in Proc. 9th Int. Workshop on Requirements Engineering and Law (RELAW 2016). Beijing, China. Sep. 2016. © IEEE

The Influence of Tax Incentive and Market Regulation Requirements on IT Companies: Empirical Evidence from the Brazilian Industry

Carlos Henrique C. Duarte
Brazilian Development Bank (BNDES), Av. Chile 100, Rio de Janeiro, RJ, Brazil
cduarte@bndes.gov.br & carlos.duarte@computer.org

***Abstract*—The effectiveness of public policy measures to foster the competitiveness of Information Technology (IT) industries has been intensively debated nowadays, since a multitude of policy instruments can be applied to different business segments and time frames, with varying levels of legal security and expected results. In this paper, we investigate the influence of some tax incentive and market regulation requirements on Brazilian IT companies, using a data set containing 942 firms. We perform empirical studies concerning the influence of two public policy measures on the labor productivity of such companies, estimated through their annual gross revenue per worker ratio. We show that there has been a positive correlation of electronic tax recognition and bookkeeping regulation measures with the labor productivity growth of locally owned suppliers of enterprise resource planning software. Likewise, we show that, on average, the productivity growth of foreign capital hardware manufacturers, while payroll tax reduction benefits were in force, was higher than in other periods. These studies illustrate the temporal and business segment influence that technology requirements might have on the productivity of IT companies when considered in public policy measures and also suggest a future research agenda.**

***Index Terms*—Tax Incentives; Market Regulation; Public Policy; Legal Requirements; Information Technology.**

I. Introduction

Information Technology (IT) was early recognized as a possible source of social and economic development in more mature and less developed economies [2]. The entry barriers in this business sector appeared to be low, requiring from local governments just investments in human resource formation and basic computation / communication infrastructures, apart from guarantees of friendly business environments and the availability of financial support. With their relative competitiveness insured, private companies would be able to explore their local markets and even export [15]. Since that time, however, the competition has been increasingly fiercer, some demanding success factors have been elicited from practice [4] and global consulting companies have even devised specific products to better advise IT clients that have worldwide footprints [7].

From the public policy maker point of view, it would be tempting to propose and implement measures aiming to make local IT companies more competitive, considering their harsh competitive environment. (The effectiveness of) this kind of measure, however, has been intensively debated nowadays, since a multitude of policy instruments can be applied to different business segments and time frames, with varying levels of legal security and expected results. Such instruments can be regarded to be of social nature, such as to rise the education level of IT workers and improve their language skills, or of economic character, for example to improve the business environment by deregulation and the existing financial instruments by making them more accessible to IT companies.

It happens that public policy measures based on social instruments tend to produce more intangible results and mature in the long term, whereas those based on economic instruments are usually more tangible and of short term maturation. Consequently, the latter receive more attention and criticism. Policy makers are frequently required to answer some recurrent questions: Is there an underlying stable legal framework? Are the instruments allowed by international agreements? What is their comparative performance? Is there any (empirical) evidence of their effectiveness? What is their (temporal) influence on specific economic sectors and segments?

In this paper, we focus on the study of specific market regulation and tax incentive instruments hoping to find answers to the last two kinds of questions. We understand by market regulation any measure that leads to constraints on the economic behavior of persons and companies, and by tax incentive any measure that fosters the behavior of such actors in particular ways. Examples are respectively safety regulation norms issued by public health care surveillance agencies and fiscal legislation produced by city, state or central governments in order to change taxation bases. Both kinds of instruments may produce effects on the productivity of IT companies, consequently affecting their competitiveness.

We investigate here the influence of some tax incentive and market regulation requirements on Brazilian IT companies. We study the consequences to these companies of electronic tax recognition and bookkeeping regulation measures, as well as payroll tax reductions. The former were a result of a series of legal norms to improve the Brazilian excise tax collection system, whereas the latter were a result of discussions between the central government and IT entrepreneur associations. We

†The assumptions, views and opinions in this paper are solely those of the author and do not necessarily reflect the official policy, strategy or position of any Brazilian government entity.

study the influence of these two public policy measures on the competitiveness of such companies, by observing variations in their labor productivity, which is calculated here in terms of the annual gross revenue per worker ratio.

We perform empirical studies using a large data set, containing the revenues, employment and calculated labor productivities of 942 Brazilian IT firms. We show that there has been a positive correlation of electronic tax recognition and bookkeeping regulation measures with the labor productivity growth of locally owned suppliers of enterprise resource planning (ERP) software. Likewise, we show that, on average, the productivity growth of foreign capital hardware manufacturers, while payroll tax reduction benefits were in force, was higher than in other periods.

From the Requirements Engineering point of view, the rationale for performing these studies is two-fold. They illustrate the temporal and business segment influence that technology requirements might have on the productivity of IT companies when considered in public policy measures. In addition, they lead us to wonder which requirements should be posed so as to attempt to maximize public policy measure effects. This view suggests a future research agenda, which is also proposed herein.

This paper is organized as follows: Section II addresses related work; Section III describes our research universe, which consists of tax incentives (III-A) and market regulations (III-B) that benefit the Brazilian IT industry, apart from its competitiveness measurement (III-C); Section IV presents our data sets and research methodology; Section V contains our data analyses and research findings on the correlation of market regulations and tax incentives with Brazilian IT company productivities; Section VI presents the threats to the validity of our work. We conclude the paper discussing our findings and suggesting future research.

II. Related Work

Given the social, economic and professional character of our research, it is not easy to perform comparisons with related work. Indeed, the corresponding research subjects appear to be relatively unexplored in Computer Science: Only in 2012 the ACM introduced a new branch unifying all these topics under the same root in its Computing Classification System [10].

Nevertheless, we identify some similarity of our work with the empirical studies on the impact of income tax exclusions and deductions reported in [3] using a microfile simulation model, which was constructed in terms of variables observed in a large sample of households. That simulation model has been used to advise the Wisconsin state government and legislature about the effects of changes in the tax policy. More recently, model based simulations of income tax have been used to advise the Luxembourg government on planned tax law reforms [21]. They are formulated in terms of policy and domain UML models, which are used to produce simulation results with the support of data and simulator generators. Although these studies also adopt empirical methods and are based on large data sets, there are two main differences in relation to our work: they are prospective rather than retrospective and use sets of software tools specifically developed for their purposes, instead of using packaged software.

A distinct but related approach is proposed in [16] and [17], which adopt measurement techniques to study the attained levels of compliance with current and proposed goal oriented regulations, such as those found in public policy statements. In [17], a measurement framework inspired by the ISO/IEC 15504 standard is proposed for modeling regulation compliance. This is illustrated by studying the elicitation and analysis of compliance in business process management applications of the financial sector. In [16], automation techniques are proposed to transform outcome based natural language regulatory texts into goal models, which allow some formal reasoning via goal model evaluation. These are illustrated by taking examples from the banking domain. The main distinction in relation of our work, apart from the adoption of specific purpose software tools, is that this approach attempts to represent and measure compliance with goal oriented regulations, whereas ours (as well as the work mentioned in the previous paragraph) is more concerned with the effects of traditional perspective norms. We also study the particular effects that technology requirements might have when considered in regulations, whereas the related work is strictly concerned with legal requirements. We believe that there can be mutual benefits of adopting both approaches in studying public policy norms, if they are used in a coordinated manner.

III. Brazilian IT Market Issues

A. Tax Incentives to the IT Industry

Some tax incentives have been devised so as to attempt to influence directly the Brazilian IT industry and consequently ensure its competitiveness. In general, they result from collective and gradual discussions between government agencies and entrepreneurial associations. Among these, perhaps the most enthusiastically greeted ones were payroll tax reductions that begun benefiting software service companies, encompassing from single software development services and consulting provision to full business process outsourcing activities[1], and later on were expanded to cover almost all the IT sector.

The initial rationale for implementing payroll tax reductions was the genuine desire of the country to be better positioned to compete for the global software development demand by insuring the relative affordability of hiring the required workforce for software development activities. Therefore, a special tax regime for software companies that export (REPES) was created in 2005, based on a change in the social benefit taxes paid by software companies. They were originally calculated in terms of employee wages, but have been modified to be computed in terms of company gross revenues. Subsequently, trade unions of IT workers and employers noticed that the widespread adoption of the same social benefit recognition methodology could end

[1]For a precise definition of software service companies, see [20], which studies software factories and related service provision structures, based on interviews, questionnaires, software metrics and software process assessments performed in a single company.

arduous discussions between IT companies and their employees aiming to make the allowed hiring methods more flexible. With the agreement of the Ministry of Labor, the benefit was generalized in 2011 to all kinds of software companies. Subsequently, with the 2012 economic crisis eroding the competitiveness of local companies, as an anti-cyclic measure, the benefit was further generalized to cover hardware sellers / manufacturers, as well as companies of other economic sectors. The PTR benefit was substantially reduced at the end of 2015, due to changes in the central government tax priorities.

We summarize the respective legislation as follows:

- **Law 11.196/2005 (Law of Good or LoG):** Establishes the special tax regime for software companies that export (REPES), granting additional benefits to local research and development activities, as well as to intellectual property protection measures abroad;
- **Law 11.774/2008 (LoG regulation):** Allows reductions in some payments of social benefits in proportion to the gross revenues obtained from exports;
- **Law 12.546/2011 (Payroll Tax Reduction or PTR):** Allows any software company to pay as a social benefit tax 2.5% of its gross revenues;
- **Decree 7.828/2012 (PTR regulation):** Defines some additional procedures for companies benefited by the PTR Law, such as to produce accounting records according to electronic bookkeeping procedures, and expands the scope of LoG to companies of selected sectors and manufacturers / re-sellers of certain products, making the gross revenue percentage paid as social benefits vary accordingly;
- **Law 1.3161/2015:** Cancels the PTR benefit to commerce and industrial companies and rises the percentage of the gross revenues paid as social benefits by other sectors;

As can be noted, more and more companies were included in the scope of the PTR benefit, something that made it difficult to understand and evaluate. However, due to the formulation of the legislation in terms of the Brazilian Corporate Tax Payer Registry Unique Identification Number (CNPJ) of each company, its main business code (according to the Brazilian National System of Economic Activity Classification – CNAE) and the Harmonized System Code of the benefited products (according to the Mercosur Common Naming Scheme – NCM, a derivation of the World Commerce Organization – WCO harmonized system), this subject became tractable with the aid of some automated support, as we shall show in Section IV.

B. Market Regulations of other Economic Sectors

The origins of the Brazilian legislation on electronic tax recognition and bookkeeping date back to the 1980s, just before the institution of the Brazilian Real (the currently adopted currency in Brazil) by the Federal Government, ending a period of currency value oscillations and high inflation. The desire to implement a rigorous tax collection policy led the central government to establish agreements with state governments to unify the product and service excise tax (ICMS, the Portuguese abbreviation) collection system adopted in the country, sharing rights and obligations with all the involved entities.

The preceding period events brought important consequences to the local IT industry. Since that time, multicurrency and highly parameterized systems are produced in the country, avoiding rework due to economic policy changes, for example. Moreover, the tax legislation itself established strict requirements for commerce automation hardware and software, which were originally of optional enforcement, but have been made compulsory later on, ensuring high compliance with taxation rules and efficiency in tax collection procedures. Clearly, this legal framework explicitly involves both manufacturers of commerce automation equipment and ERP companies, which normally provide packaged and customized software, apart from software services[2].

We summarize the respective legislation as follows:

- **ICMS Agreement 44/1987:** Specifies requirements and rules for accreditation and use of Point of Sale (PDV in Portuguese) terminals (comprising both hardware and embedded software), as well as for fiscal document contents and controls in the country;
- **ECF Agreement 01/1998:** Regulates the adoption of Tax Coupon Issuers (ECF, a kind of PDV equipment) in substitution to the invoice printers previously used in product and service sales (compulsory usage in the whole country from 2001 onwards);
- **SITEF Adjustment 07/2005:** Establishes a protocol for numbering, printing, communicating, processing and substituting Electronic Invoices (NF-E) presented to tax collection authorities and other related institutions (the decision as to make their usage compulsory was delegated to state governments, which nowadays have all enforced the norm);
- **ICMS Agreement 143/2006:** Establishes Digital Bookkeeping Procedures (EFD) in substitution to the production of traditional printed accounting books (compulsory in the whole country from 2016 onwards for small companies and from 2009 onwards for other companies);
- **Decree 6.022/2007:** Establishes a Public Digital Bookkeeping System (SPED) cooperatively managed by the central and state governments (which is of compulsory usage by all those entities that wish to validate their fiscal electronic documents, which in turn are required to adopt digital certification);

In order to illustrate the technical requirements captured in the aforementioned market regulation norms, we present in Table I a selection of these requirements. In the first column, the requirement identifier is presented, formed using the norm name appended to the respective article, item and paragraph numbers. In the second column, we classify each requirement according to its focus in the corresponding system, software or development process. In particular, system requirements may be related either to business rules or hardware specifications required in enforcing these norms. We also classify the type

[2]For a precise definition of software product companies, see [19], which differentiates the packaged software from the custom software business based on three case studies with data gathered from interviews, surveys, observations and archival records of software development teams of many companies.

TABLE I
SELECTED TECHNICAL REQUIREMENTS IN THE BRAZILIAN ELECTRONIC TAX RECOGNITION AND BOOKKEEPING LEGISLATION.

Id	Focus	Functionality	Statement
(PDV3)	System	Non-Functional	A PDV equipment will contain, at least:
(PDV3.2)	System	Non-Functional	an issuer of fiscal coupons and/or invoices
(PDV3.3)	System	Non-Functional	an issuer of analytic listings
(PDV3.8)	System	Non-Functional	a serial manufacturing number stamped into or an identification plaque attached to its surface
(PDV3.14)	System	Non-Functional	a numbered protection device to insure equipment inviolability, preventing any violation without evidence
(PDV3.20)	System	Non-Functional	a fiscal memory in PROM or EPROM with capacity to store 1825 daily accumulated gross sale values and their date and time, restart counter, equipment serial number and the issuer federal, state and city taxpayer ids
(PDV3.P5)	Software	Operational	Independently of the issued document type, the equipment transaction order number shall start from 1, be sequential and increasing
(PDV3.P16)	Software	Non-functional	In case of failure, disconnection or fiscal memory exhaustion, this event shall be immediately detected and informed in a message by the equipment, which will stay blocked, except from reading the fiscal memory
(PDV3.P19)	Software	Operational	Access to the fiscal memory is restricted to basic software, under manufacturer responsibility
(PDV4)	System	Non-Functional	The equipment will not have key, device or function that:
(PDV4.1)	System	Non-Functional	prevents the issuance of tax documents or printing analytic listings
(PDV15)	Software	Informational	Each invoice will contain at least the following data:
(PDV15.2)	Software	Informational	the transaction order number
(PDV22)	Software	Informational	In sales wherein goods are picked up by the buyer, invoices may be replaced by fiscal coupons, containing:
(PDV22.2)	Software	Informational	issuer name, address, federal and state taxpayer ids
(PDV22.3-4)	Software	Informational	issue date in the format (day, month, year) and transaction order number
(PDV26)	Software	Informational	The equipment should be able to print analytic listings reporting the issued tax documents and others
(ECF2)	Software	Informational	For proving operating costs and expenses, under the ICMS legislation, ECF issued documents must contain:
(ECF2.1)	Software	Informational	the federal taxpayer id of the acquiring person or company
(ECF2.2)	Software	Informational	the description of the goods or services object of the transaction, even if summarized or through codes
(ECF2.3)	Software	Informational	transaction date and value
(NF-E3a)	Process	Non-Functional	Each NF-E must be issued based on the proper layout by means of software developed or acquired by the taxpayer or made available by tax administration authorities.
(NF-E3b)			The following procedures apply to issued NF-Es:
(NF-E3.1)	Process	Non-Functional	the digital file of the NF-E shall be prepared according to the XML standard
(NF-E3.2)	Software	Operational	the numbering of NF-Es is sequential and must be restarted when the one billion limit is reached
(NF-E3.3)	Software	Informational	each NF-E shall contain a code generated by the issuer to compose an access key, along with the issuer CNPJ and the serial number of the NF-E
(NF-E3.4)	Software	Informational	each NF-E must contain the issuer digital signature to ensure the authorship of the digital document
(NF-E3.5)	Software	Informational	the identification of the goods in a NF-E shall contain their code, set according to the NCM
(NF-E4)	System	Functional	The digital file of the NF-e can only be used as a tax document after:
(NF-E4.1)	System	Functional	being transmitted electronically to a tax administration authority
(NF-E4.2)	System	Functional	have their use authorized by the respective protocol
(NF-E5)	Software	Operational	The digital NF-E file transmission must be made via the Internet, through security protocols or encryption, using software developed or acquired by the taxpayer or made available by tax administration authorities
(NF-E10)	System	Functional	The issuer shall keep the digital NF-E file under its own guard and responsibility, for the period established in legislation, even outside the company. It shall be made available to tax authorities whenever requested
(NF-E10.1)	System	Functional	the recipient should check the authenticity and validity of the NF-E and the existence of authorization
(EFD1.1)	System	NA	An EFD is considered valid for tax purposes only after the confirmation of reception of the respective file
(EFD1.2)	Software	Operational	The reception and validation of an EFD will be held in SPED
(EFD2)	Software	Operational	The EFD file must be digitally signed by the taxpayer or its legal representative
(EFD7)	System	NA	EFD bookkeeping replaces the standard bookkeeping and printing procedures of the following books:
(EFD7.1-5)	System	Functional	Inputs and Outputs Registration, Inventory Record, and Production and Excise Tax Calculation books
(SPED2)	System	Functional	SPED unifies the reception, validation, storage and authentication of books and documents, which are part of the accounting and bookkeeping of each company, using a single computerized flow of information
(SPED4)	System	Functional	Access to SPED information shall be granted to its members, without prejudice to compliance with commercial, tax and banking secrecy legislation

of each requirement as non-functional or functional, the latter ones segregated into informational and operational.

As can be noticed, the Brazilian electronic tax recognition and bookkeeping legislation heavily relies on IT and, as a side effect, generates a proper environment for the growth of local companies exploring the commerce automation and ERP businesses, which is analyzed in the sequel.

C. Competitiveness Measurement of IT Companies

Specific market regulations and tax benefits may be important to insure the competitiveness of IT companies. The adoption of such public policy instruments tend to mature in the short term and consequently requires effective evaluation metrics. One way to satisfy this requirement is to measure how the production capability of each company is affected, something normally modeled using production functions, which explicitly relate, in a time relying manner, dependent variables (called outputs) to independent ones (called inputs). These functions can be applied to distinct objects, such as individuals, business units, companies, economic sectors and even national accounts. Therefore, they are an adequate instrument for measuring individual effectivity and assessing comparative performance.

In a seminal work concerning the productivity of R&D intensive companies [8], the following formulation of production functions was proposed, based on the classical multiplicative model of time series:

$$Q_{o_t} = \beta_0 e^{\lambda t} K_{o_t}^{\beta_1} C_{o_t}^{\beta_2} L_{o_t}^{1-\beta_2} \text{ where:} \quad (1)$$

- Q: Economic output (e.g. sales, revenues, value-added);
- C: Capital input (e.g. machinery, facilities);
- L: Labor input (e.g. salaries, benefits, labor taxes);
- K: $\sum_i w_{o_i} R_{o_{t-i}}$, a measure of accumulated and still productive R(&D) capital, where R_o measures deflated gross investments and w_o their connections to an object o knowledge;
- λ: Rate of disembodied technical change external to all the studied objects;
- β_j: Constants (for $0 \leq j \leq 2$);

Independent variables that reflect tangible resources (C), such as investments in capital goods, are admittedly arduous to treat, due to the effect of physical or economic processes such as depreciation and inflation happening in each analyzed time period. In turn, the variables that capture intangible resources (K), such as intellectual property, although also delicate to be treated, may help explaining why sometimes increases in capital and labor inputs are not reflected in productivity growth, for example due to technological lags or debts.

We believe that the study of IT company productivities (consequently their competitiveness) should be performed based on variables which are intuitive both to policy makers and to the whole industry, not only due to their familiarity with the adopted variables, but also because of a simple productivity function formulation. Moreover, variables for which publicly available observations are at hand should be preferred, since these may be subject to third party validation. Finally, depending on the observational and comparative nature of a study, between individuals or groups of companies in specific time periods, it is not even necessary to care about currency fluctuations, inflation, depreciation or other factors that equally affect the whole population under investigation. These requirements guide us in the choice of independent variables and in the formulation of specific production functions.

The tradition in the IT sector is to adopt, in the proposition of productivity functions, restricted versions and specific interpretations of the general definition in (1). In particular, investments in capital goods tend to be negligible in IT companies when compared to research, development and labor investments. This means that, in the IT industry, β_2 is typically presumed to be nearly zero, resulting in the following simplified definition:

$$P_{o_t} = \frac{Q_{o_t}}{L_{o_t}} = \beta_0 e^{\lambda t} K_{o_t}^{\beta_1} \quad (2)$$

Since labor has paramount importance in research and development intensive companies, typical production functions adopt the cost of labor, worked hours or employed personnel as inputs (L_o). The number of produced items (such as manufactured equipment or lines of code) is normally used as a measure of output (Q_o). However, the mixture of economic and technical independent variables seems to be counter intuitive, since they belong to different abstraction levels.

Here we adopt the total number of company workers (possibly comprising trainees, employees and some shareholders) reported by each company at the calendar year end as the measure of input. This is an objective and consistent metric that allows us to compare distinct companies. Indeed, it would be possible to further specialize this metric taking into account in our analysis the skills, wages, location and other personal aspects of company workers, but we choose to adopt this simple and general formulation here to facilitate the development of our study. As a measure of output, we adopt the annual gross revenues of each company, which is also an objective and consistent metric, since every company is obliged to produce yearly financial statements presenting this kind of data according to generally accepted accounting principles. In this way, we adopt the gross revenue per worker ratio as a measure of IT company productivity. This is usually called labor productivity in the literature.

The main advantage of adopting labor productivity in relation to other corporate productivity measures is that it is easy to compute and understand. Other measures – such as those based on value added labor, capital labor, strict capital and multiple factors – can be easily misinterpreted, are more difficult to calculate or suffer influence from factors that are not always explicit[3].

IV. Research Data & Methodology

A. Research Universe & Data Collection

The main subject of our research are companies established in Brazil. We organize our observations regarding such companies in a data set containing company revenues and employment. We stratify our data set and analyses according not only to the main business segment of each company, determined from its main source of revenues, but also to the main source of invested capital.

We use in our research a data set containing 942 Brazilian IT companies. Among these, 687 are software companies and 255 are hardware companies. This seems to be a quite representative sample, considering the total number of such companies existing in the country at the end of 2014: the Brazilian Association of Software Companies (ABES) estimated in 6.994 the number of software companies [1], divided in 3.642 software product and 3.352 software service companies; and the Ministry Science, Technology and Innovation (MCTI) reported the existence of 562 hardware manufacturers in the country benefited by some tax incentive [13]. Unfortunately, neither ABES/MCTI nor any other representative institution of the sector reports a break down of their estimates on the Brazilian IT company population according to the main origin of company capital.

[3]For a detailed discussion on the advantages and disadvantages of adopting specific productivity measures, the reader is referred to the extensive study of the OECD [14] on this subject

This data set was constructed collecting employment and revenue data from multiple sources. The primary data sources were the annual financial statements of each company. In Brazil, large size companies listed in the Bovespa Stock Market upload their financial statements into the Internet site of CVM[4]. Moreover, most large and medium size companies in Brazil are obliged to publish their financial statements in large circulation newspapers. For those cases in which the author did not have direct access to the corresponding financial statements, secondary sources were used. In those cases, revenue and employment data were extracted from annual publications of market research institutions to which such data were directly disclosed by the studied companies [22], [6], [9].

As a selection criterion for inclusion of an observation in our data set, we required third party validation. This kind of validation was performed by diverse entities, such as external auditors or board members, which approved the annual financial statements analyzed in our research, and trusted market research institutions, which perform varying levels of cross checking.

B. Data Correction & Adjustment

Our data set is organized in terms of the CNPJ number of each company and also contains company name, main business code (according to CNAE) and main origin of invested capital (local or foreign). In many cases, these data were collected from secondary sources and were clearly incorrect, having to be manually fixed by the author in queries to the Brazilian Inland Revenue databases.

It happens that there is no publicly available source containing revenue and employment data that can be used for cross checking secondary source data, mainly due to public trade and tax secrecy policies. In those cases that divergent data concerning the revenue or employment of a company were collected from different secondary sources, just the smallest figures were admitted in our data sets.

Our data sets also contain categorical data, for instance the main business segment of each company, determined from the CNAE code of the company stored in Inland Revenue databases. Again, whenever the main source of company revenue was not correctly reflected in the corresponding business segment code, the latter was manually adjusted by the author. This kind of data adjustment process, as well as the subsequent filtering, computation and analysis, were all performed with automated support provided by spreadsheets [12].

We are aware that the aforementioned treatments raise important threats to the validity of our work, which are analyzed in detail in Section VI.

C. Data Classification & Filtering

After collecting, correcting and adjusting data, we performed a data classification and temporal filtering process, due to our interest in studying the temporal correlation of tax incentive and market regulation requirements with the labor productivity of IT companies in specific periods.

[4]CVM is the Brazilian Securities and Exchanges Commission.

The main business code of each company was used to create categories of software product, software service and hardware manufacturing companies. Moreover, companies were classified as foreign or locally owned based on their main origin of invested capital.

As a result, the software companies in our data set were further divided into 210 product and 477 service companies. In addition, we also determined that 77.5% of the software product companies were locally owned, whereas 80.5% of the service companies had most of their invested capital from a local origin. We also found out that 66.4% of the hardware companies were locally owned.

It is important to mention that we also adopted a temporal filter in treating our data. Although our time series on economic data begin in 1990, we chose to conduct our analyses just in the period from 2004 to 2014 for two reasons: Due to the exchange rate of the Brazilian Real in relation to the American Dollar (adopted in our revenue time series for past comparability reasons), which was relatively stable in this period, and due to the availability of data on the respective company populations collected by their representative institutions.

D. Derived Data Computation

Since our data set is sparse, in the sense that there are some missing observations in the middle of some periods, we used interpolation in order to estimate interior points in the curves of revenues and employment of each company, based on the points already present in each corresponding time series. Nearly 13.0% of our data on revenues were computed in this way, whereas 23.7% of our employment data was a result of interpolation. The labor productivity ratio of each company was calculated for all those years in which there were original or computed simultaneous observations of revenues and employment.

The decisions taken in the computation of derived data also threaten the validity of our work and, due to this fact, we analyze them in detail in Section VI.

V. Data Analysis & Research Findings

Now we study the relation of the market regulation and tax benefit requirements described in Section III to labor productivity, considering the population of Brazilian IT companies. As we have already mentioned, we conjecture that the inclusion of companies in the scope of such public policy measures contributes to increase their labor productivity growth. Given their distinct formulations, each public policy instrument requires a slightly different empirical study design.

Concerning electronic tax recognition and bookkeeping regulation measures, the affected companies are essentially manufacturers of commerce automation equipment and ERP companies. Considering this, we identified such companies in our data set and performed statistical tests comparing their labor productivity growth with those of other companies. We took into account the entire period from 2004 to 2014 to compute productivity growth ratios, since the respective regulations were enforced in this whole period. Consequently, the following hypothesis was tested in each of the four partitions of our data

TABLE II
RESULTS OF WELCH'S T-TEST ON THE AVERAGE LABOR PRODUCTIVITY GROWTH OF BRAZILIAN IT COMPANIES.

NATURE OF BUSINESS	ORIGIN OF CAPITAL	Benefited by Market Regulations			
		$n_1 + n_2$	t-value	DoF	p-value
A. Hardware	1. Local	6 + 41	−65.7849	2	1.0000
Products	2. Foreign	0 + 31	N.A.	N.A.	N.A
B. Software	1. Local	25 + 50	8.3101	44	0.0000
Products	2. Foreign	4 + 15	−31.7327	3	1.0000
C. Software	1. Local	8 + 145	46.7346	8	0.0000
Services	2. Foreign	2 + 53	−57.1857	53	1.0000

TABLE III
POWER OF WELCH'S T-TEST ON THE AVERAGE LABOR PRODUCTIVITY GROWTH OF BRAZILIAN IT COMPANIES.

NATURE OF BUSINESS	ORIGIN OF CAPITAL	Benefited by Market Regulations		
		X_1 / X_2	S_1 / S_2	Power
A. Hardware	1. Local	1.0029 / 1.0118	0.0002 / 0.0006	100.0%
Products	2. Foreign	N.A.	N.A.	N.A.
B. Software	1. Local	1.0122 / 1.0115	0.0012 / 0.0011	100.0%
Products	2. Foreign	1.0044 / 1.0181	0.0002 / 0.0001	100.0%
C. Software	1. Local	1.0275 / 1.0120	0.0003 / 0.0003	100.0%
Services	2. Foreign	1.0035 / 1.0152	0.0000 / 0.0002	100.0%

set, that is, considering the main business segment and capital origin of each company, generating six statistical test results:

(HYP1) On average, the labor productivity growth of companies in the scope of electronic tax recognition and bookkeeping regulation measures is higher than in other companies;

Since our productivity data set is skewed, we apply a logarithmic transformation to approximate a normal distribution. Next, we adopt the one tailed Welch's t-test for two independent samples with unequal sizes and variances to attempt to validate **HYP1**. The t-test suggests as a null hypothesis that the difference between the average values observed in the respective populations is equal to zero. In the context of our analyses, this means that the average labor productivity growth of companies in the scope of the studied market regulations is equal to that of other companies.

We present in Table II the inputs and computed results of our statistical tests. Each sample partition, with size $n = n_1 + n_2$, is divided in two sub-partitions, the first containing the productivity growth of companies in the scope of the studied regulation and the second one containing the productivity growth of companies that are not. For example, our sample contains 47 observations from companies that produce hardware and have a local main origin of invested capital (first line in the table, labeled with $A.1$), of which 6 are included in the scope of the studied market regulations and 41 others that are not. The respective t-values are computed taking into account the sub-partition sizes (n), as well as the labor productivity growth averages (X) and variances (S) of each sample sub-partition, in the same way that the varying degrees of freedom (DoF) numbers are computed. The resulting p-values are determined from the corresponding t-values with the computed degrees of freedom using the t-distribution.

Before drawing any conclusion from the data presented in Table II, let us investigate whether or not there is any lack of capability of the reported tests to find positive conclusions, if those were the case. In order to clarify this, we study the statistical power of our tests, which corresponds to the probability to reject a null hypothesis when it is false. In other words, the larger the power, the more likely we are to reject the null hypothesis when it is false. In tests of differences of two average values, statistical power is calculated from the sizes, averages and variances of the respective samples, as well as from the significance level at which the null hypothesis should be rejected. The data from our *post hoc* power analysis are presented in Table III.

Tables II and III show that there is not a sufficient number of observations of the group of foreign capital hardware manufacturers to permit any statistical test. Moreover, although there are enough observations and the corresponding statistics have sufficient power, our tests on foreign capital software companies and locally owned hardware manufacturers do not allow us to reject the null hypothesis, nor to confirm **HYP1**. This can be verified comparing the computed p-values with the significance level adopted in our work (0.025). It does not mean that **HYP1** does not hold for the corresponding populations, just that the statistical test could not determine it.

As can be noticed, the only tests for which **HYP1** is confirmed are those related to locally owned software companies, both for product and service suppliers. Since the only kind of software company included in the focus group of our statistical tests are ERP suppliers, our tests confirm that there is statistical support to ascertain that, on average, labor productivity growth is higher in companies affected by electronic tax recognition and bookkeeping regulation measures. The identification of ERP companies among service and product suppliers should not surprise reader, since these companies indeed provide packaged and customized software, as well as related services.

Let us turn to payroll tax reductions measures. In this case, almost all IT companies were benefited by the incentive in the short period of time comprising 2013 and 2014. Therefore, we are led to perform tests comparing the productivity growth rates of companies from 2004 to 2012 with those observed in the period from 2013 to 2014. Consequently, we test the following hypothesis in each of the four partitions of our data set, that is, considering the main business and capital origin of each company, generating six statistical test results:

(HYP2) On average, the labor productivity growth of companies is higher while their are benefited by payroll tax reduction measures than in other periods;

We also use Welch's t-test on our log-transformed productivity data set to attempt to validate **HYP2**. As a null hypothesis, we state that the average labor productivity growth of companies while they are in the scope of payroll tax reductions is equal to that outside the period of the granted benefits. The resulting sample sizes, t-values, required degrees of freedom and critical p-values are presented in Table IV.

TABLE IV
RESULTS OF WELCH'S T-TEST ON THE AVERAGE LABOR PRODUCTIVITY GROWTH OF BRAZILIAN IT COMPANIES.

NATURE OF BUSINESS	**ORIGIN OF CAPITAL**	**Benefited by Tax Incentives**			
		$n_1 + n_2$	t-value	DoF	p-value
A. Hardware Products	1. Local	22+47	−38.9728	26	1.0000
	2. Foreign	5+31	88.5780	4	0.0000
B. Software Products	1. Local	24+74	−1.4971	23	0.9260
	2. Foreign	1+19	N.A.	N.A.	N.A.
C. Software Services	1. Local	68+152	−1.1700	67	0.8769
	2. Foreign	17+53	−60.0674	16	1.0000

TABLE V
POWER OF WELCH'S T-TEST ON THE AVERAGE LABOR PRODUCTIVITY GROWTH OF BRAZILIAN IT COMPANIES.

NATURE OF BUSINESS	**ORIGIN OF CAPITAL**	**Benefited by Tax Incentives**		
		X_1 / X_2	S_1 / S_2	Power
A. Hardware Products	1. Local	0.9980 / 1.0134	0.0017 / 0.0009	100.0%
	2. Foreign	1.0346 / 1.0019	0.0008 / 0.0002	100.0%
B. Software Products	1. Local	1.0122 / 1.0139	0.0055 / 0.0005	15.9%
	2. Foreign	0.9693 / 1.0093	N.A.	N.A.
C. Software Services	1. Local	1.0124 / 1.0144	0.0142 / 0.0006	9.2%
	2. Foreign	0.9843 / 1.0078	0.0016 / 0.0003	100.0%

Before drawing conclusions from the data in Table IV, we investigate again the statistical power of our tests. The data of our *post hoc* power analysis are presented in Table V.

Tables IV and V show that there are not sufficient productivity growth observations from the investigated group of foreign capital software product companies to permit any statistical test. Moreover, the corresponding tests concerning locally owned software companies do not have sufficient statistical power. In addition, our tests on locally owned hardware manufacturers and foreign capital software service companies do not allow us to reject the null hypothesis, nor to confirm **HYP2**.

The only test which confirms **HYP2** is that related to foreign owned hardware manufacturers. Consequently, there is statistical support to ascertain that, on average, the labor productivity growth of foreign owned hardware manufacturers was higher while affected by payroll tax reductions than in other periods. What is most interesting to observe is that this group of companies was not among those originally targeted by this public policy measure, as described in Section III-A.

It is important to mention that we used in our analyses of labor productivity compound annual growth rates (CAGR) rather than absolute growth rates. We were led to adopt the CAGR measure since it reflects average annual rates instead of rates observed for the whole period of study, which cannot be computed in all cases due to missing data in the temporal boundaries of some time series.

VI. THREATS TO VALIDITY

The main threats to the validity of our work are the internal ones, related to data collection, data adjustment and derived data computation. Indeed, we adopted multiple and sometimes divergent secondary data sources, had to perform manual corrections in some collected data and made critical decisions concerning how to treat missing data. In the end, the obtained data sets were not randomly selected, nor normally distributed.

Concerning the adoption of multiple data sources, we had access to and used both primary and secondary sources, something that could compromise the quality of our data. However, the adopted sources were considered trustful in all cases, since the data provided was always subject to third party validation. The adopted data sources were published financial statements and published market research studies, which are respectively analyzed prior to their publication by auditors or board members and by market researchers. As a consequence, raw data provided directly by the studied companies was never used in our analyses.

The simultaneous adoption of primary and secondary data sources forced us to define clear criteria for data selection and was also a source of divergent observations concerning some companies. Data collected from primary sources was always preferred in our research, since no second party noise could be introduced in this way. For this reason, our data sets were populated first with data from published financial statements. On the other hand, economic data collected from secondary sources often contained mistakes and generated doubts in cases of multiple observations concerning the same company. In those cases, clearly mistaken data were manually corrected by the author, after validation with any existing alternative sources – such as Inland Revenue databases – and just the smallest figures from multiple observations were admitted, considering that a frequent source of lack of correction in market research studies is overstatement.

Missing data were also noticed in our data sets, which could lead to biased estimates, decreased statistical power, increased standard errors and weak generalizability. Concerning their treatment, we decided to use interpolation to estimate each missing revenue and employment data based on its neighborhood. This adjustment resulted in estimates for 13.0% of our data on revenues and 23.7% of our employment data. Alternative methods could be used, such as imputation [18], but interpolation appeared to have a better performance in our case, since data was not missing completely at random, existing data points were good predictors of missing data and the sizes of our data samples were expressive. Interpolation has also been used by other authors in similar situations [8].

Although our productivity data set captures the figures of companies of different natures and businesses, data samples defined out of it were not randomly selected, since the collected data reflects mostly the productivity of mid-size and large companies, for which public data is available. To confirm this, we performed a Wald-Wolfowitz runs test on a yearly basis, taking into account our revenue and employment data. In this kind of test, runs are defined as sequences of equal signals of the difference between an observed value in relation to the sample mean. As a null hypothesis, the number of runs is assumed to be a randomly distributed variable and the main hypothesis is that randomness is absent. By applying the test on our data sets, we confirmed, with the significance level of 0.025, that they were almost always not randomly selected.

Even though this could harm the possibility of generalizing our findings obtained from statistical tests, which are based on the randomness assumption, we obtained an expressive population coverage, leading to observations on productivity data of more than 10% of the estimated population. In addition, we performed alternative analyzes based on descriptive statistics and statistic control charts (not presented here), which led to the same conclusions.

Another aspect that deserves further analysis is the lack of normality in the distribution of our productivity data. In order to cope with this limitation, we applied logarithmic transformations attempting to approximate normality. Although the application of these transformations on our data was effective – since variance, skew and curtosis were significantly reduced – the transformed data was still not normally distributed, as we could confirm by the application of the Anderson-Darling test in each case. Since the sizes of our samples are expressive and we use statistic methods that compensate for distinct variances in sample sub-partitions, we accept the obtained results as if our log-transformed data were normally distributed.

Finally, geographic and temporal validity were not objects of concern here, since the analyzed public policy instruments are particular to the Brazilian IT Industry and considered almost the entire periods in which they have been in force.

VII. Concluding Remarks

We have investigated the influence of market regulation and tax benefit requirements on the labor productivity growth of Brazilian IT companies. We showed that there has been a positive correlation of electronic tax recognition and bookkeeping regulation measures with the productivity growth of locally owned suppliers of ERP software. Likewise, we showed that, on average, the labor productivity growth of foreign capital hardware manufacturers, while payroll tax reduction benefits were in force, was higher than in other periods.

By applying empirical methods to study these relationships, we were able to uncover the technology requirements, formal definition, business segment coverage and temporal span of the respective public policy measures, as well as to obtain evidence of the categories of companies that pass in the respective statistical tests. These illustrate the temporal and business segment influence that technology requirements might have on the productivity of IT companies when considered in public policy measures. It remains to be shown whether or not there is a causal connection between the respective requirements, measures and the identified better labor productivity growth performance.

Our studies also lead us to wonder which technology requirements should be posed so as to attempt to maximize public policy measure effects. The ideas that technology requirements can be implemented by legal means, so as to enforce their fulfillment, and that regulations can be formulated in goal oriented ways, in conversations between regulators and regulatees, respectively appear in [11] and [17]. As far as the productivity growth of IT companies in Brazil is concerned, we have provided statistical evidence that the inclusion of the requirements in Table I in the respective legislation (corresponding to a substantial part of these norms) is correlated to the productivity growth of local ERP companies, but the affordability requirement underlying the legislation described in Section III-A had limited correlation with the productivity growth of hardware manufacturers, just for foreign capital ones and during a short period of time. While it is not clear if it is the whole set of requirements or just a few of them that positively benefit the former business segment, the latter segment generates relatively less employment and perhaps this was an *ad hoc* additional reason for local policy makers to reduce the respective tax incentives recently. The reader should not perceive this public policy making approach as a means to create market barriers or provide unfair advantages to specific companies, economic sectors or to a whole national economy, so long as regulations and incentives equally benefit local and foreign capital companies and are connected to research and development efforts or anti-cyclic measures, which aim to equalize competitive conditions and are allowed by international trade agreements (such as those supported by the WCO), something that we have argued herein concerning the studied norms.

We perceive the reported studies as part of a broader research agenda which aims to investigate the connections among technology requirements, public policy instruments and corporate productivity. We have already investigated in [5] the connections between quality assurance measures and the labor productivity of Brazilian software companies. A specific research direction in our agenda is the study of the connections between norms that govern the Brazilian financial system and the productivity of local banking automation companies. We believe that some of the lessons learned by putting forward this agenda will be of local and general interest alike.

References

[1] ABES. *The Brazilian Software Market [Online]*. Brazilian Association of Software Companies, 2005-2015. Volumes from 2004 to 2014. Available: `https://www.abessoftware.com.br/dados-do-setor/ anos-anteriores`.

[2] Subbash C. Bathnagar and N. Bjorn-Andersen, editors. *Information Technology in Developing Countries*. Elsevier, 1990.

[3] David Burress. The Wisconsin tax model system. In Richard H. Austing, Dennis M. Conti, and Gerald L. Engel, editors, *Proceedings 1978 ACM Annual Conference*, volume II, pages 531–538. ACM, December 1978.

[4] Erran Carmel. The new software exporting nations: Success factors. *The Electronic Journal on Information Systems in Developing Countries*, 13(4):1–12, May 2003.

[5] Carlos Henrique C. Duarte. On the relationship between quality assurance and productivity in software companies. In *Proc. 2nd International Workshop on Conducting Empirical Studies in Industry (CESI 2014). Hyderabad, India*, pages 31–38, June 2014.

[6] Exame (Informática). The bigger and better Brazilian companies, 1999-2015. Volumes from 1998 to 2014. In Portuguese.

[7] Johan Gott. *A.T. Kearney Global Services Location Index: Making Offshore Decisions*. A. T. Kearney, 2003.

[8] Zvi Griliches. Productivity, R&D and basic research at the firm level in the 1970s. *The American Economic Review*, 76(1):141–154, March 1986.

[9] Informática Hoje. Anuário [Online], 2005-2015. Volumes from 2004 to 2014. In Portuguese. Available: `http://www.forumeditorial.com.br/`.

[10] Zvi Kedem et al., editors. *ACM Computing Classification System*. ACM, 2012.

[11] Alzbeta Krausova, Fabio Massacci, and Ayda Saidane. Legal patterns implement trust in IT requirements: When legal means are the "best" implementation of IT technical goals. In *Proc. 2nd International Workshop on Requirements Engineering and Law (RELAW 2009)*, pages 33–38, September 2009.

[12] David M. Levene, Mark Berenson, David Stephan, and Timothy C. Krehbiel. *Statistics for Managers using Microsoft Excell*. Prentice Hall, 5th edition, 2008.

[13] MCTI. *Report on the Results of the Information Technology Law [Online]*. Ministry of Science, Technology and Innovation, 2006-2015. Volumes from 2005 to 2014. Available: `https://sigplani.mcti.gov.br/`.

[14] OECD. *Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth*. Organization for Economic Cooperation and Development, 2001.

[15] Larry Press. Software export from developing nations. *IEEE Computer*, 26(12):62–67, December 1993.

[16] Rouzbahan Rashidi-Tabrizi, Guss Mussbacher, and Daniel Amyot. Transforming regulations into performance models in the context of reasoning for outcome-based compliance. In *Proc. 6th International Workshop on Requirements Engineering and Law (RELAW 2013)*, pages 34–43, July 2013.

[17] Andre Rifaut. Compliance management with measurement frameworks. In *Proc. 4th International Workshop on Requirements Engineering and Law (RELAW 2011)*, pages 15–24, August 2011.

[18] Bernard Rosner. *Fundamentals of Bioinformatics*. Books and Coole, 7th edition, 2010.

[19] Steve Sawyer. Packaged software: Implications of the differences from custom approaches to software development. *European Journal of Information Systems*, 9:47–58, September 2000.

[20] Harvey P. Siy et al. Making the software factory work: Lessons from a decade of experience. In *Proc. 7th International Symposium on Software Metrics (METRICS 2001)*, pages 317–326. IEEE Computer Society, April 2001.

[21] Ghanem Soltana, Mehrdad Sabetzadeh, and Lionel C. Briand. Model-based simulation of legal requirements: Experience from tax policy simulation. In *Proc. 24th International Requirements Engineering Conference (RE 2016) (to appear)*. IEEE Computer Society, September 2016.

[22] Valor Econômico. Valor 1000, 2004-2015. Volumes from 2003 to 2014. In Portuguese.