sectorial classification based on the technological/factor intensity was used to estimate the *probit* model that refers to the probability of exporting, which is part of our empirical strategy that will deal with the self-selection in the decision to export.

Finally, the variables of the Pintec that were used to characterize the companies' technological learning patterns are the variables that refer to the importance attributed to innovation activities and to sources of information for innovation. In every case, the company receives 1 if it attributes high or medium importance to innovative activities such as R&D or expenses in machinery and equipment related to innovation, or to sources of information such as universities and research institutes, and clients and suppliers, for example. In short, all the variables used in this study and its definitions can be found in Table 1 as follows.

Table 1 – Construction of the variables used in the study

Variable	Concept	Variable Code	Base of Origin	Periods
rookie	Variable that indicates if a company has began to export in 2001 or in 2002. For a company to be considered estreante, it can not have exported in the two previous years. Once into international markets, it needs to keep exporting by two consecutive years and at least in one the years of 2004-2005.	-	Secex, 1999-2005	2001 or 2002
PO	Average employees of the firm	X02	PIA, 1999-2005	Before (1999-2000), During (2001-2002) and After (2003-2005)
Productivity	Value Added in Industrial Transformation/PO (VTI)/PO	X32/X02	PIA ,1999-2005	Before (1999-2000), During (2001-2002) and After (2003-2005)
RLV	Net revenue from sales	X14	PIA, 1999-2005	Before (1999-2000), During (2001-2002) and After (2003-

				2005)
Market Share, Market Share^2	Market share on Net Revenue from Sales, at CNAE 3 level	-	PIA, 1999-2005	Before (1999-2000), During (2001-2002) and After (2003-2005)
Sectorial Controls	Sectorial Control *Dummies*, based on factor and technological intensity of the sector	-	PIA, 2001	During (2001-2002)
Firm's age	Age of the company (check out Costa et al. 2006).	-	Rais, 2001	During (2001-2002)

Sources of information for innovation

Fpedi	Importance attributed to intramural R&D as a source of information for innovation (=1 if medium or high, 0 otherwise)	V108	Pintec, 2003-2005	After (2003-2005)
Fofi	Importance attributed to other internal sources as a source of information for innovation (=1 if medium or high, 0 otherwise)	V109	Pintec, 2003-2005	After (2003-2005)
Ffmq	Importance attributed to the suppliers of machinery and equipment as a source of information for innovation (=1 if medium or high, 0 otherwise)	V110	Pintec, 2003-2005	After (2003-2005)
Fcc	Importance attributed to clients and costumers as a source of information for innovation (=1 if medium or high, 0 otherwise)	V111	Pintec, 2003-2005	After (2003-2005)
Fconc	Importance attributed to competitors as a source of information for innovation (=1 if medium or high, 0 otherwise)	V112	Pintec, 2003-2005	After (2003-2005)
Feci	Importance attributed to consulting companies as a source of information for innovation (=1 if medium or high, 0 otherwise)	V113	Pintec, 2003-2005	After (2003-2005)
Fuip	Importance attributed to universities and research institutes as a source of information for innovation (=1 if medium or high, 0 otherwise)	V114	Pintec, 2003-2005	After (2003-2005)
Fccp	Importance attributed to training institutes as a source of information for innovation (=1 if medium or high, 0 otherwise)	V115	Pintec, 2003-2005	After (2003-2005)
Fiec	Importance attributed to metrology and other centers of basic industrial technology as a source of information for innovation (=1 if medium or high, 0 otherwise)	V116	Pintec, 2003-2005	After (2003-2005)
Fpat	Importance attributed to licenses	V117	Pintec, 2003-	After (2003-

	and patents as a source of information for innovation (=1 if medium or high, 0 otherwise)		2005	2005)
Fconf	Importance attributed to conferences as a source of information for innovation (=1 if medium or high, 0 otherwise)	V118	Pintec, 2003-2005	After (2003-2005)
Ffeira	Importance attributed to fairs as a source of information for innovation (=1 if medium or high, 0 otherwise)	V119	Pintec, 2003-2005	After (2003-2005)
Frii	Importance attributed to information networks as a source of information for innovation (=1 if medium or high, 0 otherwise)	V120	Pintec, 2003-2005	After (2003-2005)
	Innovation Activities			
Pedi	Importance attributed to intramural R&D as an innovation activity (=1 if medium or high, 0 otherwise)	V24	Pintec, 2003-2005	After (2003-2005)
Pede	Importance attributed to extramural R&D as an innovation activity (=1 if medium or high, 0 otherwise)	V25	Pintec, 2003-2005	After (2003-2005)
Pedoc	Importance attributed to the acquisition of other kinds of knowledge as an innovation activity (=1 if medium or high, 0 otherwise)	V26	Pintec, 2003-2005	After (2003-2005)
Aqmed	Importance attributed to the purchase of machinery and equipment as an innovation activity (=1 if medium or high, 0 otherwise)	V27	Pintec, 2003-2005	After (2003-2005)
Trein	Importance attributed to training of the workforce as an innovation activity (=1 if medium or high, 0 otherwise)	V28	Pintec, 2003-2005	After (2003-2005)
Imec	Importance attributed to the expenses with the introduction of innovation into the market as an innovation activity (=1 if medium or high, 0 otherwise)	V29	Pintec, 2003-2005	After (2003-2005)
Prii	Importance attributed to the expenses with industrial projects as an innovation activity (=1 if medium or high, 0 otherwise)	V29	Pintec, 2003-2005	After (2003-2005)

Source: Created by the authors

Productivity and Learning-by-exporting

As mentioned before, the objectives of this article are: (i) to show the efficiency gains post-entrance, in the international market, of *rookie* exporting companies and (ii) to characterize the technological process that permeates the learning process of these corporations.

The first step in this direction is to find evidence that, effectively, there is an improvement process in the performance of companies resulting from the learning obtained from the exporting activity (learning by exporting). This point is essential for our project, since these *ex-post* productivity gains do not constitute a robust result in the existing empirical literature.

On the other hand, a well known fact in all empirical literature is the existence of *ex-ante* productivity gains, that is, companies that before beginning to export already show considerable productivity gains. This stylized fact corroborates the idea that exporting companies incur in costs when they begin to export, and, therefore, only the more productive companies self-select themselves as exporting companies.

The objective of this section is to provide a few preliminary results regarding the existence or not of *ex-ante* and *ex-post* productivity gains for Brazilian exporting companies. Consequently, *ex-ante* and *ex-post* productivity gains are obtained for new starters in the export market. These gains capture the productivity differential of *rookie* companies in contrast with those companies that did not present any exporting activity in the period analyzed.

The identification strategy for obtaining the productivity premium of the exporting companies, the standard of productive performance and the existence of a learning process (*learning by exporting*) follows the seminal work developed by Bernard and Jensen (1999). The formatting of the database for the execution of the empirical strategy obeyed the following steps:

1. The companies that started exporting in 2000 were selected (after we ensured, through SECEX's database, that these companies did not present any records of exports between 1996 and 1999), and, only after this, did we follow up their productive performance histories throughout the period between 1999 and 2003.
2. This performance history is compared, year after year, to a group of companies that did not present any records of exports in any of SECEX's databases throughout the period between 1996 and 2005. This comparison of the productivity between exporting companies starting in the market in 2000 and non-exporting companies occurs via a performance equation with the following specification:

$$\ln \pi_i = \alpha EXP_i + x_i \beta + e_i \,, \quad (1)$$

in which $\ln \pi_i$ is the natural logarithm of the value added per employee for the company i - as *proxy* for productivity -; EXP_i is a binary variable equal to one in case the company is a *rookie* exporting company starting in 2000, and equal to zero in otherwise; x_i is a vector of control variables for the company i[4]; e_i is the random error term; and α and β are parameters to be estimated by Ordinary Least Squares (OLS), where α represents the productivity premium, that is, the productivity differential (conditional to vector x_i) of the *rookie* exporting company, the ones that began in 2000, in relation to the non-exporting companies.

The companies enter the sample based on the following process:

[4] The variables that compose vector x_i include: the firm's scale (log of employment), average schooling of the workforce, market share at CNAE 3 level, its location (*dummies* for the Brazilian state) and its sector (*dummies* for CNAE3).

- We check all the companies that compose the PIA in year *t* (*t* = 1999 to 2003);
- By using SECEX's information, we attribute to each company the following status: (A) exporting companies starting in 2000, (B) other exporting companies and (C) non-exporting companies;
- We eliminate from the sample all the exporting companies[5] with status (B);
- This procedure ensures that a sample only presents companies with status (A) and (C) for each year, but nothing guarantees that all the companies are listed in all the years of the analysis. Example: in SECEX's databases, we obtained the number of companies that started exporting in 2000; yet these companies can appear in PIA's database between 1999 and 2001, and get out of the database in 2002 or in 2003[6].

Regression (1) was run for each of the years between 1999 and 2003 so that the comparisons of the exporting premiums were obtained before (*ex-ante* premium), during, and after (*ex-post* premium) the companies entered the export market in 2000. The sample sizes for each year are 11,959 (1999), 12,396 (2000), 13,792 (2001), 14,598 (2002) and 15,675 (2003). The percentage of exporting companies debuting in the foreign market in 2000 corresponds to approximately 5.6% of the sample.

The results of the regressions mentioned above are reported in the second column (II) of Table 2. When checking the conditional productivity premium (α) for the period between 1999 and 2003, we obtain the following results: (a) new exporters in 2000 presented productivity premium already 30% higher than non-exporting companies, even before they started exporting (see year 1999); (b) this differential increases almost two percent in the year the companies begin exporting (see year 2000); (c) three years after the entrance of these exporting companies, the conditional productivity differential reaches the α value of 43%, clearly signaling the learning effect (*learning by exporting*).

Table 2: Productivity premium, new exporters in 2000

Year (I)	Exporting Firms (II)	Stayers (III)	Quitters (IV)	Movers (V)
1999	0.3058	0.3429	0.3194	0.2252
2000	0.3197	0.3800	0.2103	0.3104
2001	0.3735	0.3803	0.2362	0.3413
2002	0.4123	0.5255	0.2140	0.3138
2003	0.4364	0.5504	0.1696	0.3583

Source: Created by the authors based on the PIA, Pintec, Secex and Rais

The same exercise was done for the companies that started exporting in 2001. In this case, sample sizes for each year are 11,819 (1999), 12,181 (2000), 13,664 (2001), 14,549 (2002) and 15,659 (2003). The percentage of exporting companies starting in the market in 2001 corresponds to approximately 5% of the sample[7].

The results of the exercise above are reported in column (II) of Table 3. We noticed that the results related to the productive performance and the existence of the learning process by exporters are robust. Firms seem to begin to export when they present a conditional productivity premium (α) of approximately 30% and this premium increases considerably after their first year in foreign markets.

[5] Exporting companies are those that showed some kind of exporting activity in SECEX's databases in the period 1996-2005.

[6] It is important to mention that when a company exits the PIA it does not necessarily mean that the company went bankrupt. The PIA is a censitary database for companies with more than 30 employees, but it also includes a random sample of companies with less than 30 employees. A company can leave the sample in any given year or period of time simply due to a reduction in the number of employees.

[7] The percentage of debuting companies in 2000 is 0.6% higher than the percentage of companies starting in 2001. This difference in the propensity of entering the exporting sector can be a result of the effects of the currency devaluation in 1999.

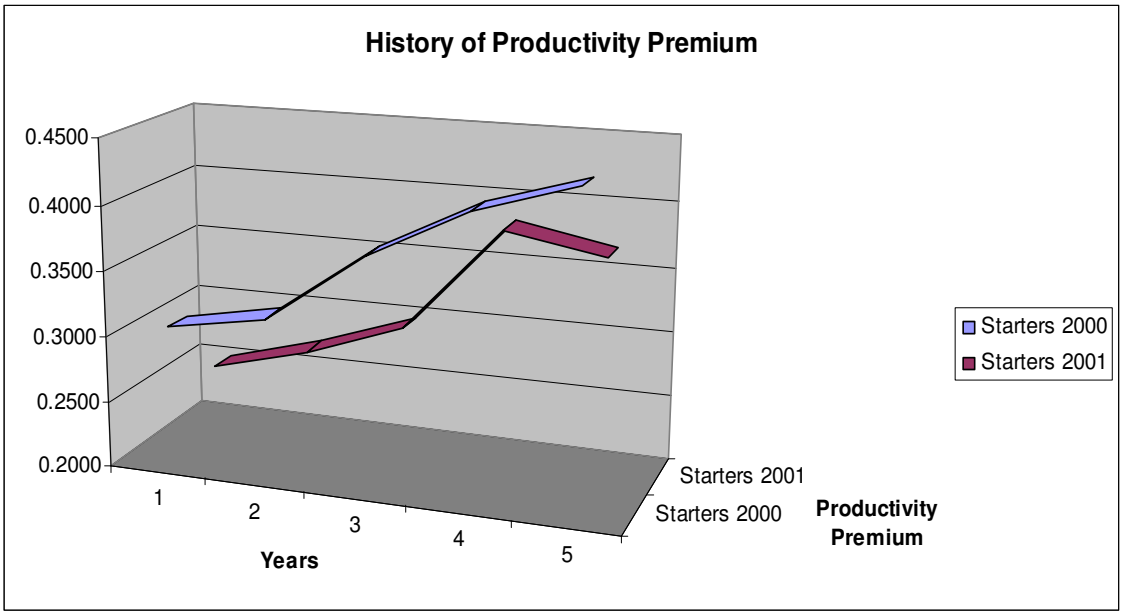
Table 3: Productivity premium, new exporters in 2001

Year (I)	Exporting Firms (II)	Stayers (III)	Quitters (IV)	Movers (V)
1999	0,2508	0,2636	0,2444	0,1873
2000	0,2712	0,3123	0,2188	0,2512
2001	0,2980	0,3259	0,1880	0,3500
2002	0,3811	0,4810	0,1778	0,2816
2003	0,3674	0,4584	0,1758	0,3215

Source: Created by the authors based on the PIA, Pintec, Secex and Rais

The results are also shown in Graph 1, which illustrates the respective temporal paths of the conditional productivity premium for the companies starting in 2000 and 2001. Inclination changes of the productivity curves after the entrance of these companies in the exporting sector show the dimension of the learning effect.

Graph 1: History of the productivity premium of new exporters in 2000 and 2001.



Source: Created by the authors based on the PIA, Pintec, Secex and Rais

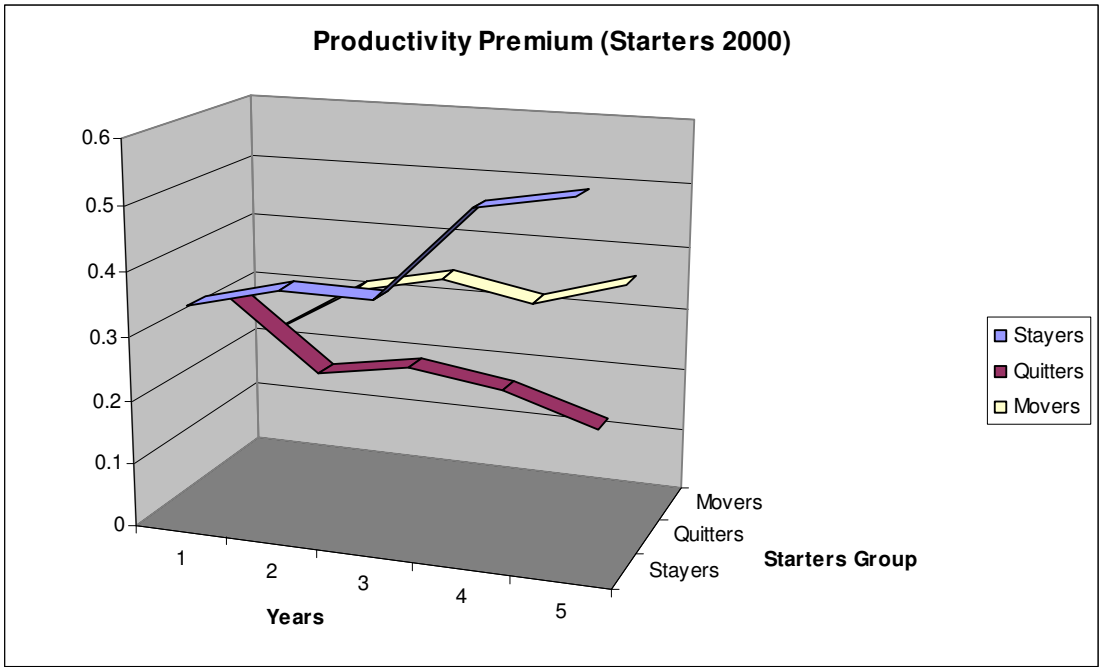
Given the evidence on productivity gains based upon the learning process, a question that needs to be asked is: why do some companies present unstable behavior or abandon the exporting activity permanently? The initial hypothesis to be tested is that not all companies starting in the market will present productivity gains resulting from learning.

The next objective of this section is to replicate the previous strategy for the companies starting in the market and classified in different levels of survival: Stayers (firms with permanent exporting activity after entering), Movers (firm with unstable exporting activity after entering), and Quitters (firms leaving exporting after entering). The results for *rookie* companies starting in 2000 are reported in columns (III)-(IV) of Table 2. When checking out the conditional productivity premium (α) for each one of these groups of companies, we identified the following results: (a) all new exporters presented, in the year before their entrance, a productivity premium of not less than 20%; (b) these premiums increased in the entrance year for the group of exporting companies

classified as Stayers and Movers; (c) this premium strongly decreased in the debut year for the group of exporting companies classified as Quitters; (d) the productivity premium did not stop growing in the subsequent years for Stayers; (e) the productivity premium stopped at 30% for the Movers; (f) Quitters could not recover, in the subsequent years, the productivity premium observed before their debuts.

These results were replicated for the group of exporting companies starting in 2001. In this group of companies, Stayers correspond, in average, to 2.2% of the sample, while these percentages are 1.5% for the Quitters, and 1.2% for the Movers[8]. Such results are reported in columns (III)-(IV) of Table 3, confirming the soundness of the results discussed in the previous paragraph. The stylized facts mentioned above can also be easily observed if we analyze the changes of the inclinations of the curves in Graphs 2 and 3 below.

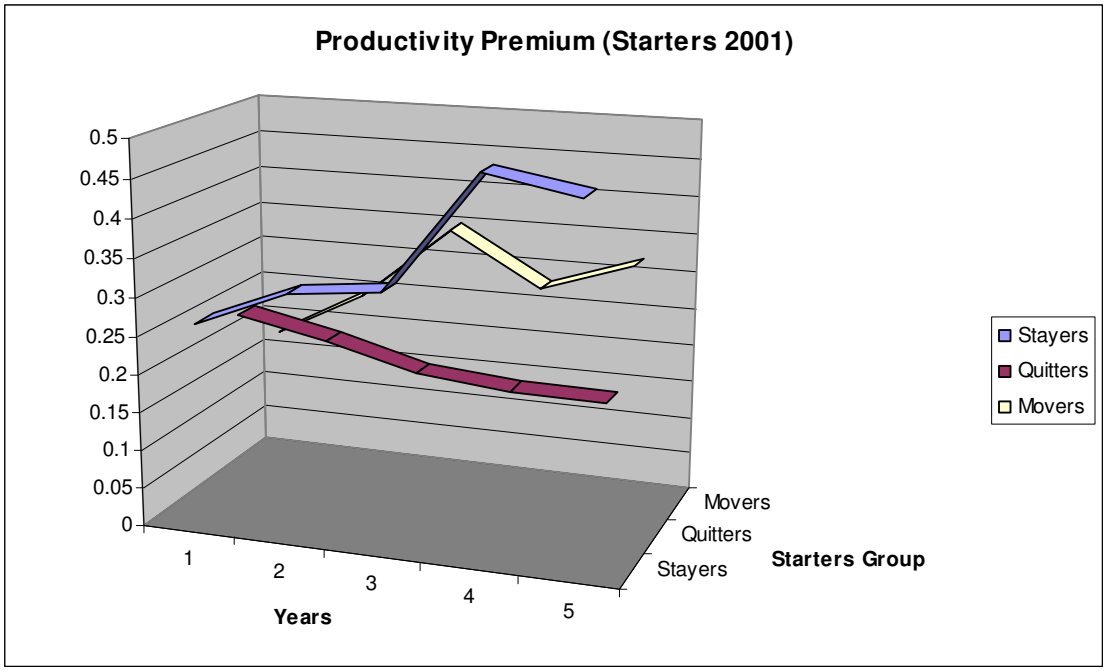
Graph 2: History of the productivity premium of new exporters in 2000, by category.



Source: Created by the authors based on the PIA, Pintec, Secex and Rais

[8] It is interesting to notice that the difference of 0.6% in the number of companies starting between 2000 and 2001 is basically concentrated on the participation of Movers companies. If this difference is really due to the currency devaluation of 1999, we can infer that the positive impact of currency devaluations over exporting companies happens basically due to entrance patterns by this type of company.

Graph 3: History of the productivity premium of new exporters in 2001, by category.



Source: Created by the authors based on the PIA, Pintec, Secex and Rais

A summary of the results above provide the following answers to the questions listed in the introduction of this subsection:

1. Firms showed a very well defined standard of entrance into foreign trade: a productivity premium, α, not lower than 20% in the year that preceded their debuts;
2. Although there is homogeneity in the pattern of their entrance, it is not possible to infer it from α, *ex ante*, which companies will flourish in the international market. It is important to mention that the expression flourish refers to the productivity gains resulting from their learning (*learning by exporting*);
3. Analyses of α, *ex post*, allow us to identify a specific permanence pattern of these companies in the exporting sector: (a) companies that have a loss of productivity as soon as they enter the export market will end up leaving it and will not be able to recover their previous levels of performance, the ones they had before their debut; (b) companies that present productivity gains immediately after their entrance will definitely not leave the international market, but what will ensure their uninterrupted permanence will be their learning gains (*learning by exporting*).

It is important to mention that the analysis presented throughout this section presents a few limitations; the non-control of the obliquity of the selection of the exporting companies is one of them. So we recommend caution in the interpretation of the results as being conclusive to corroborate the hypothesis of *ex-post* gains at their entrance in terms of the Brazilian economy. However, the results presented here are very similar to the results obtained by Araújo (2006).

The work developed by Araújo (2006) required a *survey* of the empirical literature on *ex-post* gains and it also provided results for the Brazilian economy. Two aspects that deserve to be highlighted in Araújo's work are: (i) as already mentioned, the results presented by the author are similar to ours, but the author uses *propensity score matching* techniques to deal with the selection problems that have not been overcome in this article; (ii) evidence of *ex-post* gains is stronger for the least developed economies.

The *rationale* for *ex-post* gains for new exporters in developing countries is provided by Blalock and Gertler (2004). The argument is that in less developed countries, where access to technology is more restricted and firms are farther away from the world's technological borders, exposure of the companies to the international market can, in fact, represent higher marginal benefits.

This argument is the main object of the analysis found in the following sections, in which the entrance dynamics and consequent differentiated absorption of technology are analyzed.

Empirical strategy and the treatment of self-selection in exporting activity

If exporting activity was random in our sample, estimating such average effects derived from exports would be very simple: one would only have to test *ex-post* the average differences for the variable supposedly affected by exports, for both groups of cases (composed by the companies that exported) and of controls (composed by those that did not export). This could be done by using a regression model such that the only explaining variable is a *dummy*, indicating if a company started or not in the international market.

However, as already discussed, our problem is not so simple because the companies that start operating in the international market do self-select themselves – and this has to be taken into account regardless of the methodology chosen. When the determinants of the self-selection are known and sufficient or in case we know which variables lead a company to enter the market, one can rely on the so called quasi-natural experiments. Basically, what this methodology does is to sophisticate the t-tests of the natural experiments in order to compare companies with similar characteristics.

Hence, in order to evaluate the existence of productivity, employment, turnover and *market share* gains after firms enter the international market and their relation to the technological standards of the companies, we are interested in regressions such that performance variables are regressed on their debut, variables that represent the technological standards of the companies, and the interaction of these variables with their entrance. So as to achieve this, we adopt the following empirical strategy:

1. As we know the variables that determine the self-selection of the companies that begin exporting, the first step consists in estimating the probabilistic model that estimates the probability of a company being a *rookie* in the international market. This model is fundamental for the control of the self-selection of the companies starting in the market in the steps below.
2. Having the estimated probabilities in hand, two conditions are necessary for the treatment of the self-selection: the balancing condition, which is related to the equality in mean tests of the variables used in the probabilistic model between the companies debuting in foreign markets and those that are not; and the common support condition, which means that, in the study, there are no neighborhoods of $p(X)$ in which one can only find *rookies* and not their *non-rookie* counterparts, and vice-versa.
3. Since we have 14 variables of sources of information and 7 innovative activities, we tried a dimensional reduction of these two groups of variables by principal component analysis. It is important to highlight that principal component analysis not only reduces the dimension of the explanatory variables of our models, but it also indicates different technological standards described by the components.

These components characterize the process of *learning by exporting* when it is present.

4. With these components in hand, the last step consists in estimating the regression models as follows:

$$performanc\text{æ}_i = \beta_0 + \beta_1 estreante_i + \mathbf{cpsources}_i + \mathbf{cpactivities}_i + estreante_i * \mathbf{cpsources}_i + estreante_i * \mathbf{cpactivities}_i + \varepsilon_i \quad (2),$$

in which $performance_i$ is the performance variable we are interested in, $estreante_i$ is the debut variable, $\mathbf{cpsources_i}$ is the group of components that represents the importance of the sources of information for innovation and $\mathbf{cpactivities_i}$ is the group of components that refers to the importance of innovation activities.

Nonetheless, this equation is weighed by the following weighting factor:

$$w_i = \frac{estreante_i}{p(X_i)} + \frac{1 - estreante_i}{1 - p(X_i)} \quad (3),$$

in which $estreante_i$ is exactly what was defined previously and $p(X_i)$ denotes the estimated entrance probability. Intuitively, the weights are such that bigger weight is given to the companies that did not start exporting though they had a high probability of doing it – counter-factual for those not starting in the market – and for the companies that started operating though they had a low probability – counter-factual for those not starting in the market. Technically, what we will do follows the *inverse probability weighting* (IPW) technique for the control of the selection bias.[9]

Hence, the *probit* model used for estimating the probabilities of a company starting to export – the first step of our empirical strategy – is described in Table 4. In it a debut in exporting in 2001 or 2002 is regressed on the logarithm of employees and productivity in the two previous years (in order to avoid endogeneity), on the age of the company, *market share* in the quadratic form (also in the two previous years) and sectorial controls. Generally speaking, the coefficients of the model are as expected, except for the productivity, which, conditional to the other variables, did not seem to influence the probability of starting in the market. Particularly, the size of the company measured by the logarithm of employees strongly influences the probability of exporting.

[9] Regarding this technique and its validity, please read Hirano and Imbens (2001) and Abadie (2005).

Table 4 –*probit model* for the probability of exporting

Variáveis independentes	Coef.	Estat. t
Log (PO) – before	0,817	6,02
Firm's age	-0,035	-3,65
Log (productivity) – before	0,031	0,32
Mkt Share – before	79,353	2,69
Mkt Share ^2	-670,871	-2,40
Factor intensity 2 (dummy)	0,639	2,16
Factor intensity 3 (dummy)	1,079	4,19
Constant	-6,104	-5,61
N	1044	
lrchi²	107,7700	
prob>chi²	0,0000	
pseudo R²	0,1298	
Log Likelihood	-361,27	

Source: Created by the authors based on the PIA, Pintec, Secex and Rais

The balancing condition was respected within five quintiles of the estimated probability, and the level of significance of the tests *t* was 0.005. The balancing condition implied in the limitation of the sample of the regression models for the companies whose value of *p(X)* was between 0.222 and 0.832 (328 companies).

Characterization of technological efforts and its influence over the performance of exporting companies starting in the market

In this section, we will try to characterize the patterns of the technological effort of the companies in our sample by using analytical procedures of principal component analysis (PCA). Thus, for each group of variables - 14 variables of sources of information and 7 innovation activities, described in Table 1 – a PCA based on the correlation matrixes of these variables was carried out, separately, in a way to produce orthogonal importance indexes among themselves, reducing, consequently, the dimension of the matrix of explanatory variables of the regression models in the following section.[10]

In the analysis of main components for the variables of sources of information, presented in Table 5 below, we selected four of its first thirteen components, responsible for approximately 58% of the total variation of the standardized variables. The components in this analysis are the following:

- The first component describes an index of general importance, in which for all variables a positive coefficient is designated, being, however, the five sources of information that are more strongly correlated to this vector are metrology and other centers of basic industrial technology, consulting companies, training centers, universities and research institutes, and

[10] The PCA aims at explaining the structure of variation and co-variation, or correlation, by means of a few linear combinations produced based on the original variables. It is the method based on the spectral decomposition of these matrixes, being the linear combinations carried out based on the estimated eigenvectors.

also information networks. That is, the main sources in this component are all outside the company and almost all of them refer to institutions that render technological services.

- The second component bears a bigger positive association with other sources within the company and with licenses and patents and consulting companies, and a negative association with sources of information from competitors, fairs and exhibitions, and information networks.
- The third index of importance of sources of information is a linear combination in which the sources that are more strongly correlated are the competitors, suppliers of machinery and equipment, and licenses and patents, while the main negative correlations are given by the sources of intramural R&D and exhibition fairs.
- Finally, the fourth index bears strong correlation with conferences, other sources of information inside the company, licenses, permits and patents, and fairs and exhibitions, while it has a negative correlation, mainly, with the sources of information that come from training centers, and clients, and consumers.

That is, while the first index reflects a general indicator, in which sources outside the company prevail, the other indexes represent combinations of internal sources and combinations of external sources not only compound of strictly technology companies.

For the PCA regarding the level of importance of innovation activities, three components were selected, being representative also of approximately 58% of the variation of the standardized variables. The components in this analysis are as follows:

- In a similar way to the results of the analysis carried out for the variables of sources of information, the first component reflects a general index of importance in which all innovative activities are positively correlated and with relatively high values, being the main three activities evaluated in this vector of intramural R&D, industrial projects and training.
- The second component, on the other hand, is an index that reflects the importance of the purchase of machinery and equipment, and training, and the other variables presented negative correlations.
- Finally, the third component expresses the importance of expenses with the introduction of technological innovations into the market that are, according to IBGE, commercial activities, directly linked to the release of a product that is technologically new or improved.

In the following section, these are the components that will serve as explanatory variables of the effects of *learning by exporting*.

Table 5 – Principal Component Analysis – Importance of Information Sources and Innovation Activities

Components	*Eigenvalues*	*Proporition*	*Variables*	*comp1*	*correlation*	*comp2*	*correlation*	*comp3*	*Correlation*	*comp4*	*correlation*
1	2.769	0.213	fpedi	0.247	0.410	0.157	0.190	-0.217	-0.244	0.121	0.127
2	1.472	0.113	fofi	0.142	0.237	0.434	0.526	0.043	0.048	0.373	0.391
3	1.265	0.097	ffmq	0.180	0.300	-0.135	-0.164	0.484	0.545	0.163	0.171
4	1.101	0.085	fcc	0.277	0.460	-0.142	-0.173	0.118	0.133	-0.436	-0.458
5	0.988	0.076	fconc	0.166	0.276	-0.288	-0.350	0.588	0.661	-0.074	-0.077
6	0.887	0.068	feci	0.352	0.585	0.266	0.323	0.023	0.026	0.158	0.165
7	0.846	0.065	fuip	0.348	0.579	0.241	0.292	-0.150	-0.169	-0.228	-0.240
8	0.726	0.056	fccp	0.351	0.584	0.135	0.164	0.173	0.195	-0.326	-0.342
9	0.700	0.054	fiec	0.378	0.629	0.170	0.207	-0.172	-0.193	-0.277	-0.290
10	0.638	0.049	fpat	0.072	0.119	0.320	0.388	0.424	0.477	0.332	0.348
11	0.567	0.044	fconf	0.310	0.516	-0.179	-0.217	-0.181	-0.203	0.398	0.417
12	0.535	0.041	ffeira	0.270	0.450	-0.419	-0.508	-0.216	-0.243	0.280	0.294
13	0.535	0.039	frii	0.312	0.519	-0.424	-0.515	-0.109	-0.122	0.131	0.138
N	3863	n. of comp.	13	trace	13						
1	3.490	0.698	epedi	0.515	0.962	-0.162	-0.179	-0.156	-0.073		
2	1.232	0.246	eaqmed	0.415	0.775	0.563	0.624	0.143	0.067		
3	0.218	0.044	etrein	0.443	0.827	0.504	0.559	0.069	0.032		
4	0.057	0.112	eimec	0.397	0.741	-0.520	-0.577	0.726	0.339		
5	0.003	0.001	eprji	0.458	0.855	-0.365	-0.405	-0.650	-0.304		
N	235	n. of comp.	5	trace	5						

Source: Created by the authors based on data from the PIA, Pintec, Rais and Secex.

Learning-by-exporting and technological standards

In this section, we show the results of the econometric analysis of the *ex-post* effects of the firm's debut in the international market over its productivity, turnover, employment and *market share*. In order to achieve this, we estimated a group of models, regarding the level of these variables in the period 2003-2005 depending on the company's debut in the international market. Hence, considering that *rookie* companies and non-*rookies* that find themselves in the region of common support have similar characteristics, in the models shown below we tested if the companies that started exporting are more productive, bigger or have a bigger *market share* in the period 2003-2005 than those that did not start exporting. It is important to highlight that all these equations are weighted by the weights described by equation (3), which aims at controlling the self-selection of the companies starting in foreign markets.

These considerations made, in Table 6, we can find the results of the regressions presented for the variables of the logarithm of turnover, productivity, logarithm of employees and for the variable of *market share* for the companies that respected the condition of common support (in our case 328 companies). These regressions have as explanatory variables, in addition to the distinctive debut *dummy* variable, the importance indexes of the sources of information and innovative activities presented previously and their interactions with the debut *dummy*.[11] Given the presence of several irrelevant variables in the models estimated and the reduced size of the sample, the level of statistical significance of 10% for the inference over the significance of the coefficients will be considered.

First of all, we have the result that the *rookies* that remain in the international market become bigger and more productive when compared to a group of companies that did not even start in the market. As before, the companies that start operating in the international market see their turnover and productivity grow, respectively, 105.8% and 47.2% - more than those that did not start in the period 2003-2005, and employ 40.7% more than those that did not start.[12] However, in spite of these gains, significant *market share* gains were not found. This result is compatible with the stylized facts found in international literature and in Araújo's results (2006) for the Brazilian case.

Regarding the technological components, in an interesting and congruent way, in all models – except in the model for the *market share* – the component regarding the importance of sources of information such as competitors, machinery and equipment suppliers, and licenses and patents is negative for the group with all the companies. However, this negative relationship seems to be valid only for the companies in the control group, that is, innovative companies that did not begin to export. The coefficients, in all models, for the interaction between the debut *dummy* and this component have positive signs with magnitude at least close to, or higher than, the coefficient that was not interacted. On the other hand, the second index of sources of information, associated positively to the other sources of internal information, licenses, permits, patents and consulting companies presents a positive and significant coefficient for the new exporters. That is, a differential as regards the importance attributed to these sources of information was observed. While the companies in the control group disregard a specific group of sources outside the company, the same does not happen with the companies in the treatment group; being even possible to verify that the range of sources of

[11] The authors are aware of the possible endogeneity in the relations established, being, therefore, this work a first exercise that should be interpreted carefully.

[12] Naturally, one should not compare these results in a quantitative way with those in section 3, since the sample and the debut variable differ a lot.

information of these companies is broader, which may possibly reflect in the quality of the technological effort made by these companies.

Regarding innovative activities, also in a general way (for all companies) a positive relation between the economic performance variables and the general index of importance of innovative activities with emphasis on intramural R&D is verified. Only in the model for the logarithm of the turnover the coefficient of this general index, when interacted with a the debut *dummy,* presented a negative coefficient of the same magnitude of the non-interacted one, demonstrating that this indicator is not capable of explaining the average turnover of the companies starting in the international market. Even so, it is possible to conclude that, on the whole, there is a positive relation between economic performance and this indicator of technological effort. What is different among the models estimated is the positive impact associated with innovative activities such as purchasing machinery and equipment and training over the variables that represent the size of the companies, that is, the logarithm of turnover and employment. It is important to remember that these activities are strongly related to innovation that is in progress, which denotes a positive relation between this form of innovation and the growth of companies.

Thus, what was done in this part of the work allows us to conclude that differentials in the impact of the technological behavior of new-exporting innovative companies and those that do not begin to export is in the broader collection of sources information used by the former, which goes in line with the hypothesis of *learning by exporting.*

Table 6 – Linear models for performance variables

	Log(turnover) - after		**Log(Productivity) - after**		**Log(PO) - after**		**Market share - after**	
	Coef.	T stat.	coef	T stat.	coef	T stat.	coef	T stat.
estreante	**0,722**	**2,789**	**0,387**	**3,243**	**0,342**	**3,243**	-0,001	-0,5681
pcf1	0,090	1,025	-0,006	-0,121	0,084	-0,121	0,000	0,3147
pcf2	0,026	0,187	-0,075	-0,841	0,034	-0,841	0,000	-0,0203
pcf3	**-0,346**	**-2,391**	**-0,213**	**-2,331**	**-0,186**	**-2,331**	-0,003	-1,0442
pcf4	0,126	0,872	0,011	0,141	0,083	0,141	0,002	1,5207
pcai1	**0,400**	**3,917**	**0,180**	**2,628**	**0,192**	**2,628**	0,002	1,7827
pcai2	**0,226**	**1,733**	0,052	0,647	0,201	0,647	0,000	-0,2621
pcai3	0,078	0,518	0,111	1,156	0,093	1,156	0,002	1,2394
Estr*pcf1	0,004	0,035	0,009	0,149	-0,007	0,149	0,002	1,2937
Estr*pcf2	**0,371**	**2,056**	**0,205**	**1,899**	**0,208**	**1,899**	0,004	1,4177
Estr*pcf3	**0,342**	**1,665**	**0,314**	**2,636**	**0,119**	**2,636**	0,009	1,7790
Estr*pcf4	-0,053	-0,217	-0,072	-0,675	0,017	-0,675	-0,002	-0,7242
Estr*pca1	**-0,411**	**-3,023**	0,028	0,343	-0,391	0,343	-0,001	-0,4680
Estr*pca2	0,104	0,533	-0,113	-1,067	0,213	-1,067	0,001	0,8971
Estr*pca3	0,204	0,882	-0,042	-0,311	-0,043	-0,311	-0,001	-0,3820
_cons	16,438	93,000	10,290	121,104	5,166	121,104	0,007	3,7985
	N	328	n	327	n	328	n	328
	f(15,312)	6,3400	f(15,311)	4,5400	f(15,312)	5,7500	f(15,312)	1,9900
	p>F	0,0000	p>F	0,0000	p>F	0,0000	p>F	0,0154
	R^2	0,2417	R^2	0,2261	R^2	0,2623	R^2	0,1920
	Root ME	1,5877	Root ME	0,8303	Root ME	0,9730	Root ME	0,0170

Conclusions and extensions of this article

In this article, we initially show that there are productivity differentials between firms that begin exporting and those that never export in Brazil in the first years of this decade. These productivity differentials are not only *ex-ante* – which goes in line with the literature – but also *ex-post*, in accordance to the learning-by-exporting hypothesis. Additionally, we provide some evidence that the productivity gains are even larger for those firms that remain as exporters for more time.

However, instead of accepting the hypothesis of technological *learning by exporting* as the explanation for these differentials of competitiveness, we related the existence of these *ex-post* gains to variables of the Pintec 2003-2005 relative to the sources of information for innovation and to the importance of innovative activities. Nonetheless, this exercise had to be restricted to a small sample of innovative firms in the period 2001-2005.

We concluded that companies that start operating in the international market present a pattern of knowledge accumulation that is significantly different from companies that never start exporting: while, for both types of companies, innovative activities are, in general, (but with emphasis on intramural R&D) important, for companies that start operating in the international market the scope of sources of relevant information for innovation is bigger, including other sources within the company, licenses, patents, consulting companies, competitors and suppliers of machinery and equipment. In spite of the limitations of the exercise carried out here, the hypothesis of *learning by exporting* resulting from better access to input and imported equipment, bigger exposure to competition and the possibility of technological cooperation with other companies of the productive chain seems to be true in Brazil.

In future versions of this article, we intend to advance in three directions:

(i) To apply econometric techniques to correct the panel attrition between both rounds of the Pintec (2001-2003 and 2003-2005), as well as to extend the conclusions of the study also to non-innovative companies, and capture the effect that exporting can have over the decision to innovate (in this case, the greatest difficulty is that non-innovative companies do not answer the part of the questionnaire that refers to the variables which we are interested in – then, we have a *missing variable problem*);

(ii) To treat a possible endogeneity between technological standards and performance variables, besides creating estimates that are not sensitive to fixed effects (estimates in first difference);

(iii) To incorporate into the analysis of technological standards the difference between sources of information and/or national and foreign cooperation.

Bibliography

ABADIE, A. Semiparametric Difference-in-Differences Estimators. *Review of Economic Studies* 72, p. 1-19, 2005.

ARAÚJO, B. C. P. O. Análise empírica dos efeitos e*x-post* das exportações sobre a produtividade, o emprego e a renda das empresas brasileiras.. In DE NEGRI, J., DE NEGRI, F. e COELHO, D. *Tecnologia, exportação e emprego.* Brasília, IPEA, 2006.

AW, B. Y., HWANG, A. R. Productivity and the export market: a firm-level analysis. *Journal of Development Economics*, v. 47, p. 313-332, 1995.

BERNARD, A. B., JENSEN, J. B., Exceptional exporter performance: cause, effect, or both? *Journal of International Economics*, v.47, p.1-25, 1999.

BLALOCK, G., GERTLER, P.J. Learning from exporting revisited in a less developed country. *Journal of Developmenr Economics*, v.75, p.397-416, 2004.

CLERIDES, S., LAUCH, S., TYBOUT, J. R. Is learning by exporting important? micro-dynamic evidence from Colombia, Mexico and Morocco. *The Quarterly Journal of Economics*, v. 113, n. 3, p. 903-947, 1998.

COSTA, G., ALVES, P., BITTENCOURT, M., ARAÚJO, K. e DOYLE, H. *Avaliação de uma Proxy para a Idade da Firma Utilizando Amostragem Complexa.* Texto para Discussão 1240, IPEA, 2006.

ELLERY JR, R. e GOMES, V., *Perfil das Exportações, Produtividade e Tamanho das Firmas no Brasil*. Texto para Discussão 1087, IPEA, 2005.

HIRANO, K. e IMBENS, G. Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services & Outcomes Research Methodology* 2, p. 259-278, 2001.

TYBOUT, J. R. Plant and firms level evidence on "new" trade theories. *In*: CHOI, K., HARRIGAN, J. (Eds.) *Handbook of International Trade*, p. 1.243-1.278, Oxford: Basil Blackwell, 2003.