v. 3, n. 22, 2023

Journal of
Engineering
Research

GOODS AND SERVICES TAX DYNAMICS: A MULTIPLE LINEAR REGRESSION STUDY FOR BRAZILIAN STATES

Patrícia Ferreira Motta Café
Finance Secretariat (DF)

Luciano Emídio Neves da Fonseca
University of Brasília (UnB-FGA)

Ricardo Chaim
University of Brasília (UnB-FGA)

Abstract: This research paper aims to model the dynamics and characteristics of the tax on the circulation of goods and services (ICMS) and estimate the potential ICMS tax capacity for Brazilian states based on a set of socioeconomic variables. As ICMS serves as the primary source of public revenue for Brazilian states, it is crucial to determine the maximum tax collection capacity for the tax authority, considering the economic and social characteristics of each state.

The behavior of ICMS was analyzed by considering a combination of classic variables commonly used in similar studies. Moreover, an econometric model utilizing multiple linear regression, employing the Ordinary Least Squares Method, was developed to estimate the potential ICMS revenue for all Brazilian states. Statistical criteria were employed to select the most appropriate estimation mode while also adhering to the principle of parsimony by choosing the simplest model that meets the selected criteria. The Tax Effort Index for each state was calculated from the ratio between the effective and the potential ICMS revenue, providing a valuable tool for analyzing revenue performance in policy-making processes. Lastly, the study produced an SD model depicting the dynamics of the Brazilian good and services tax.

Keywords: ICMS, Econometric model, Multiple linear regression, Tax effort index, SD model.

INTRODUCTION

Taxes are the primary source of revenue to fund government operations and provide public services to society. In Brazil, the main taxes are levied on a classic basis such as income tax, consumption tax, and property tax. Due to Brazil's federalist system, taxes are collected at the three levels of government: federal, state, and municipal.

Generally, the Federal Government handles social contributions and income tax while the States are responsible for taxing goods, and the Municipalities impose taxes on services and urban properties. In 2021, the federal government was responsible for 66% of total revenue, while states and municipalities for 27% and 7% respectively. The Federal District is the only federated unit that collects both state and local taxes.

Within the Brazilian tax system, taxes on goods and services contributes to 44% of the total tax revenue. Among these taxes, the ICMS state tax is the one that most raises revenues, corresponding to 22% of the total and 96% of state revenues. The focus of this work will be on the ICMS tax.

According to Prado (2009), despite the significant role of ICMS in the Brazilian taxation system, the States face a fragile situation. Most states operate with limited budgets and are burdened with debt. Moreover, horizontal cooperation among states is lacking, and the "fiscal war" exemplifies this issue. After the Constitution of 1988, states suffered the most in terms of federal transfers due to the expansion of social contributions for municipal programs, which were not shared with the states. Over time, state governments have primarily relied on blue-chip sectors like electricity, telecommunications, and fuel for approximately 40% of their ICMS revenues, leaving little room for growth.

Consequently, state tax administrations have no choice but to collect ICMS as efficiently as possible, as this tax is the mainstay of their budgets. Monthly, state technicians evaluate tax collection by gathering modality and economic activity and establishing monthly and annual comparisons. However, the assessment of the state capacity of collecting ICMS should go beyond the simple analysis of historical collection series, because these series obviously do not include uncollected components due to the effect of tax expenditures, administrative and judicial litigations, elision and/or evasion. These uncollected taxes constitute the so-called tax gap, which is an object of many tax administration studies. Consequently, estimating the maximum capacity of tax revenue and knowing how much the effective collection represents in relation to this potential is an important management tool for tax administration. This ratio between actual and potential ICMS revenue will be here called Tax Effort Index (TEI).

The focus of this study is to explore a set of socioeconomic variables related to the 26 Brazilian States and the Federal District, collectively referred to as "States" hereafter, which can potentially explain their ICMS revenue from a structural standpoint. The ICMS variable will be regressed against explanatory variables of social and economic nature for all states, using an econometric tool with cross-section data for the year 2014. Consequently, models estimating potential revenue will be developed, and the fiscal effort index of each state will be calculated by comparing potential revenue with actual ICMS revenue.

The paper's organization is as follows: Section 2 will present a discussion, including a literature review of other studies that deal with the same theme. Section 3 will discuss the methodology to be used in the SD and statistical models. Section 4 will analyze the collected variables that can be used as explanatory variables in the estimation models of potential ICMS tax capacity as well as the behavior analysis of explanatory variables in relation to ICMS. Section 5 will test tax capacity models using the econometric package Gretl (Cottrell and Lucchetti, 2016), and the most appropriated model will be selected. Section 6 will discuss results obtained for the potential ICMS revenue using the selected model, a SD model of the state tax capacity will be presented as well as results for the fiscal effort of each State; and finally, Section 7 will present final conclusions.

DISCUSSION

The potential revenue of a specific country or state represents the maximum amount of revenue that the government could generate considering their socioeconomic conditions and the legal framework of their tax system. According to Viol (2006), there are two concepts of potential tax capacity: the legal and the structural perspective.

The legal potential tax capacity is related to what the government demands from taxpayers based on the existing tax legislation. The potential revenue would then be that maximum possible revenue resulting from the complete application of the current tax system. To measure the legal potential, one must consider the taxable basis outlined in the legislation and the applicable current tax rates. Carvalho et al. (2008) introduced a deterministic method for measuring tax capacity.

On the other hand, the structural potential tax capacity is more challenging to define and measure accurately. It involves estimating econometric models where taxes serve as the dependent variable, while other explanatory variables reflect the socioeconomic characteristics of the country or state. Estimating the structural potential requires considering various literature works that have employed this approach, both internationally and in Brazil.

From the concept of potential tax capacity, one can derive the concept of Tax Effort Index (TEI), or degree of effectiveness, according to some authors. This index is calculated from the ratio of tax revenue, which effectively enters in the public coffers, and the potential revenue, which is estimated by an appropriate structural econometric model. The TEI is used to make comparisons of fiscal efforts between countries, as well as among federal units of a given country.

At the international level, several authors have studied variables and tax capacity models of countries. Using cross-section data in 1964, Lotz and Morss (1969) were the first authors to confirm the positive influence of per capita income and degree of openness of economy. Shin (1969) discussed the significance of per capita income, agricultural product and population growth variables in the analysis of cross-section data. Chelliah (1971) showed that ratio of extractive industry product variable was highly significant, degree of openness was significant and per capita income was not significant. Bahl (1971) confirmed the significance of the agricultural product and the mining industry product, and the tax capacity related negatively with the first and positively with the second. Tait, Grätz and Eichengreen (1979) updated the results of Lotz and Morss, as well as Chelliah using cross section data in 1974, and they concluded that the variables of the most explanatory power were mining industry product and degree of openness. Mann (1980) studied the tax capacity of Mexico, using time series, and he concluded that degree of openness, per capita income and agricultural products were significant at certain periods of time while only per capita income was significant and inversely related to tax capacity at more recent time. Piancastelli (2001) used both cross section data with the average for the period 1985-95, as well as panel data, concluding that for the total sample studied, per capita income and degree of openness of economy were significant. However, when the sample was divided into low- and middle-income countries, he found that only degree of openness became significant for the low income group while agricultural and industry products influenced tax capacity negatively and positively, respectively, for the middle-income group. Cafe (2003) estimated tax capacity of industrialized countries and Latin America countries, concluding that per capita income and degree of openness were significant and positively related to tax capacity of the full sample of countries, while for separated groups, there was an improvement in the linear adjustment when the agricultural product variable was added to the model.

Several studies establish comparisons among the tax capacity of the Brazilian States. Reis and Bianco (1996) used production function models with panel data for the years 1970, 1975, 1980, 1985 and 1990, obtaining expected results for GDP, urban population, and inflation. Marinho and Moreira (2000) estimated the potential tax capacity of the Northeast States for various taxes in the period between 1991 and 1996, also using models of production function with panel data, obtaining significant and direct relationships between ICMS and per capita income, urban population and degree of urbanization, and negative relations with exports and inflation. Vasconcelos *et al* (2006) used panel data from 1986 to 1999 to estimate the potential tax burden of Brazilian States, concluding that industry and services product and GDP per capita were significant and they had positive signs as expected. Carvalho *et al* (2008) estimated the Amazon States tax capacity between 1970 and 2000, in the census years, also using production function models with panel data and they concluded that the economic and demographic variables used in the model were important to access potential revenues of States. However, they obtained a negative not expected sign for industrial product and a not significant relationship. Cafe (2011) estimated the potential tax capacity of Brazilian States in the 2003-2007 period, using linear regression models, and she concluded that GDP, urban population, and industry variables were positive and significant.

SYSTEM DYNAMICS MODELING COMBINED WITH STATISTICAL METHODS

The study considered many variables found in the literature that could affect the tax capacity like the Gross Domestic Product (GDP), which measured economic development stage; Exports and Imports, which measured degree of openness; Sector Products (added-value indices) that measure the degree of industrialization and urbanization; and population size. Many other variables such as level of economic inequality (Gini Index), debt (Default Rate), employment (Formal Jobs) and size of the private sector, may also influence state tax capacity. However, a full list would produce extremely complex and probably inaccurate models.

The analysis of variables considered interventions and expected signs (as shown in the Casual Loop Diagram in Figure 1), as well as level of correlation among them, always considering specific economic and tax aspects of Brazilian states. An explanation of the methods used, besides the analysis of each variable and concerns about their interrelation are shown below.

Multiple linear regression models and the method of Ordinary Least Squares (OLS) were used to estimate the potential ICMS tax capacity of Brazilian States. For that, the potential revenue of each State was estimated by the following equation, according to Wooldridge (2010):

$$\hat{y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_K x_{Ki} + \mu_i$$

where,

$\hat{y}_i$ - Estimated value for ICMS revenue of each State;

i - Index which represents each State (from 1 to 27);

K - Index which represents the number of explanatory variables;

x_{Ki} – Explanatory variable K of State i;

β_K – Parameter to be estimated for each explanatory variable K;

β_0 – Intercept;

μ_i - Residue of the State index i.

The value $\hat{y}_i$ calculated by this equation, estimates the actual value y_i of potential ICMS revenue for each State. The difference between the estimated and the actual value is represented by a residue μ_i. Therefore:

$$\mu_i = y_i - \hat{y}_i$$

where,

$\hat{y}_i$ - Estimated value of potential ICMS revenue of each State;

y_i - Actual value of ICMS revenue of each State;

μ_i - Residue of the State index i.

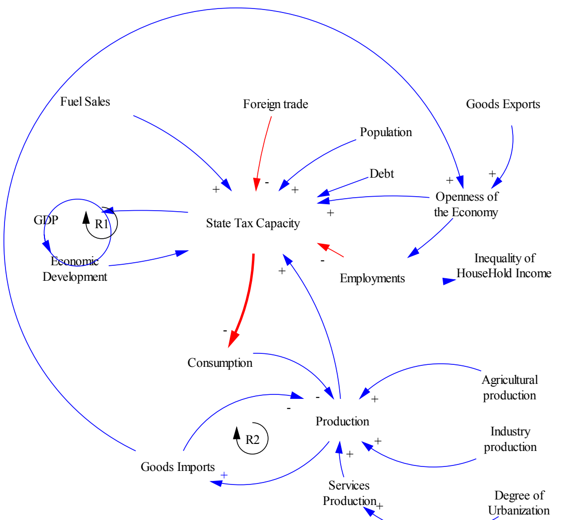

Figure 1: ICMS Causal Loop Diagram

A vector of β parameter (one parameter β_K for each explanatory variable) must be chosen to make the smallest possible error in the estimation of the ICMS potential revenue y_i. The estimations of is accomplished by solving a set of $\hat{\beta}$ overdetermined normal equations, which have the following solution:

$$\hat{\beta} = (X^TY)^{-1}X^TY$$

where,

$X_{27 \times (K+1)}$ is the design matrix with all measured values of the explanatory variables. The lines correspond to the index of each State, and the columns correspond to the index of each explanatory variable.

$Y_{27 \times 1}$ is the column vector with measured values of ICMS revenue of each State.

$$Y_{27x1} = \begin{bmatrix} y1 \\ y2 \\ ... \\ y27 \end{bmatrix}$$

After the estimation of β parameters, the following tests were applied:

1) **Reset Test** of Ramsey regression specification error, as in Wooldridge (2010), especially for omission of variables. This test includes quadratic and cubic terms in the model, and it verifies if the coefficients of these terms are significant, via F test:

Ho: $\beta_5 = \beta_6 = 0$

H1: $\beta_5 \neq 0$ or $\beta_6 \neq 0$

If Ho is rejected, the model is poorly specified.

2) **Breusch-Pagan Test** of heteroscedasticity, as in Wooldridge (2010). This test checks if the variance is affected by some independent variables (μ^2 x β) via F test:

Ho : $\beta_1 = \beta_2 = \beta_3 = ... = 0$

Ho is the null hypothesis of homoscedasticity.

3) **Variance Inflation Factors (VIF)** where values above 10.0 may indicate a multicollinearity problem, as pointed by Miloca S. and Conjo P. (2011). If a variable is a linear combination of others, the fit degree R^2 tends to 1.0, since VIF is given by:

$FIV = 1/(1-R^2)$

if $R^2 \rightarrow 1$, $FIV \rightarrow \infty$

In this study, Gretl and Vensim software packages were used for analysis and selection of the most appropriate model in statistical terms, as well as for running the tests presented above.

ANALYSIS OF SOCIOECONOMIC VARIABLES

Several socioeconomic variables were collected for all Brazilian States, which may serve as explanatory variables to estimate the potential ICMS tax capacity for the year 2014. Table 1 consolidates this information.

Figure 2 presents a preliminary graphic analysis of ICMS, the dependent variable of the model. A compensation factor of 1/3 was applied to ICMS data from São Paulo State, to avoid distortions. São Paulo is clearly an outlier, with production index and tax revenues way beyond the Brazilian average. The same compensation factor will be applied to the variable population and to all those variables related to the economic performance of São Paulo (GDP, Ag_P, Ind_P, Serv_P, Fuel, X and M).

From Figure 2, the ICMS revenue curve of Brazilian States follows an exponential trend with degree of adjustment R^2 of 0.96. Then the variable logarithm will be used for the linearization of the model.

The correlation level between the dependent variable (ICMS) and each explanatory variable was studied. Table 2 shows the values of linear correlation (LC), which serves as an indication of what variables should be included in potential tax capacity models.

Variables	LC
Serv_P	0,98
GDP	0,97
Fuel	0,97
Ind_P	0,92
EAP	0,92
M	0,89
X	0,86
Income	0,60
Agr_P	0,58
FJ	0,50
DR	-0,41
Gini	-0,16

Table 2 – Linear correlation between ICMS and explanatory variables represented by their initials according to Table 1.

The linear correlation values vary in magnitude between 0,16 and 0.98. The positive sign points to a direct linear correlation while the negative sign for an inverse linear correlation. All variables will be regarded as potential explanatory variables in the ICMS tax capacity models.

Table 2 shows that the relationship between ICMS and variables which indicate economic performance were positive and high, above 0.85, except for Agriculture Product (Ag_P). For illustration purpose, Figure 3 shows the relationship between ICMS and industry product (Ind_P) with positive linear correlation of 0,92. This positive relationship was expected, as pointed out by Varsano et al (1998).

Rio de Janeiro State is an outlier in the relationship between ICMS and Industry Product, due to the oil production industry. Although the State is a major oil producer, taxation is at destination in interstate operations with lubricants and oil fuels, according to the Brazilian Federal Constitution. Thus, ICMS is charged where consumption occurs, which induces a gap between industry product and ICMS revenue in Rio de Janeiro.

The curves of variables related to economic activity plotted against Brazilian States ordered by ICMS revenue follow an exponential trend, which suggests the use of logarithm of these variables to obtain linear relationships in the estimation models of potential ICMS tax capacity. All these variables, with the exception of Agriculture Product, have good explanatory power of the dependent variable ICMS, since the values of fit degree R^2 are above 0.75.

The level of linear correlation among the variables that are indicative of economic activity was also studied and a high correlation was found, as expected. Thus, the models will contain only one indicative variable of economic activity to avoid collinearity between explanatory variables, respecting this classic hypothesis of the method.

The Economically Active Population (EAP) also has a strong correlation to ICMS, with correlation index of 0.92. The positive relationship is obviously expected: the larger the population, the greater the tax capacity, as cited by Varsano et al (1998). The EAP curve has also an exponential behavior, shown in Figure 4, also indicating the use of its logarithm for the linearization of the estimation model.

The degree of adjustment R^2 of 0.84 of ICMS versus EAP indicates a good explanation power of EAP variable in the ICMS potential estimation model, as shown in Figure 5.

Gini index measures the degree of inequality in the distribution of per capita household income among individuals. Its value can vary theoretically from zero, when there is no inequality, to one when inequality is maximum. The correlation coefficient found between ICMS and Gini index is -0.16, indicating an inverse relationship between these two variables.

There is a positive linear correlation of 0.86 and 0.89 between ICMS and export (X) and import (M) variables, respectively, as shown in Table 1, despite the Complementary Law No. 87/1996, known as Kandir Law. Although this Law exempts from taxation goods and services for export, increase in ICMS revenue following both exports and imports was observed. This positive correlation is probably due to the direct effects of imports and secondary effects of exports, moving the economy, creating jobs and increasing income to purchase goods.

Varsano et al (1989) argue in favor of an inverse relationship between potential ICMS revenue and trade balance, and consequently exports. On the other hand, Vasconcelos et

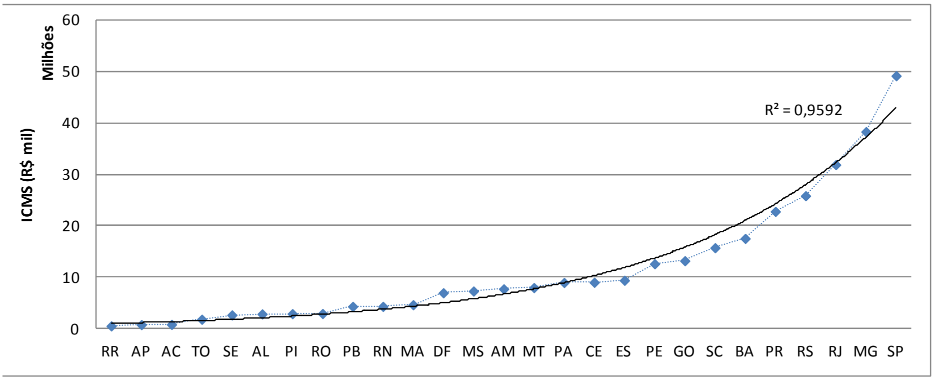

Figure 2 – ICMS revenue 2014 in R$ billions versus Brazilian States ordered by ICMS revenue, considering one third of São Paulo ICMS revenue.

Variables	Initials	Units	Source
Goods and Services Tax	GST	R$ (thousands)	CONFAZ
Gross Domestic Product	GDP	R$ (millions)	IPEA/IBGE
Economically Active Population	EAP	Individuals	IBGE
Agricultural Product	Ag_P	R$ (millions)	IBGE
Industry Product	Ind_P	R$ (millions)	IBGE
Sevice Product	Serv_P	R$ (millions)	IBGE
Gini Index	Gini	Index	IBGE
Fuel Sales	Fuel	m^3	ANP
Exports	X	US$ (thousands)	BCB/MDIC/Secex
Imports	M	US$ (thousands)	BCB/MDIC/Secex
Formal Jobs	FJ	Units	BCB/TEM
Default Rate	DR	Percentage	BCB
Average Monthly Income	Income	R$	IPEADATA/PNAD/IBGE

Table 1 – Socioeconomic variables to estimate potential ICMS tax capacity.

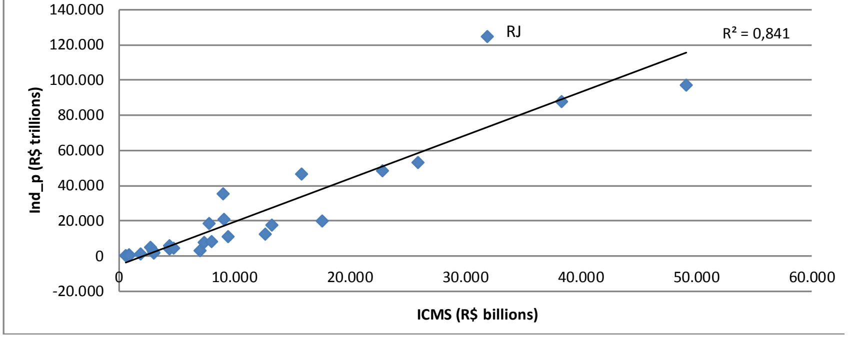

Figure 3 – Industry product (Ind_P) in R$ trillions versus ICMS in R$ billions of Brazilian States. The blue points are observations for each State. The continuous line represents a linear model adjusted to the observations with R^2 = 0,841. Notice that Rio de Janeiro State is an outlier.

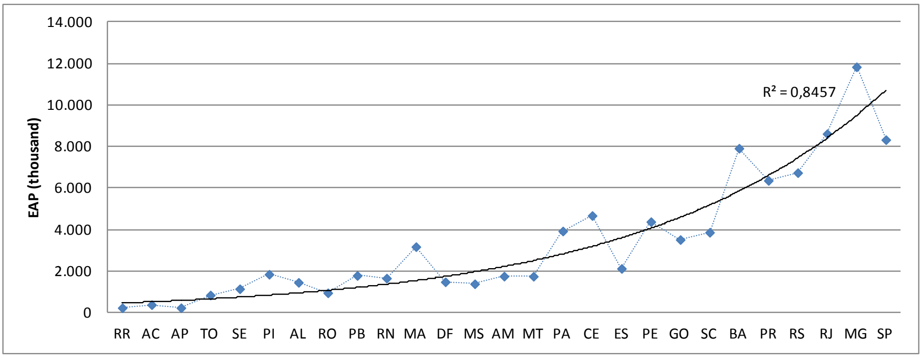

Figure 4 – Economically Active Population (EAP) in thousands of individuals versus Brazilian States ordered by ICMS revenue. The blue points represent the observations for each State. The continuous line represents an exponential model adjusted to the observations with $R^2 = 0,846$.

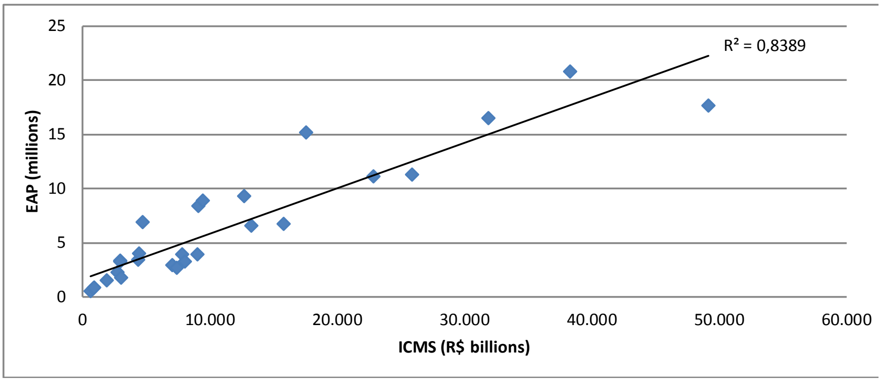

Figure 5 – Economically Active People (EAP) in millions of individuals versus ICMS in R$ billions of Brazilian States. The blue points are the observations for each State. The continuous line represents a linear model adjusted to the observations with $R^2 = 0,8389$.

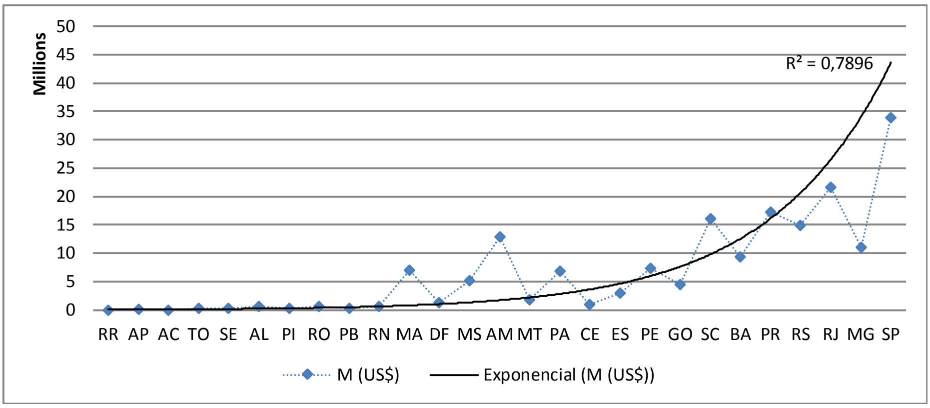

Figure 6 – Imports of goods (M) versus Brazilian States ordered by ICMS revenue. The blue points are the observations for each State. The continuous line represents a 2nd degree polynomial model adjusted to the observations with $R^2 = 0,7896$.

Journal of Engineering Research ISSN 2764-1317 DOI 10.22533/at.ed.3173222314064

*ln(ICMS) = 4,768 +0,762*ln(GDP) + 0,234*ln(EAP) - 2,7348*(Gini)*				
n = 27 observations (Brazilian states)				
	coefficient	**standard error**	**t satistics**	**p-value**
const	4,76842	0,730709	6,526	1,17E-06
l_GDP	0,761815	0,0692918	10,99	1,24E-10
l_EAP	0,233736	0,0773604	3,021	6,10-03
Gini	-2,73481	0,932574	-2,933	7,50E-03
R^2 adjusted	0,982145			
F(3,23)	477,7343	**p-value (F)**	7,52E-21	

Table 3 – Selected Model for Potential ICMS Tax Capacity Estimation

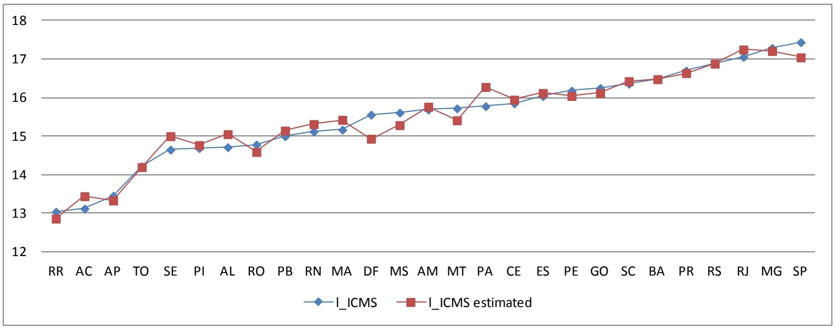

Figure 7 – Comparison between logarithm of effective ICMS (in blue) and logarithm of ICMS values estimated by the model (in red).

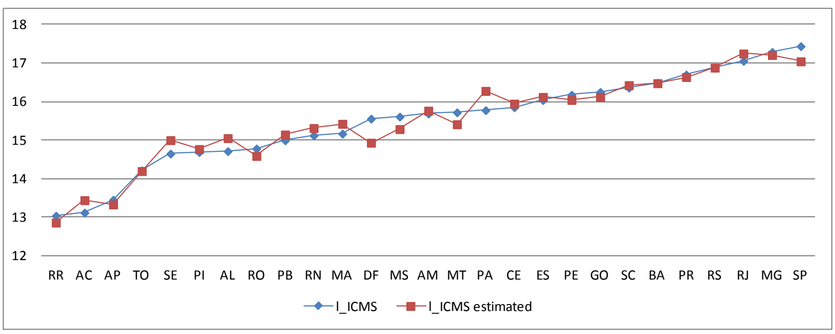

Figure 7 – Comparison between logarithm of effective ICMS (in blue) and logarithm of ICMS values estimated by the model (in red).

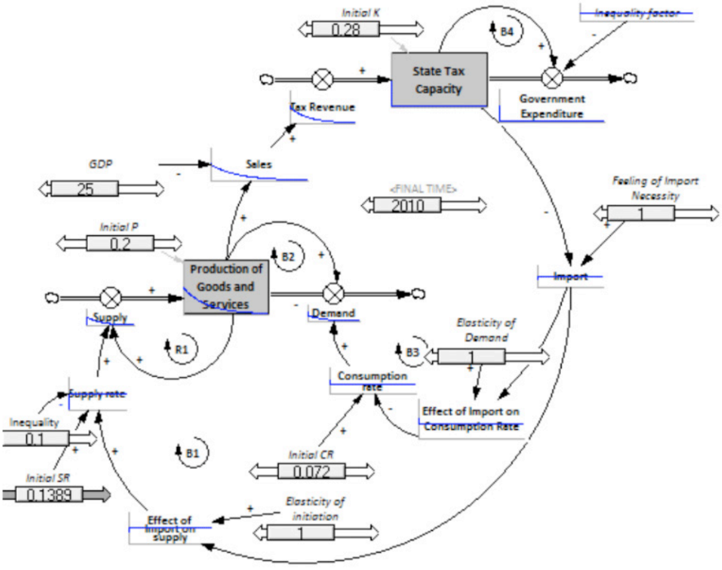

Figure 9 - Stock and Flow Diagram of the state tax capacity problem

al (2006) point out that international trade is an important source of income, especially in developing countries.

There is a good degree of adjustment between ICMS and foreign trade variables, with R^2 greater than 0.75. In addition, X and Y curves tend to a 2nd degree polynomial equation, suggesting the use of square root of these variables in order to linearize the models. Figure 6 illustrates the curve of Imports (M).

The variable of formal jobs generated in the year (FJ) is the difference between admissions and layoffs in 2014. This variable is an indirect indicator of economic performance and therefore establishes a positive correlation with ICMS of 0.50. However, the ICMS versus FJ showed a degree of adjustment R^2 of less than 0.5, indicating that FJ has a low explanation power in the ICMS estimation model.

According to Table 1, the variable default rate on credit operations (DR) has a negative correlation with ICMS of -0,41, indicating an inverse relation. The negative sign of this correlation was expected, since the default rate is an indicative, in reverse, of family income and its consumption power.

SELECTION OF ICMS TAX CAPACITY MODEL

Several models have been tried with Gretl program (Cottrell and Lucchetti, 2016), including the explanatory variables described in the previous section, and containing only one variable of economic performance. The goal in any attempt was to obtain a model that contains all significant variables to at least 10%, expected signs and satisfactory behavior in Reset, Breush-Pagan and collinearity tests.

The F statistic showed significance at 1% of the set of variables in all tested models, indicating that the set of explanatory variables can effectively be used to estimate the dependent variable (ICMS). However, only three models were able to meet the conditions placed above. The selected model incorporates three significant variables: Gross Domestic Product (GDP), population (EAP) and Gini coefficient (Gini). These variables demonstrated statistical significance at a level of at least 1% and exhibited the expected signs. The model showed good fit and good results in all tests. Since there is more the one model that presented good results in all statistic tests, this one was selected using the Akaike, Shwarz and Hannan-Quinn criteria. Table 3 presents the selected model.

RESULTS

The Tax Effort Index (TEI) was calculated from the estimated ICMS tax capacity. The largest deviations obtained between actual and estimated ICMS are reflected in this index since it is defined as the ratio between effective and potential ICMS. Figure 8 shows TEI curve for Brazilian States, whose index values range between 0.80 and 1.41. Index values below 1.0 indicate that States can increase its ICMS revenue, while values above 1.0 indicate that States collect higher revenue than what would be expected from their bases.

Figure 7 shows the comparison between logarithm values of effective ICMS and ICMS values estimated by the selected model. The estimation has an average deviation of 0,01% from the logarithm of effective ICMS, which corresponds to 0,99% from the effective ICMS.

Table 4 presents the effective and the potential ICMS tax revenue (year 2014) predicted by the econometric model which includes Gross Domestic Product (GDP), Economically Active Population (EAP) and Gini Index, as well, the Tax Effort Index for the 27 Brazilian States. Figure 8 shows the Tax Effort Index for the set of Brazilian States to help visualize their behavior.

UF	ICMS revenue (R$ thousand)		Tax Effort Index
	Effective	Potential	
AP	861.450	1.070.925	0,80
MA	4.715.626	5.780.039	0,82
DF	7.031.138	8.357.760	0,84
SC	15.769.833	18.567.402	0,85
RJ	31.886.537	37.425.693	0,85
RR	606.921	698.763	0,87
SP	122.836.140	138.736.045	0,89
AL	2.927.852	3.265.054	0,90
PR	22.815.805	25.187.339	0,91
PA	9.067.286	9.853.340	0,92
SE	2.732.000	2.902.787	0,94
PI	2.979.359	3.034.906	0,98
CE	9.455.729	9.539.847	0,99
AC	896.908	892.783	1,00
TO	1.894.720	1.857.080	1,02
GO	13.252.854	12.711.632	1,04
MG	38.288.044	36.112.911	1,06
RS	25.854.212	24.226.969	1,07
RN	4.389.457	4.109.319	1,07
ES	9.025.731	8.306.675	1,09
PB	4.391.930	3.996.106	1,10
MT	8.038.089	7.232.896	1,11
BA	17.575.936	15.794.274	1,11
PE	12.659.801	11.263.545	1,12
RO	3.006.652	2.653.030	1,13
MS	7.366.784	5.294.727	1,39
AM	7.788.737	5.537.286	1,41

Table 4 – ICMS effective e potential, Tax Effort Index

According to Table 4, 14 States achieved a TEI equal to or above 1.0, indicating a higher tax effort, while12 States had a TEI below 1.0, indicating a relatively lower tax effort. Amazonas and Mato Grosso do Sul presented the highest TEI of 1.41 and 1.39, respectively. On the other hand, Amapá, Maranhão and Distrito Federal presented the smallest TEIs of 0.80, 0.82 and 0.84, respectively.

On the other hand, Distrito Federal presented a low TEI de 0.84 due to the size of the public sector, which responds for about 40% of its economic activity and it is not in the incidence field of the ICMS taxation.

These econometric analyses give insights to produce the following stock and flow diagram to represent the variables and their interrelation. The diagram shown in Figure 9 considers the actual structure of the model, including stocks, flows, estimated parameters and external inputs [Sterman, 2000].

More supply explains the production of goods and services, which induces an increase in sales and in state tax capacity. The demand is also based on the total amount of goods and services available, on the size of population and on consumption rate. The last one is influenced by imports and by state tax capacity itself, as a consequence of the openness of the economy.

CONCLUSION

Measuring the structural potential tax capacity based on the socioeconomic characteristics of state or country is not a straightforward task. It requires characterizing the tax base by considering various factors, such as the capacity of tax contribution (GDP and population), composition of the economy (products form different sectors), foreign trade, degree of urbanization and other relevant variables. These factors need to be incorporated into models that aim to estimate state or country tax capacity.

This paper examined a comprehensive set of variables that might explain tax capacity, including GDP, sector products and fuel sales variables used to measure economic performance. Exports and imports of goods were used to measure the degree of openness of the economy. As far as socioeconomic parameters, besides size of population widely used, number of formal jobs, default rate on credit operations, Gini index of income inequality and finally, a proxy of private sector were adopted.

The preliminary study of the dependent variable (ICMS) and the behavior of potential explanatory variables were very useful for building models. This study consisted of a graphical analysis of variables for the set of States, verifying trends and presence of outliers. In the case of ICMS variable, for example, exponential trend was observed, as well as presence of an outlier that would be São Paulo revenue, which is far higher than the other state revenues. Thus, measures were taken to correct possible distortions that could happen, such as using ICMS logarithm aimed the variable linearization, and employing a third of São Paulo ICMS revenue. It was also found exponential behavior of GDP, population, sector products and fuel sales, indicating the use of these variables logarithm as well. Regarding trade variables, it was found quadratic trend, indicating the use of these variables square root.

Additionally, analysis of correlation level and signs between ICMS and explanatory variables was important to select which variables had power to explain it and had the expected signs. Thus, positive and over 0.8 correlations were found between ICMS and GDP, population, industry and service products, fuel sales, exports and imports. Negative and between 0.5 and 0.8 correlations were found between ICMS and Gini index, and positive between ICMS and formal jobs and agriculture product. Finally, low correlations below 0.2, negative and positive, were found between ICMS and default rate and private sector proxy, respectively.

On the other hand, high levels of correlation between independent variables possibly introduced collinearity issues in models. In this study, it was observed a high degree of correlation among the variables that were indicative of economic activity, leading the use of only one variable of this group in the estimation models.

Several models using the method of OLS in Gretl were tested. In general, it was obtained a good degree of fit in almost all models tested and they all showed significance of the set of variables using the F test. However, few models presented all variables significant at least 10%, in addition to be simultaneously successful in Reset test for good model specification, Breusch-Pagan test for heteroscedasticity and collinearity test.

Choosing the best model should be based on objective criteria, for example, containing all statistically significant variables, expected signs of coefficients and good degree of adjustment, besides meeting the requirement tests and the classic assumptions of the OLS method. It should be taken into account the principle of parsimony in all cases to choose the simplest model that still meets the requirements. According to that, the selected model was one that explains ICMS logarithm by industry product logarithm and population logarithm.

The results of the potential tax capacity analysis were used to calculate the Tax Effort Index (TEI) for Brazilian states. The highest TEI values were observed in Amazonas (1.41) and Mato Grosso do Sul (1.39). These states exhibited exogenous factors that positively influenced their index, as previously discussed in relation to their economic base. On the other hand, the lowest TEI values of 0.80, 0.82, and 0.84 were observed in the states of Amapá, Maranhão, and Distrito Federal, respectively. This can be attributed to the fact that certain significant economic activities in these states, particularly in the case of the Distrito Federal, are not subject to ICMS legislation, as previously mentioned.

The Tax Effort Index is a useful tool for analyzing fiscal performance, which allows comparisons between countries or states. It can even be considered for feasibility studies of tax burden raising or even as a guide to tax enforcement actions. However, TEI should not be used mechanically as an absolute truth. Its calculation is linked to econometric models' estimation, which always requires additional analysis and verification of results.

Finally, the present econometric study was based on cross section data for the set of Brazilian states. For further studies, it is desired to include time dimension by using panel data to establish a more adjusted comparison with the result from SD model.

REFERENCES

BANCO CENTRAL DO BRASIL. **Economia Regional.** Available in: <https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries >. Access in: 10 Nov. 2015.

BAHL, R. W. A Regression Approach to Tax Effort and Tax Ratio Analysis. **IMF Staff Papers**, Washington, v. 18, n. 3, p. 570-612, Nov. 1971.

CAFÉ, P. F. M. **Reforma Tributária: Panorama Internacional e Aspectos Quantitativos. 2003**. 111 f. Dissertation (Master's Degree in Economcs) – Economy Departament, Universidade de Brasília, 2003.

CAFÉ, P. F. M. Um Estudo do Potencial de Arrecadação do ICMS e do Índice de Esforço Fiscal dos Estados Brasileiros. **Caderno de Finanças Públicas**, Brasília, n. 11, p. 5-296, Dec. 2011.

CARVALHO, D. F.; OLIVEIRA, C. C. R. de; CARVALHO, A. C.. Desigualdades Inter-Regionais, Capacidade Tributária e Esforço Fiscal dos Estados da Amazônia (1970-2000): uma abordagem econométrica de fronteira estocástica. **Novos Cadernos NAEA**, [S.I.], v. 10, n. 2, p. 5-44, Dec. 2007.

CHELLIAH, R. J. Trends in Taxation in Developing Countries. **IMF Staff Papers**, Washington, vol. 18, n. 2, p. 254-331, Jul. 1971.

CONFAZ - CONSELHO NACIONAL DE POLÍTICA FAZENDÁRIA. **Boletim do ICMS**. Available in: <https://www.confaz.fazenda.gov.br/legislacao/boletim-do-icms>. Access in: 25 Jun. 2015.

COTTRELL, A and LUCCHETTI, R, "Gretl User's Guide: Gnu Regression, Econometrics and Time-series Library"; <http://gretl.sourceforge.net/gretl-help/gretl-guide.pdf>, Access in: 1 Mar., 2016

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Sistema de Contas Regionais.** Available in: < http://www.ibge.gov.br/home/estatistica/economia/contasregionais/2013/default.shtm>. Access in: 25 Jun. 2017.

_______. **Sistema IBGE de Recuperação Automática – SIDRA. População.** Disponível em: < http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?z=t&o=25&i=P&c=3153>. Access in: 8 Nov. 2017.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Base de dados Regional. Contas Regionais.** Available in: http://www.ipeadata.gov.br/. Access in: 20 Jan. 2017.

_______.**Base de dados Social. Renda Nível: Estados.** Disponível em: http://www.ipeadata.gov.br/. Access in: 25 Fev. 2016.

MANN, A. J. Economic Development and Tax Structure Change: Mexico, 1895-1975. **Public Finance Review**, [S.I.], v. 8, n. 3, p. 291-306, Jul. 1980.

MARINHO, E. L. L.; MOREIRA, A. F. Esforço Fiscal e Carga Tributária Potencial dos Estados do Nordeste. **Revista Econômica do Nordeste**, Fortaleza, v. 30, special n., p. 634-651, Dez. 1999.

MILOCA, S. A.; CONEJO P. D. Multicolineariedade em Modelos de Regressão. In: XXII SEMANA ACADÊMICA DA MATEMÁTICA, 2008, Cascavel. **Anais eletrônicos...** Cascavel: Unioeste, 2008. Available in: <https://www.ime.usp.br/~yambar/MI404Metodos%20Estatisticos/Aula%208-9%20Regress%E3o%20mult%20dim/inete%20adicional%20-%20multicolinearidade%20em%20modelos%20de%20regressao.pdf> Access in: 8 Fev. 2016

PIANCASTELLI, M. Measuring the Tax Effort of Developed and Developing Countries. Cross Country Panel Data Analysis, 1985/95. **Texto para Discussão IPEA**, Rio de Janeiro, n. 818, p. 1-18, Set. 2001. Available in: <http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/ingles/dp_103.pdf> Access in: 25 Sep. 2015.

PRADO, S. R. R. do. A "questão federativa" no Brasil: ampliando a agenda. In: PISCITELLI, R. B. et al. **Reforma tributária:** a costura de um grande acordo nacional. São Paulo, Editora Atlas, 2009. chap. 5, p. 170-181.

REIS, E. J.; BLANCO, F. A. Capacidade Tributária dos Estados Brasileiros, 1970/90. **Texto para Discussão IPEA**, Rio de Janeiro, n. 404, p. 1-26, Fev. 1996. Available in: <http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_0404.pdf>. Access in: 30 Jan. 2016.

SECRETARIA DE FAZENDA DE ESTADO DO DISTRITO FEDERAL. **Arrecadação de Origem Tributária do Distrito Federal.** Gerencia de Estudos Econômicos e Política Fiscal, Dec. 2014.

SECRETARIA DO TESOURO NACIONAL. **Estimativa da carga Tributária Bruta do Governo Geral 2021.** Disponível em: < https://sisweb.tesouro.gov.br/apex/f?p=2501:9::::9:P9_ID_PUBLICACAO:43205#:~:text=Em%202021%2C%20a%20carga%20tribut%C3%A1ria,PIB%20em%20rela%C3%A7%C3%A3o%20a%202020.> Acesso em: 23/03/2023.

SHIN, K. International Difference in Tax Ratio. **The Review of Economics and Statistics**, Massachusets, v. 51, n. 2, p. 213-220, May 1969.

Sterman, J. (2000). Business dynamics: systems thinking and modeling for a complex world, Irwin McGraw-Hill.

TAIT, A. A.; GRATZ W. L. M.; EICHENGREEN, B. J.. International Comparisons of Taxation for Selected Developing Countries, 1972-76. **IMF Staff Papers**, Washington, v. 26, n. 01, p. 123-156, Mar. 1979.

VASCONCELOS, J. R. de; PIANCASTELLI, M.;MIRANDA, R. B. Esforço Fiscal dos Estados Brasileiros. **Texto para Discussão IPEA**, Brasília, n. 1045, p. 1-39, Set. 2004. Available in: <http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_1045.pdf>. Access in: 30 Set. 2015.

VARSANO, R. et al. Uma Análise da Carga Tributária no Brasil. **Texto para Discussão IPEA**, Rio de Janeiro, n. 583, p. 1-55, Aug.. 1998. Disponível em: http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_0583.pdf. Access in: 07 Fev. 2016.
Ventana Systems. (2015). Manual of Vensim version 5. Retrieved from http://www.vensim.com/

VIOL, A. **A Definição e o Cômputo da Arrecadação Potencial. Análise do Potencial Econômico-Tributário e de seus Condicionantes**. In: Assembleia Geral do Centro Interamericano de Administrações Tributárias (CIAT), 40.: Florianópolis, Apr. 2006.

WOOLDRIDGE, J. M. **Introdução à Econometria:** Uma Abordagem Moderna. 4 ed. São Paulo: Cencage Learning, 2014. 701 p.