Economics
The Open-Access, Open-Assessment E-Journal
Vol. 13, 2019-31 | May 17, 2019 | http://dx.doi.org/10.5018/economics-ejournal.ja.2019-31

Regional tax effort in Spain

Anabel Zárate-Marco and Jaime Vallés-Giménez

Abstract
This work examines in depth the hypotheses explaining the tax capacity of regional governments, also determining their tax effort and explanatory factors. The study is done for the Spanish regions, using different techniques which have rarely been applied in this area. The results show that these jurisdictions have exercised their tax autonomy responsibly, in response to different budget and demographic factors and to the economic cycle. Also, an asymmetrical tax behaviour linked to income is observed: some regions have practically exhausted the possibilities of current sub-central taxes, while others still have ample fiscal space.

JEL H71 H2 C23 C51
Keywords Regional tax effort; regional taxes; tax potential; frontier techniques

Authors
Anabel Zárate-Marco, University of Zaragoza, Spain, azarate@unizar.es
Jaime Vallés-Giménez, University of Zaragoza, Spain

The authors would like to thank the three anonymous referees and the co-editor of the journal for their suggestions and comments, which have undoubtedly improved the draft. They gratefully acknowledge the funding support of the Spanish Institute for Fiscal Studies, the Government of Aragon and the European Regional Development Fund (Public Economics Research Group) and the Ministry of Science, Innovation and Universities (Project RTI2018-095799-B-I00). The usual disclaimer applies.

Citation Anabel Zárate-Marco and Jaime Vallés-Giménez (2019). Regional tax effort in Spain. *Economics: The Open-Access, Open-Assessment E-Journal*, 13 (2019-31): 1–32. http://dx.doi.org/10.5018/economics-ejournal.ja.2019-31

Received October 24, 2018 **Published as Economics Discussion Paper** November 8, 2018 **Revised** February 25, 2019 **Accepted** April 11, 2019 **Published** May 17, 2019

1 Introduction

Historically, the study of tax effort at the sub-central level has related mainly to two issues. On one hand, the main subject of analysis in barely decentralised governments is the high degree of financial dependence on transfers from the central government, and the pernicious effects of the lack of fiscal accountability, as seen in Spain in the first two decades of the regional decentralisation model. On the other, the literature associated with equalisation transfers mainly studies the analysis and construction of indicators of tax need and potential tax revenue, but not tax effort, even though this is an indicator sometimes considered in the formulas determining the amount of these transfers, as happened explicitly in the first stages of the Spanish regional financing model. We see, therefore, that the international literature rarely quantifies the real exercise of fiscal accountability at the sub-central level, unlike the high level of attention to this matter for central governments, and only occasionally does it propose to determine the real causes explaining the degree to which this sub-central tax autonomy is exercised.

In fact, as tax decentralisation progresses and important taxes are assigned to regional governments, giving them greater regulatory power over essential elements of those taxes (e.g., tax credits and tax rates), ever-larger differences are created between both tax rate levels and the configuration of many of these taxes. This process of increasing territorial differentiation in tax matters is concerning: it can mean a considerable increase in the costs of tax collection and tax compliance, it facilitates competition to attract mobile tax bases, it makes the tax differences between territories less transparent, and it makes it more difficult to calculate theoretical tax revenue and tax effort, and thus the necessary equalisation transfers. These concerns are also present in Spain, as the Informe de la Comisión de Expertos para la Revisión del Modelo de Financiación Autonómica (2017) warned.

Alongside this, in the current context of mutual reproaches between levels of government, caused by budget imbalances and the strict financial restrictions associated with the effects of the economic crisis, examining regional tax behaviour will let us test the veracity of the claims of this level of government to be the victim, or the central government's accusations of a lack of regional fiscal accountability, an aspect which is also present in the Spanish case, with accusations of financial disloyalty flying between these levels of government.

For this reason, the goal of this work is to quantify the use regional governments make of their potential tax capacity, that is to say, the tax effort, and examine the causes explaining their tax effort, based on an empirical exercise for the Spanish regions during the period 2002–2012. This work presents several methodological peculiarities differentiating it from the emerging international literature on tax effort at the sub-central level. First, we use several frontier methods, some of which have rarely been applied in this field. Second, we extend the explanatory hypotheses of potential tax revenue, by combining general indicators with other specific indicators of tax capacity. And finally, we use a temporal approach with panel data considering total tax revenue, rather than the cross-section estimates more often found in this type of study, or the partial works which consider only one tax.

The work is structured as follows. In the next section, we define the concept of tax effort, and review the available literature on the subject and the different methodologies applied in its study. In the third section, we briefly explain the Spanish regional funding system which is the subject of our study. In the fourth section, we propose explanatory hypotheses for the tax potential (or tax capacity) and tax effort of those regions. The fifth section presents the estimates and calculations of tax effort and its determinants. The results show that the Spanish regions have exerted a strong tax effort, and that in general, they do not have much room to manoeuvre to raise tax revenues, although we find significant asymmetry placing some regions at the limits of their potential tax income, while others still have considerable fiscal space. The work concludes with a section of final thoughts.

2 Review of the literature on tax effort

The concept of tax effort is subjective and hard to evaluate, as it is not directly observable. This is attested by the fact that several approaches to it have been suggested in the literature, but none has been universally accepted as satisfactory. The most widely recognised tendency in the literature considers tax effort to be the degree to which a jurisdiction effectively uses its tax capacity. In other words, it can be expressed as the quotient between the real tax revenue obtained by a jurisdiction and its potential tax revenue or tax capacity. The tax capacity of a jurisdiction can be defined as the volume of tax resources which a government can obtain when making full use of its regulatory power over the taxes within its reach, with effective management of them (legal tax capacity). However, an economic approach is normally used, which determines the maximum tax revenue a jurisdiction can obtain given its economic, social, institutional, and demographic characteristics (economic tax capacity).[1]

In this way, mathematically, the numerator of the tax effort (the exercised tax capacity or real revenue collected) depends on the action of the government, as higher tax rates, or more intense efforts in tax management and inspection leading to lower tax evasion, raise the effective tax collected. Meanwhile, the denominator of the tax effort would correspond to the potential tax resources which a government could obtain using the tax bases available to it (tax capacity). This denominator would be independent of the action of the government (Martínez-Vázquez and Boex, 1997), and as it is unobservable, this figure is difficult to quantify. The goodness of the tax effort indicator would thus depend on the quality of the measurement of the denominator (or tax capacity).

In practice, various ways have been put forward to measure this tax capacity or denominator of tax effort. This has sometimes been done through macroeconomic indicators such as the income or wealth of the territory, which is a straightforward and inexpensive option. Unfortunately, this method may not reflect the true capacity of sub-central governments to obtain resources through their own taxes, and may not capture the uneven distribution of specific tax bases, such as those linked to mineral resources, for example.[2] An alternative method to approximating tax capacity is the Representative Revenue System (Advisory

[1] The theoretical aspects of the tax capacity and tax effort concepts can be reviewed in Frank (1959), ACIR (1988), Gold (1986), Kincaid (1989), Bird and Slack (1990), Dahlby and Wilson (1994), and Cyan, Martínez-Vázquez and Vulovic (2014).

[2] Mikesell (2007) and Costa (2008) conduct an exhaustive review of the advantages and disadvantages of each method of estimating tax capacity.

Commission on Intergovernmental Relations-ACIR, 1988). This method calculates the potential revenue governments would obtain from their own taxes if they applied a standard tax rate (López-Casanovas and Castellanos, 2002; López-Laborda, 2015) to the different tax bases available to them, managing the taxes with a standard efficiency level. This line would be followed by Hy et al. (1993) and Yilmaz et al. (2002) and more recently by the proposal of De la Fuente (2013), who applies the average tax rate imposed by the set of regions. Although the Representative Revenue System has significant theoretical advantages, it has the disadvantage that the regional tax bases are not always available, making it necessary in many cases to resort to proxies of those tax bases, with varying success, depending on the information available in each case. This in turn means making subjective decisions in the process of evaluating the different types of taxes. Another disadvantage is that if the decentralised tax bases are not closely linked to regional income, resources may be transferred from low-income regions to rich ones through equalisation grants.[3] And depending on the standard tax rate used, it can lead to a certain instability in the revenues of sub-central governments by incentivising strategic behaviour, such as setting sub-optimal tax rates (Herrero and Martínez-Vázquez, 2007; López-Laborda, 2015).

Nevertheless, the econometric method has been the most popular for empirical studies of tax capacity. This is an indirect method used to estimate tax capacity based on regressions of observed tax revenue according to objective, non-manipulable indicators, used as proxies for the tax bases. The first econometric applications used the least squares analysis (OLS) to estimate the average tax capacity of a jurisdiction, given its tax bases, economic structure, institutional aspects, and demographic trends. More recently, a novel methodology based on the stochastic frontier of production possibilities, proposed by Aigner et al. (1977), has begun to come into use. This methodology applies maximum likelihood techniques to estimate a stochastic tax frontier, as it is understood to provide a better fit for the potential tax capacity of a jurisdiction than the average behaviour provided by the OLS approach. In this way, the tax capacity of a jurisdiction will be considered as the maximum revenue level it could obtain with a virtuous use of its tax bases and efficient management of its taxes, taking as a benchmark the best results reached by the set of jurisdictions with similar conditions over the whole period considered.

The literature includes various papers which calculate tax effort based on tax capacity measured with one or other of these techniques. First, Frank’s (1959) index is a pioneering example of the school of thought which uses a macroeconomic approximation of tax capacity as the basis for calculating tax effort. This index defines tax effort as the quotient between the fiscal pressure of the jurisdiction and its per capita income (macroeconomic indicator of the tax capacity), but as tax pressure is the proportion of national income collected in taxes, mathematically Frank’s Index is transformed into the quotient between per capita collection and the square of per capita income. This means that the drawback of this index is that it shows excessively high values in low-income territories, even if their citizens pay low taxes. An application of this index for local governments in Spain can be reviewed in Cordero et al. (2010).

[3] Insofar as the tax bases of the taxes reflect patterns of consumption or use of resources by the regions, rather than their available resources or purchasing power, these may not be valid factors for calculating their tax capacity, as they actually reflect the economic decisions of the citizens (Barro, 1986, and Bird, Martinez-Vazquez, and Torgler, 2008).

Second, there are notable works by Bahl (1972) and Hy et al. (1993) which estimate tax efforts based on tax capacities calculated according to the Representative Revenue System. The Relative Tax Effort Indexes of Zabalza and Lasheras (2000) and López-Casasnovas and Castellanos (2002) for Spain can also be included in this line of approach. In fact, these Relative Tax Effort Indexes are calculated as the quotient between effective collection in each region and the expected collection based on standard or joint collection behaviour (or tax capacity). This expected collection is obtained by multiplying a theoretical tax rate (which might be the overall average rate, the observed rate, or the regulatory rate) by the expected tax base, territorialised through variables intended to approximate the tax revenue capacity of each region.

Third, there are papers which calculate the tax effort by comparing real tax revenue with the estimated revenue calculated using econometric methods. With the OLS estimate of tax capacity, the error term can be positive or negative, so that the tax effort calculated based on that estimate can be over 100%; in other words, a jurisdiction may deviate above or below the predicted average. There are many papers using this methodology to study tax effort at the national level (from the pioneering Lotz and Morss, 1967, to the recent work of Le et al., 2012); however, there are hardly any studies at the sub-central level (Herrero and Martínez-Vázquez, 2007, and Cordero et al., 2010, for Spain). Instead, the SFA estimate of tax capacity is based on the idea that no economic agent can be located beyond the frontier, so that the tax effort obtained by comparing real tax revenue with the frontier or the potential revenue estimated with stochastic frontier analysis cannot exceed 100%. Thus, any deviation from the frontier represents each jurisdiction's margin for manoeuvre to raise its revenue to the "potential" maximum. This methodology has been used in a few studies of tax effort at the central level, such as Pessino and Fenochietto (2010), and Cyan et al. (2014); and in the regional-level works of Jha et al. (1999), Ramírez and Erquizio (2011), Garg et al. (2017), and even the preliminary approximation of Medina (2012) for the Spanish regions. Also relating closely to our research is the literature using stochastic frontier techniques to analyse efficiency in the collection of the taxes which a jurisdiction has established (Alm and Duncan, 2014).

Fourth, although to a much more limited extent, other techniques can be found in tax effort studies, which although not parametric, share the frontier approach discussed above (in SFA), and are frequently found in evaluations of the efficiency of units of production. One of these is Data Envelopment Analysis (DEA), which has been used to measure the tax effort of municipalities in Colombia (Departamento Nacional de Planeación, 2005) and in several states in India (Thirtle et al, 2000, and Rajaraman and Goyal, 2012). Data Envelopment Analysis (DEA) is a linear programming technique which facilitates the construction of an enveloping surface for calculating a synthetic indicator of relative efficiency. Regions obtaining the maximum level of tax revenue from the tax bases they use would be on the frontier, so that the tax collecting margin of other regions could be measured by their distance from the frontier. This is a technique which lets tax effort be determined without the need to explicitly calculate tax capacity. Given that this is a non-parametric method, there is no need to know the functional form of the input-output relationship; neither is it a statistical method, so there is no need to set a probabilistic distribution of inefficiency.

Its non-convex version, the Free Disposal Hull (FDH) model, has been less used - as far as we know, only Mattos et al. (2011) have used it to measure efficiency in the collection of municipal taxes in Brazil. FDH is a special case of DEA, with the distinguishing feature of not requiring convexity, and that the benchmark units for comparison of the collections are real decision units, giving meaning to the comparison between decision units. With FDH, a jurisdiction which does not exploit its maximum tax-gathering potential will be compared with a real jurisdiction which obtains more tax income in similar socioeconomic conditions, rather than a virtual one constructed from linear combinations (as in the DEA model).

Recently, extensions to this methodology have appeared, known as Order-m (Cazals et al., 2002) and Order-α (Aragon et al., 2005) partial frontier approaches. These approaches do not envelop all the data, and although they are beginning to be used in business efficiency studies, they have not yet made an impact in the field of fiscal federalism, except for the contribution of Vallés-Giménez and Zárate-Marco (2017). These non-parametric partial frontier methods are interesting because they permit atypical efficient or super-efficient observations - i.e., beyond the estimated tax revenue frontier, making it possible to greatly reduce sensitivity to errors of measurement and outliers. Thus, this method can reduce the possible impact on tax revenue comparisons of, for example, regions with significant unusual taxation features or with natural resources which facilitate high tax revenues.

To sum up, a comparative review of the literature enables us to conclude that there is ample international empirical evidence for calculating tax effort at the national level, although there are still few empirical studies of sub-central governments. The problem of the availability of applied research is exacerbated by the lack of necessary information in many countries, especially for works that consider several years at the same time (panel data). Meanwhile, the literature offers a wide range of techniques for calculating tax effort, allowing us to take advantage of the merits of each one and check the robustness of the results obtained, while being aware of the limitations of each approach.

3 The Spanish regional funding system

Our study will consider the set of Spanish regions in the common funding system (therefore excluding the Basque Country and Navarre, the two regions with their own special funding system), as it presents a series of characteristics making it ideal for analysing these aspects. On one hand, in the Spanish regional funding system, jurisdictions have a high degree of tax autonomy, which allows for a large enough fiscal space for heterogeneous tax behaviour to appear within a common national framework. Each region has the autonomy to establish its own taxes and specify certain elements in the assigned taxes (tax rates, tax credits, allowances), so the regional tax scenarios vary widely. On the other hand, until 2002 the equalisation transfers model for Spanish regions considered their tax effort as one of the factors determining the allocation of grants, though the tax effort indicator used was not calculated appropriately.[4] Also, the aggregate budgetary balance at the regional level masks the coexistence of heavily indebted regions (e.g., Catalonia and Valencia) alongside financially sound ones (e.g., Galicia and the Canary Islands). This heterogeneity is extremely interesting, because by studying how the different Spanish regions have used their tax authority, we can assess whether they are justified

[4] Since then, tax capacity has been the basis for the regional financing system.

in borrowing to mitigate the shortfalls of the regional funding system, or whether the central government should intervene to correct the low sub-central tax effort.

Comparatively, the Spanish regions are among the sub-central jurisdictions with the most autonomy to manage their spending and revenue, although historically they have had a smaller margin of discretion. They can support their spending policies with transfers from higher levels of government through various funds, borrowing (with strict limits), and a plethora of their own and assigned taxes. The regions have gradually introduced their own taxes, where they are responsible for every aspect of performance, application, and regulation. These are usually environmental taxes with low potential revenue, given the restrictions of the Regional Funding Act (LOFCA) for establishing taxes on matters already taxed by other levels of government. For a detailed list of the tax measures adopted by each region, see the annual reports published by the Ministry of Finance and Civil Service (Ministerio de Hacienda y Función Pública- MHFP).

In contrast, the taxes whose collection is, totally or partly, assigned by the central government to the regions (hereafter, assigned taxes), provide them 90 per cent of their non-financial revenues, and as shown in Table 1, the regions have a certain degree of regulatory or decision-making power over many of those assigned taxes, which they can exercise within limits. The regions have more regulatory power over Inheritance Tax (ISD), Estate and Property Transfer Tax (ITPAJD), Wealth Tax (IP), Gambling Taxes (TJ), and Personal Income Tax (IRPF), only 50% of which is assigned, than over the Special Tax on fuels (IH) and the Tax on Certain Forms of Transport (IDMT), where the regions have some power over the tax rate. Meanwhile, they have no regulatory power over Value Added Tax (VAT), Special Taxes on the manufacture of alcoholic drinks and tobacco, or the Special Tax on electricity (IE). In these taxes the regions only have territorialised shares of 50%, 58%, and 100% of revenue, respectively. The regions can also set surcharges on certain State taxes. And additionally, as shown in the last column of Table 1, they have powers relating to management of the IP, ISD, ITPAJD, TJ, IDMT, and the Tax on Retail Sales of Certain Hydrocarbons, included in the IH since 2013.

Thus, the volume of tax revenue in the Spanish regions may be influenced by at least five factors. First, regional governments have discretion to set tax rates for the majority of assigned taxes. Second, in many cases they also have regulatory power to establish other tax elements, such as exemptions, rebates, and tax deductions. Third, the regions manage, monitor and inspect taxation, tasks which may affect the total tax base and thus, tax revenue. Fourth, they can use techniques which combine the previous three mechanisms, so that if the amount of the tax base is increased they can reduce the tax rate applied, or establish more deductions or rebates, for example. And finally, they can establish their own taxes, and surcharges on certain other taxes. All of this means that the regions have the power to make decisions on approximately one third of total regional tax revenue, and that there are sufficient optional elements to give them a significant capacity to affect regional taxes. For this reason we think it is important to quantify the available or unexploited tax margin within the reach of this level of government.

Table 1: Taxes assigned to the Spanish regions, regulatory and management powers*

TAX	ASSIGNED REVENUE	REGULATORY POWER	APPLICATION AND MANAGEMENT
Wealth Tax (IP)	100%	Minimum exempt, tax rate, deductions and rebates.	Regional
Inheritance Tax (ISD)	100%	Rebates, tax rate, existing assets, deductions and allowances.	Regional
Estate and Property Transfer Tax (ITPAJD)	100%	Tax rate, deductions and rebates (except in IOS modality).	Regional
Gambling Taxes (TJ)	100%	Exemptions, tax base, tax rates and flat rates, accrued and allowances.	Regional
Personal Income Tax (IRPF)	50%	Regional tax rate and deductions, regional personal and family minimum (±10%)	Central
Value Added Tax (VAT)	50%	No, as required by the EU. Shared tax.	Central
Special taxes on alcohol and tobacco	58%	No, as required by the EU. Shared tax.	Central
Special tax on fuels	58%	Regional tax rates.	Central (Tax on retail sales of certain fuels: Regional)
Special tax on electricity	100%	No, as required by the EU. Shared tax.	Central
Tax on certain forms of transport (IDMT)	100%	Tax rates (+15%).	Regional

* Non-assigned taxes include Corporate Tax, Non-Resident Income Tax, Carbon Tax, tax on insurance premiums, etc.

Source: By the authors.

4 Factors explaining the stochastic tax frontier and tax effort of the Spanish regions

Our study is based on the implementation of the Stochastic Frontier Approach (SFA). The SFA is a parametric technique which is increasingly popular around the world in empirical studies seeking to know what percent of total tax capacity governments are capturing through their tax efforts. To answer this question requires identifying governments' maximum potential tax revenue, and the SFA becomes the most appropriate technique for that purpose, because that is what SFA does: it identifies the maximum potential of tax collection, while non-frontier analyses only estimate the mean of the tax collection. Moreover, SFA allows us not only to determine the level of regional tax effort, but also to identify the factors explaining the different levels of tax effort of the regions, estimating them simultaneously with potential or frontier tax revenues and their determining factors.

Statistically, SFA involves specifying a regression model for the tax frontier, with two error terms. With an output approach, this could be represented as follows:

$$ln\ TAX_{it} = \beta_0 + \Sigma^m{}_{k=1}\beta_k\ lnx_{kit} + v_{it} - u_{it} \qquad [1]$$

where TAX_{it} will be the tax revenue for the region *i* in the year *t*, with $i = 1, 2, \ldots, 15$ and $t = 2002, \ldots, 2012$; x_{kit} represents a vector of values corresponding to *k* relevant variables to explain the tax capacity of region *i* in the year *t*; β_k corresponds to a vector of parameters to be estimated; and β_0 is the common constant for all the regions in the year *t*. The error term, v_{it}, represents the usual statistical noise, i.e., everything beyond the control of the region, and is assumed to be independent and identically distributed as a $N(0, \sigma_v^2)$. The second error term, u_{it}, represents the error in obtaining the maximum amount of revenue for given inputs or tax bases, and would be the function of variables z_{it}, which may vary over time and would include observed heterogeneity.

$$u_{it} = \delta z_{it} + w_{it}, \qquad [2]$$

where δ is a coefficient vector to be estimated and w_{it} is the error term.

Endogenous variable

We took as endogenous variable (output) the regional tax revenue, TAX, considering all tax sources (own and assigned taxes, with and without regulatory power) which may be distributed very unevenly among the regions due to uncontrollable external factors, as this is the only way to avoid the risks arising from the possible substitutability and interdependence of the different forms of obtaining tax revenues by this level of government (ACIR, 1988).[5]

Factors determining tax potential or tax capacity

To choose the inputs or explanatory variables of the tax potential, we considered the available empirical evidence on central and sub-central tax behaviour, and performed a series of regression analyses to select the variables which best measure the tax capacity of the Spanish regions, taking always into account that tax capacity is independent of government decisions or actions, which excludes the consideration of variables such as tax rate. Specifically, we estimated the real revenue collected for each assigned tax and for the total aggregate, according to the main macroeconomic regional indicators which can explain that revenue, and alternatively, a series of proxies of their respective tax bases (as the territorialised tax base information needed to estimate revenue from the taxes considered does not exist).[6] Given that

[5] Although regional governments do not have the power of decision over a given tax, if we do not include the revenues derived from its tax base, x, and that tax base were unevenly distributed among the regions, we would be undervaluing the real tax capacity of the regions with a relatively large base x, and overvaluing those in the opposite situation.

[6] We decided not to include the subject of vertical tax externalities (e.g., Dahlby and Wilson, 2003; Andersson et al, 2004), given that although in the Spanish case the main tax types are shared between the central and regional levels of government (Personal Income Tax, Value Added Tax, and excise taxes are partly assigned), it is really only in Personal Income Tax where tax bases with regulatory power for the regions are shared. Also, the fact that Personal Income Tax has gradually been decentralised, and that during the initial decentralisation period the regions merely

assigned taxes represent approximately 90 per cent of the non-financial revenues of the regions,[7] the explanatory variables of the assigned taxes will also provide a good explanation of the endogenous TAX or total tax revenue. Table A.1 of the Appendix shows the definition of each variable used, the source the data were obtained from, and the main descriptive statistics of the variables. The matrix of correlations is presented in Table A.2. All the nominal variables were deflated according to the corresponding regional price index.

These estimates can be seen in Table A.3 in the Appendix. The upper part of the table shows the estimated revenue (for each assigned tax and for the aggregate) according to the specific tax bases, and the lower part shows the results for the general indicators of tax capacity (income and population). We took variables in logs. The estimation uses the robust errors method proposed by Driscoll and Kraay (1998) and adapted by Hoechle (2007). This methodology generates robust estimates of tax capacity and can be used when the residuals are nonspherical, and without the need for the residuals to be homoscedastic or for absence of serial and contemporary correlation. In our case, this is recommended by the Wooldridge, Wald, and Pesarán tests, respectively. Other methods which would let us simultaneously eliminate the problems mentioned are Parks-Kmenta feasible generalized least squares, and Beck and Katz's panel corrected standard errors, although the former cannot be used when T<N, as in our case, and the latter perform better with smaller samples.

The results of these estimates demonstrate that the two general indicators of tax capacity often used in the literature, population (POP) and income (INCOME), provide a fairly satisfactory explanation of revenue for most individually considered taxes and for the aggregate; although it seems to be necessary to include a specific proxy for the tax bases linked to wealth and gambling in order to mitigate the defects shown by general indicators in the individual estimation of these taxes (IP and TJ, respectively).

Thus, as seen in Table 2, to estimate the tax potential or equation [1] of the stochastic frontier model, we used as explanatory variables or inputs the two general indicators of tax capacity (POP and INCOME), the proxies of the wealth tax base (stock of private capital, STOCKP) and the gambling tax base (regional expenditure on gambling, GAMBLINGEXP), and different features of the institutional context arising from the heterogeneity of the sample, which this technique lets us capture with dummy variables in the tax frontier equation. With the dummy variable DPROV, we identify the single-province regions, which enjoy both regional and provincial revenues, as they assume the responsibilities of the Provincial Governments; with the qualitative variable CAN, we identify the region of the Canary Islands, given the unique features of its tax system, associated with its characteristics as an ultra-peripheral region of the European Union (article 349 of the Treaty on the Functioning of the European Union); and with the trend variable, TEND, we capture the impact of the passage of time on tax revenue and the learning effect in the regions, which have seen their tax autonomy increase significantly from 2002 (the first year of our sample) when more taxes were assigned in line with Law

occupied the "fiscal space" vacated by the central government and scarcely exercised their tax autonomy, makes the existence of vertical tax externalities very unlikely for the period being studied.

[7] Non-financial revenues are the income the region receives in the form of taxes, current and capital transfers, income from assets, and income from the sale of real investments (Chapters 1-7 of the revenue budget).

Table 2: Results of the estimates of tax potential

	OLS$_{FE}$	**SFA**	**SFA-INCOME$_{IV}$**	**SFA-INCOME$_{LAG}$**
		Coef z	Coef z	Coef z
Tax frontier				
INCOME	0.8307** 6.94	0.8325** 10.43	0.6394** 6.12	0.8644** 7.97
POP	-0.0557 -0.51	0.3087** 5.56	0.3237** 5.07	0.2160** 3.15
STOCKP	0.1839 1.70	-0.1503 -1.79	0.0864 0.85	-0.0572 -0.55
GAMBLINGEXP	0.0437 0.68	-0.0482 -1.33	-0.0903* -2.17	-0.0798* -1.97
CAN	-0.3661** -10.30	-0.3847** -5.69	-0.4298* -9.48	-0.3969** -9.71
DPROV	-0.0089 -0.38	-0.0293 -1.57	-0.0465* -2.09	-0.0804** -3.35
IP0911	-0.2469** -4.02	-0.1517** -13.2	-0.1581** -12.43	-0.1816** -13.68
TEND	0.0506** 4.53	0.0379** 11.85	0.0282** 8.49	0.0297** 8.56
CONS	4.7971** 8.39	8.7714** 21.93	7.9885** 16.28	8.6022** 18.06
Unexploited tax potential				
dPOLITCOLOUR		0.1141* 2.11	0.0803 1.45	0.1352 1.8
dSINT		0.0537 1.29	0.0622 1.41	0.0718 1.3
ACTIVISM1		-0.2265** -2.79	-0.2557** -2.72	-0.2676** -2.53
ACTIVISM2		-3.11E-06** -3.14	-3.30E-06** -3.2	-3.63E-06** -2.95
TRANSFREV		0.0005** 6.43	0.0005** 5.73	0.0006** 4.43
PATREV		-2.1831 -1.00	-1.1433 -0.49	-2.5578 -0.87
NFEXP		-0.0003** -4.1	-0.0003** -3.67	-0.0004** -3.11
FEXP		-0.0006 -1.75	-0.0006 -1.6	-0.0009 -1.84
DENSITY		0.0011** 3.74	0.0011** 3.45	0.0015** 3.51
POPGROWTH		0.00001 0.01	-0.0009 -0.31	-0.0002 -0.06
QMANAG		0.0015 0.98	0.0022 1.3	0.0022 1.06
CRISIS		0.3753** 4.6	0.3732** 4.14	0.3936** 3.33
GDPGROWTH		-0.01386 -1.64	-0.0161 -1.73	-0.0289* -2.24
CONS		0.0127843 0.04	-0.0321 -0.1	-0.059381 -0.14

R^2	97.64			
θ		0.0683** 8.03	-0.0626** -6.93	-0.0700** -6.7
σ_u^2		0.1359** 8.52	0.1414** 8.23	0.1549** 6.62
σ_v^2		0.0168** 3.02	0.0206** 2.69	0.0224** 3.11
Ho: $\gamma = \sigma_u^2/\sigma_v^2 = 0$		8.0876** 491.17	6.8523** 356.12	6.9141** 290.95

* $p < 0.1$, ** $p < 0.05$

21/2001. We also include a qualitative variable (IP09-11) which captures the years 2009–11, when in practical terms no Wealth Tax (IP) was collected.

At the top of Table A.4 we have listed the SFA papers using these variables to estimate tax capacity. This table shows how we extend the explanatory hypotheses of potential tax revenue, by combining general indicators (usual in comparative literature) with other specific indicators of tax capacity which seek to show characteristics specific to our country or our legislation, and which have not been used to date in works of this kind. We tried including other variables (the unemployment rate, the weight of the agricultural sector, etc.) but they were not significant or did not improve the model. We also tested the regional tax fraud levels, based on the estimates of the Finance Ministry Union (GESTHA), but this variable was not significant either, probably due to the lack of an official estimate of suitable quality.

Explanatory hypotheses of tax effort

In taxation, the difference between real output and the output determined by the tax frontier can only be interpreted as the unobtained part of the potential tax revenue, but it cannot be considered a measure exclusively associated with inefficiency. Therefore, we have proposed that heterogeneous behaviour in the use of potential tax revenues, or equation [2], could be caused by several factors.

Political variables: We think that each regional government may deliberately make a high or low tax effort. To take this into account, on one hand, we have included two qualitative variables: dPOLITCOLOUR, intended to show whether tax effort depends on the ideology of the party in power in the regional government, and dSINT, which identifies when the regional and federal governments are ruled by the same party and thus in harmony with each other. On the other hand, we have included two variables intended to capture the political will of a set of regions which have been strongly committed to exercising their tax autonomy (although we must take into account that the regions are left with hardly any taxable areas where they can apply their own taxes). The variable ACTIVISM1 identifies the tax effort of regions which the Order-α technique puts outside the partial revenue frontier due to more active behaviour in terms of tax effort (this can be seen in Table 3 of the section on results). As we have explained, partial frontier methods let us identify regions whose tax behaviour is atypical or super-efficient - in other words, located beyond the estimated potential tax revenue frontier, due to their

Table 3: Analysis of potential endogeneity (Wu-Hausman and Durbin)*

Variable	Wu-Hausman-F (Prob>F)	Durbin-X_i^2 (Prob>X_i^2)	Sargan-X_i^2 (Prob>X_i^2)
INCOME	0.023399 (0.8786)	0.024891 (0.8746)	7.21677 (0.0653)
POP	2.65872 (0.1050)	2.78252 (0.0953)	2.70662 (0.4391)
dPOLITCOLOUR	0.185597 (0.6672)	0.203905 (0.6516)	5.42109 (0.0665)
dSINT	2.11375 (0.1481)	2.31513 (0.1281)	0.456581 (0.4992)
TRANSFREV	0.781599 (0.3781)	0.858357 (0.3542)	3.23695 (0.1982)
NFEXP	0.7715 (0.3812)	0.844308 (0.3582)	1.00085 (0.6063)

* To test the endogeneity of each of these variables, we have used in each case the equation of the model in which they act as explanatory variables. That is to say, we have used the equation [3] of the text and consequently the TAX variable as dependent to test the endogeneity of INCOME and POP. To test the endogeneity of the other variables (dPOLITCOLOUR, dSINT, TRANSFREV and NFEXP) we have used the equation [4] and therefore the tax effort estimation as dependent variable.

specific fiscal characteristics: either because the territorial distribution of some tax bases is extremely uneven, or because there happens to be an especially intense taxable activity there (e.g., property sales in coastal tourist areas), among other possible causes. These regions are Catalonia, Madrid, Andalusia, and Valencia.[8] The variable ACTIVISM2 captures the volume of revenues from the region's own environmental taxes, which is an unmistakable sign of the will to exercise the tax autonomy the law confers on the regions to obtain tax revenue.

Budget variables: The region's tax behaviour may also be determined by its budget situation, and so our model takes into account the existence of income sources other than taxes, and the spending needs of the regions. On the income side, we included the variable TRANSFREV, which measures the volume of income from transfers which the Autonomous Regions receive from the State and from higher levels of government, and which may generate fiscal illusion and thus lead to low tax effort; and the variable PATREV, which shows income from the management of regional assets, specifically, revenue from assets and the disposal of real investments, which may complement or replace tax income. As indicators of the will to exercise self-government and the need for income, we have included the volume of non-financial current expenditure (NFEXP) and financial expenditure (FEXP), which will enable us to check whether the regional funding model has allowed a certain margin of financial autonomy, requiring greater tax effort from jurisdictions with higher spending.

[8] We have considered the alternative of working with a dummy variable which gives the value 1, to the four regions mentioned (Catalonia, Madrid, Andalusia, and Valencia), and 0 to the others. However, this gives worse results, as there are notable differences between these regions in terms of their atypical tax effort (showed in table 3), which is captured better with the proposed proxy.

Demographic variables: We have also considered demographics with the variables of density (DENSITY) and changes in population (POPGROWTH). Population density can influence tax effort in different ways. On the one hand, a higher concentration of people should make taxation easier (Cyan et al., 2014), and facilitate the use of economies of scale in tax management; but, on the other hand, larger population density could also encourage informal activities that are difficult to tax (Mkandawire, 2010). Population growth rate is associated with higher inefficiency in the tax system because it is difficult to administer a rapidly rising population of taxpayers (Bahl, 2004 and Le et al, 2008), so the tax system may lag behind the ability to capture new taxpayers (Bird et al, 2008). Until 2009, the effects of the variable POPGROWTH may have been reinforced by the fact that the Spanish funding system set the population at the level of the base year considered, obliging jurisdictions with faster demographic growth to make a greater tax effort.

Inefficiency: Regional differences in tax behaviour may arise from inefficiency in the tax collection and management process, which could be due to poor management, the use of obsolete technology, a lack of suitable human resources, corruption, tax evasion, etc. To capture this inefficiency, we have included the variable QMANAG, which we have constructed as the quotient between the non-financial current revenues the region really receives and the revenues it budgets for. Thus, the closer the real revenues come to the revenues budgeted by the regional government, and therefore the greater the value of this relationship or quotient, the greater will be the tax collection efficiency, or the control of revenues by the regional government, which will favour a higher value of tax effort. On the other hand, real revenues which are very different from the initially budgeted revenues could be a symptom of inefficiency in the tax collection and management process.

Economic cycle: Finally, we have included the variable CRISIS to reflect the impact of the economic and property recession on regional finances and tax effort, and the variation rate of GDP in each region (GDPGROWTH) to see how the tax effort varied with the different amount and intensity of each region's reactions to the cycle.

The bottom of Table A.4 shows the SFA papers using these variables to estimate tax effort.

5 Results of the estimation for the tax frontier and regional tax effort

In light of the hypotheses described above, we used panel data (2002-2012) to estimate equations [3] and [4] of the stochastic frontier model proposed by Greene (2005). This is a time-varying model, the True Random Effect (TRE) model (although the fixed effects approach gives similar results), estimated by applying the maximum likelihood procedure, with the following econometric specification:

$$TAX_{it} = \alpha_i + f(INCOME, POP, STOCKP, GAMBLINGEXP, CAN, DPROV, IP09\text{-}11, TEND) + v_{it} - u_{it} \quad [3]$$

$$u_{it} = g\,(dPOLITCOLOUR, dSINT, ACTIVISM1, ACTIVISM2, TRANSFREV, PATREV, NFEXP, \quad [4]$$

FEXP, DENSITY, POPGROWTH, QMANAG, CRISIS, GDPGROWTH) + w_{it}

Estimation using *stochastic* frontier analysis (*SFA*) lets us solve the main problem presented by the conventional approach to tax effort through a mean behaviour regression, by avoiding, as Rao (1993) indicates, letting tax effort become part of the random residue. We must also take into account that the goal of the paper is not only to determine the tax frontier, but also to simultaneously examine the determinants of the regions' tax effort relating to the observed heterogeneity.

The heterogeneity observed in u_{it} is taken into account with equation [4]. We include the efficiency determinants, z_k, as heteroscedastic variables in the inefficiency function. An advantage of this specification is that it lets us estimate the tax effort function simultaneously, as a one-step procedure, with the parameters of the stochastic frontier. This technique has an advantage over the alternative two-stage method where the first stage involves estimation of a conventional frontier model with environmental variables omitted, and the second stage involves regressing these predicted tax efforts on the environmental variables. The two-stage procedure can lead to inconsistency in assumptions about the distribution of the tax effort since the estimates of u_{it} will be biased by the omission of environmental variables in the first step regression. Thus, failure to include environmental variables in the first stage leads not only to biased estimators of the parameters of the deterministic part of the fiscal frontier but also to biased predictors of tax effort (see Kumbhakar and Lovell, 2000).

This specification considers that inefficiency may vary over time, and the inefficiency term excludes unobserved and time-invariant heterogeneity. In other words, on one hand, it is assumed that any structural inefficiency can be corrected during the sample period, so that persistent inefficiency is not included in the inefficiency term. And on the other hand, it is assumed that unobserved and time invariant heterogeneity is captured by the frontier, and thus does not appear as inefficiency. Thus, all the time-invariant effects are treated as unobserved heterogeneity, and are captured by the specific stochastic term of each region, α_i, which is an i.i.d random component. This means that the random effects model has two major advantages: it controls for any omitted variable biases, and it avoids heterogeneity biases in the estimates of technical inefficiency.

The results of estimating the model above, for which we took variables in logs and used the STATA statistical package, are shown in the third column of Table 2 under the heading SFA. The estimated λ is the ratio between the inefficiency and measurement error variability (the so called signal-to-noise ratio σ_u/σ_v), providing information on the relative contribution of both error components in total error term. Thus, as the estimator λ is significant and very high, it is indicating the presence of technical inefficiency, and SFA is confirmed as a suit method for the study, in other words, the need to include unrealised tax effort, *u*, in the tax capacity function. Thus, approaching tax capacity through a conventional mean behaviour function estimated by ordinary least squares (OLS) is not suitable, as λ is indicating that the deviations from the frontier are not only due to the estimation error, but that many of the disparities in terms of tax collection depend on the decisions made by the regional governments themselves, and on inefficiency. In fact, if we divide the variance of *u* by total variance ($\gamma=\sigma^2_u/\sigma^2_\varepsilon$), we obtain that 98.54% of the error term is due to unrealised tax effort. Additionally, as indicated by Belotti et al. (2013), the significance of the parameter θ, which measures the estimated standard deviation of the unobserved heterogeneity, validates the Greene (2005) approach, in which the unobserved heterogeneity of regions must be separated from the inefficiency effects.

To determine whether the endogeneity problems affect certain variables (INCOME, POP, political variables, TRANSFREV, and NFEXP), we have applied the two-stage Hausman procedure and calculated the Durbin (1995) and Wu-Hausman statistics (Wu, 1974 and Hausman, 1978), which can be seen in Table 3.[9] To do this, we have used instrumental variables such as the lagged variables and their variation rates, the expenditure declared by households or the weight of the agricultural sector in regional income. The output of Wu-Hausman and Durbin tests show that we cannot reject the null hypothesis of exogeneity of the variables, and the Sargan test presents strong evidence that we cannot reject the null hypothesis that the overidentifying restrictions are valid. The validity of the instruments used ratifies the Wu-Hausman and Durbin tests.

To make sure that endogeneity was not a problem, we also applied (with the command *xtsfkk* in Stata) a practical maximum-likelihood-based approach proposed by Karakaplan and Kutlu (2017) and Karakaplan (2018) to control for the endogeneity. Their approach, among other things, treats endogeneity as an omitted variable problem and adjusts the error term by a correction factor to solve the problem of inconsistent parameter estimates due to endogenous regressors (for a detailed study of this methodology, see Karakaplan and Kutlu, 2017). Also, their approach lets us check the endogeneity of the variables, testing the joint significance of the components of the η term[10], based on similar ideas to the standard Durbin-Wu-Hausman test for endogeneity[11]. The results of this estimation can be seen in Table 4. The model that ignores endogeneity is named Model EX and the one that captures endogeneity is Model EN. It can be seen that the variable INCOME can present endogeneity problems (at 5% significance, $\eta 1$=-2.102*). Nevertheless, as the technique proposed by Karakaplan and Kutlu (2017) is robust against this problem, it generates unbiased results (Model EN), and it can be seen that the results of both models (EX and EN) are very similar to each other, showing practically the same variables as our initial model as significant, and with the same sign.

In any case, as the proposal of Karakaplan and Kutlu (2017) does not let us simultaneously estimate the tax frontier and the hypotheses explaining the differences in tax effort, we have also tried instrumentalising the variable INCOME, which is the one which can present endogeneity problems (according to Karakaplan and Kutlu, 2017). To do this, we have used two alternative methods, which are conventional and the simplest to implement (Arellano and Bond, 1991, and Wooldridge, 2013) and which meet the conditions of consistency (Baer et al., 1998), which require alternative methods to generate similar distributions of efficiency and similar rankings of Autonomous Regions, identify the same regions as “better” or “worse”, etc. On one hand, we have instrumentalised the variable INCOME, just as we did for the Durbin and Wu-Hausman statistics, with the income lagged and their variation rate, and the weight of the agricultural

9 We ran a conventional endogeneity test through a linear model where one or more of the regressors are endogenously determined. We did this using two-stage least squares (2SLS).

10 Karakaplan and Kutlu (2017).

11 To do this we used the same instrumental variables as for the Durbin and Wu-Hausman statistics.

Table 4: Results of the estimates of endogenous SFA in the style of Karakaplan and Kutlu (2017)

	Model EX	Model EN
Dep.var: TAX		
Constant	5.918*** (0.955)	6.367*** (0.961)
INCOME	0.837*** (0.217)	0.983*** (0.214)
POP	-0.051 (0.085)	-0.062 (0.081)
STOCKP	0.178 (0.241)	0.029 (0.239)
GAMBLINGEXP	0.034 (0.090)	0.046 (0.087)
CAN	-0.372*** (0.058)	-0.368*** (0.054)
DPROV	-0.014 (0.039)	-0.024 (0.037)
IP0911	-0.247*** (0.031)	-0.244*** (0.031)
TEND	0.050*** (0.008)	0.054*** (0.008)
η 1 (INCOME)		-2.102* (0.830)
η 2 (POP)		4.824 (3.516)
η Endogeneity Test		X^2=9.5 p=0.009

Notes: Standard errors are in parentheses. Asterisks indicate significance at the 0.1% (***), 1% (**) and 5% (*) levels.

sector in regional income, and we then added it to our initial stochastic frontier estimation, implemented with the *sfpanel* command of STATA (Belotti et al., 2013). The results can be seen in the fourth column of Table 2, under the heading SFA-INCOME$_{IV}$. On the other hand, as an alternative, we added the variable INCOME lagged by one period to our initial stochastic frontier estimation, presenting the results under the heading SFA-INCOMELAG, in the fifth column of Table 2. It can be seen that the results obtained with these two estimations (SFA-INCOME$_{IV}$ and SFA-INCOME$_{LAG}$) are very similar to those obtained initially (SFA), which would confirm that endogeneity is not a serious problem, and show the robustness and consistency of the model used.

Focusing on the results obtained (Table 2), we can see they are consistent with what we would expect from a theoretical point of view, and with the available but scanty empirical evidence (e.g. Pessino and Fenochietto, 2010, Cyan et al, 2014, and Garg et al., 2017). Although the first instrumentation of the variable INCOME reduces its influence on the tax frontier – from 83% in SFA to 63.95% in SFA-INCOME$_{IV}$-, it seems clear that INCOME is the variable which most determines the explanation of regional tax capacity. As the variables are in logarithms, the coefficients provided by our models are elasticities, so that the results would indicate that when regional income increases by one percentage point, tax capacity rises by 0.63 to 0.86 percentage points. This would be consistent with the fact that a large part of regional tax revenue comes from income taxes[12]. Other works (Pessino and Fenochietto, 2010, or Garg et al., 2017) have also shown this variable to be the most relevant in explaining the stochastic frontier. Population (POP) also plays an important and direct role in explaining tax capacity, with an effect of 21-32%, and the suppression of Wealth Tax caused a reduction of regional tax

12 The website of Spain's Finance Ministry (*http://www.hacienda.gob.es/Documentacion/Publico/Tributos/Estadisticas/Recaudacion/2014/Analisis_estadistico_recaudacion_2014.pdf*) shows the weight of different taxes in the regional budget structure.

potential in the 2009-2011 period. The model also shows the lower tax capacity (39-42%) of the Canary Islands (CAN), given the region's peripheral and isolated nature, its adverse climate and geographical conditions, and the scarcity of natural resources. The fact that VAT and most of the excise taxes do not apply in their territory and are substituted by a regional-unique indirect tax, usually with much lower tax rates, may also be influencing this result. The single-province regions (DPROV) also show lower tax potential, although they can use both provincial and regional revenues. And finally, there is evidence of a growth trend (TEND) in tax potential, due to the passage of time and the learning process after the regions were given regulatory power over assigned taxes, especially from 2002.

Although we have already indicated that the significance of estimator λ confirms that SFA is a suitable method for this analysis, we can test the robustness of the sign of the coefficients of the variables used to estimate tax capacity with the Driscoll-Kraay robust errors method (OLS), adapted by Hoechle (2007), which can be seen in the second column of Table 2, under the heading OLS. Driscoll-Kraay standard errors are well calibrated when the regression residuals are cross-sectionally dependent, so this method would ratify the validity of the explanatory hypotheses of the stochastic frontier

Our models also support most of the explanatory hypotheses of tax effort. Many of the results are in line with the available theoretical and empirical evidence (e.g. Pessino and Fenochietto, 2010, Cyan et al, 2014, and Garg et al, 2017). The variables which capture self-governance and the exercise of fiscal responsibility are decisive for the explanation of tax effort, as are the regions' needs for non-financial expenditure (NFEXP) and for financial spending (FEXP), although these last are only significant at 6-7%. On the side of revenues, the grants received (TRANSFREV) generate fiscal laziness in regions due to the fungibility effect.

Conversely, when the region presents a solid commitment to the exercise of its taxing power or a desire to collect taxes (ACTIVISM1 and ACTIVISM2), the margin of available unexploited tax revenue becomes narrower, i.e. its tax effort increases. The model also reveals that political variables (dPOLITCOLOUR and dSINT) are not relevant in the explanation of tax effort. The management inefficiency proxy (QMANAG) also lacks significance.

A positive sign for the variable CRISIS shows that regional taxes have acted as automatic stabilisers, a result which is consistent with the economic recession and property market crisis, and with the progressiveness of income tax and its relative weight in the total tax revenue. The negative sign of the variable GDPGROWTH would ratify this result, indicating that the more the region has grown, the greater is its tax effort. Finally, the positive and significant coefficient of the variable DENSITY indicates that there are economies of scale in tax management and collection, linked to the concentration of the population.

Columns 3 to 5 in Table 5 present the calculations of the tax effort according to these stochastic frontier approaches. The magnitude of the coefficients obtained in these estimates provides average tax effort levels of around 88% by the Spanish regions, a result which demonstrates that in general, the tax behaviour of sub-central governments has been responsible. This allows us to refute the central government's repeated complaints that the regions are not responsible in tax matters. This result is similar to that of De la Fuente (2013), and Herrero and Martínez-Vázquez (2007), and the results of the Finance Ministry from its calculations using the RRS method, although it is far from the tax-room-to-manoeuvre of Medina (2012) and López-Laborda (2015).

Table 5: Use of tax revenue capacity in Spanish regions

	OLS_{FE}	SFA	$SFA\text{-}INCOME_{IV}$	$SFA\text{-}INCOME_{LAG}$	**Non-parametric frontier methods**		
					Order-α(98)*	**Order-m(160)***	**FDH**
Andalusia	0.9757	0.9094	0.9094	0.9173	1.3668	1.0350	0.9737
Aragon	1.0661	0.9065	0.9065	0.9199	0.8865	0.8302	0.8268
Asturias	1.0939	0.8293	0.8293	0.8650	0.8646	0.8648	0.8646
Balearic Islands	1.2571	0.9224	0.9224	0.9305	0.9488	0.9493	0.9488
Canary Islands	1.0681	0.8944	0.8710	0.9047	0.7017	0.6415	0.6319
Cantabria	1.0550	0.8696	0.8696	0.8548	0.8945	0.8945	0.8945
Catalonia	0.9959	0.9518	0.9534	0.9551	1.2137	0.9022	0.8471
Extremadura	1.1238	0.7332	0.7385	0.7620	0.8445	0.8448	0.8445
Galicia	0.9214	0.7873	0.7794	0.8369	0.9937	0.8750	0.8621
Castilla-León	0.9506	0.8355	0.8293	0.8654	1.0210	0.8067	0.7707
Madrid	1.0163	0.9403	0.9412	0.9332	1.5163	1.0693	0.9948
Castilla-La Mancha	0.9145	0.7959	0.8632	0.8389	0.9106	0.8234	0.8171
Murcia (Region of)	0.9731	0.8941	0.8847	0.9016	0.7594	0.7283	0.7242
Rioja (La)	1.4991	0.8964	0.9073	0.9173	0.9372	0.9372	0.9372
Valencian Community	1.0830	0.9424	0.9352	0.9417	1.3418	1.0360	0.9762
TOTAL	1.0662	0.8705	0.8760	0.8896	1.0134	0.8825	0.8609

* Although it does not appear in this document, we carried out a sensitivity analysis of the efficiency indicators, calculating the frontiers as m = 90, 120, … 200 and α = 90, 95, ... 97. These estimates make the presented results more robust and are available to readers on request.

Source: By the authors.

To check the robustness of these results, we have also calculated the regional tax effort with the Driscoll-Kraay robust errors method and with some of the nonparametric frontier methods explained in Section 2 (i.e., Order-m and Order-α partial frontier methods and the Free Disposal Hull). The results are displayed in columns 2 and 6–8 of Table 5, and show that the average tax effort of the regions ranges from 86.09 to 106.6 per cent, depending on the technique used. These results confirm that hardly any tax room for manoeuvre margin is available, and reveal a highly responsible use of tax autonomy by the Spanish regions. As suggested by Badunenko et al. (2012), the high value of λ could mean that both the SFA estimates and the non-parametric techniques are good, ratifying the similarity of the results.

However, all the cases show evidence of significant asymmetries between regions in the exercise of tax autonomy, reflecting the different territorial priorities at this level of government. Although some divergence can be seen in the results, depending on the analytical technique used, this disparity is due to the different approach underlying each method. In particular, the average behaviour approach (OLS) is very different from the frontier methods (stochastic or non-parametric), although these last are also slightly different, depending especially on whether they permit the presence of super-efficient units.

Leaving aside the average behaviour approach (OLS), which the statistic λ has been shown not to be relevant to this estimation, we observe a clear consensus among all the frontier techniques used as to the jurisdictions making the greatest tax effort: Madrid, the Balearic

Islands, Valencia, and Catalonia. There is slightly less agreement on the regions with the laziest tax behaviour, which we attribute to the inability of non-parametric frontier models to include dummy variables to control for the unique features of the institutions of each region. Murcia, Extremadura, and to a lesser extent Castilla-La Mancha and Galicia, would be the regions making the least tax effort, with some uncertainty regarding La Rioja.

To analyse these tax discrepancies in more depth, we projected the situation of the Spanish regions in terms of tax effort, according to SFA, and per capita tax revenue, in Figure 1. This graph lets us classify the 15 Spanish regions in four groups, which are bounded by average tax effort and average per capita tax revenue. It differentiates between regions with especially low per capita tax collection because their tax effort is low -below average- (e.g. Extremadura and Galicia) and regions with low actual tax revenue but with high tax effort -above average-, because their tax capacity is limited (Andalusia and Murcia). Note that all these jurisdictions with low per capita tax revenue (zones I and II) are those with the lowest per capita income. In contrast, the richer regions show higher tax revenue and greater tax effort (Madrid, Catalonia, Balearic Islands, and Aragon). At the same time, if we look at the situation in these regions in terms of debt (Table 6), we can see how, on one hand, there are regions which, despite a very high tax effort (regions in zones II and III), have not been able to halt the continuous deterioration of their financial situation. Regions such as Catalonia, Valencian Community and Balearic Islands, which make a high tax effort, lead the debt rankings (Table 6), a clear sign that there are underlying design problems in the Spanish regional financing system. But on the other hand there are regions which, although they still have ample room to manoeuvre in terms of tax effort and revenue, have exacerbated their credit situation at a rate which is clearly faster than average (e.g., Castilla-La Mancha and Castilla-Leon), which would justify the central administration's accusations of lack of accountability directed towards some regional governments.

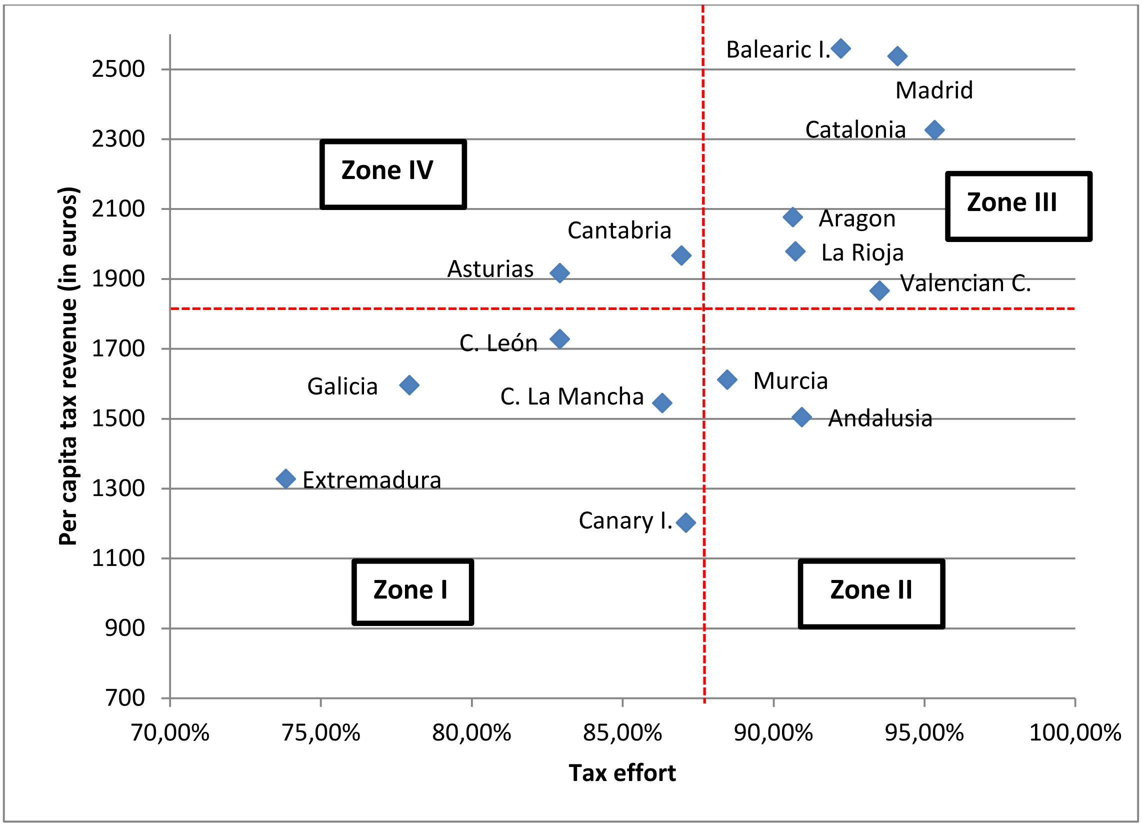
6 Final considerations

We calculated the tax effort of the Spanish regions using different techniques which have rarely been applied in this field, giving us mostly similar results. Thus, this work is contributing to introduce new methodologies in the empirical field, and also to improve the few available works on sub-central tax effort.

For the factors explaining potential tax revenue, the general indicators of tax capacity (income and population), the severe economic crisis, the temporary suppression of Wealth Tax, and the institutional differences between the regions (particularly in peripheral or single-province regions) are determinants in the explanation of the tax frontier.

For tax effort, all the techniques confirm that the regions have exercised their tax responsibilities responsibly, which would refute the central government's complaints accusing

Figure 1: Tax effort (SFA-INCOME$_{IV}$) and tax collection per capita in the Spanish regions



Source: By the authors.

the regional governments of irresponsible tax behaviour. The variables capturing financial autonomy and self-government ratify this conclusion. The greater tax effort made by the regions that want to support more or better non-financial current expenditure, and the active commitment to tax collection of regional governments despite the severe restrictions on their taxing power, all clearly indicate an appropriate and serious exercise of their fiscal responsibility. The significance of population density demonstrates the existing differences in terms of scale economies in tax management and control. We have also found evidence that the design of the transfer system introduces perverse incentives, reducing the tax effort of the regions enjoying the largest grants. Tax effort was also reduced during Spain's severe economic crisis, revealing a countercyclical policy which has been considerably influenced by political ideology, although this result is also consistent with progressive taxation.

Nevertheless, results show asymmetries in the tax effort realised. Some jurisdictions (the poor ones) have a substantial margin for manoeuvre in taxation while others (usually more prosperous regions) are practically at the limits of the possibilities offered by the sub-central funding system. The construction of this map of tax effort lets us see that some regions have acquired debts despite having plenty of room to manoeuvre in terms of unexploited tax potential, and also despite being subjected to strict balanced budget restrictions. This reflects the problems of the system coordinating policies on loans, its faulty design, and the lack of

Table 6: Changes in the credit situation of the Autonomous Regions (2002–2012)

Debt Stock/Income (in %)				**Debt Stock /Population (in euros)**			
Regions ordered according to the Debt Stock/Income ratio in 2012	**2002**	**2012**	**Average annual increase rate**	**Regions ordered according to the Debt Stock/Population ratio in 2012**	**2002**	**2012**	**Average annual increase rate**
Valencian Community	9.7	31.2	22.16	**Catalonia**	1,652	6,998	32.36
Castilla-La Mancha	3.1	27.2	77.74	**Valencian Community**	1,646	6,029	26.63
Catalonia	7.1	26.8	27.75	**Balearic Islands**	740	5,522	64.62
Balearic Islands	3.5	23.9	58.29	**Castilla-La Mancha**	436	4,865	101.58
Murcia (Region of)	3.7	17.4	37.03	**Cantabria**	542	3,444	53.54
Cantabria	3.2	16.7	42.19	**Aragon**	886	3,422	28.62
Galicia	8.2	15.4	8.78	**Rioja (La)**	583	3,278	46.23
Andalusia	7.1	15.1	11.27	**Madrid**	1,528	3,259	11.33
Castilla-León	3.4	14.8	33.53	**Murcia (Region of)**	551	3,165	47.44
Extremadura	6.4	14.4	12.50	**Castilla-León**	550	3,150	47.27
Aragon	4.6	14.2	20.87	**Galicia**	1,167	3,014	15.83
Rioja (La)	3	13.6	35.33	**Andalusia**	962	2,510	16.09
Asturias	5.1	12.5	14.51	**Asturias**	776	2,505	22.28
Canary Islands	3.1	11.8	28.06	**Canary Islands**	529	2,226	32.08
Madrid	6.4	10.7	6.72	**Extremadura**	738	2,212	19.97
MEAN	5.2	17.7	24.04	**MEAN**	886	3,707	31.84

Source: By the authors, based on data provided by Banco de España.

penalties and suitable incentives, attested by the lack of a relationship between the financial burdens of the debt and the tax effort of the regions.

For all these reasons, a legal reform of the regional tax system to raise the tax frontier according to the principle of financial self-sufficiency seems to be necessary in the short term, in order to respond to the financial problems of regions that currently find themselves with no room for manoeuvre in terms of tax revenue, due to being very close to the tax frontier. For this purpose, it could be desirable to increase the regulatory powers of the regions; facilitate their access to new tax bases - their own or at other levels of government; or extend their tax space with types of taxes that are currently shared, such as personal income tax. This would make it possible to raise the potential tax revenue of all the regions and correct the endemic financial problems that have historically afflicted part of Spain's regional financial system, while leaving the final decision on how to use this potential in the hands of each region.

Acknowledgements We would like to thank the three anonymous referees and the co-editor of the journal for their suggestions and comments, which have undoubtedly improved the draft. We gratefully acknowledge the funding support of the Spanish Institute for Fiscal Studies, the Government of Aragon and the European Regional Development Fund (Public Economics Research Group) and the Ministry of Science, Innovation and Universities (Project RTI2018-095799-B-I00). The usual disclaimer applies.

References

Advisory Commission on Intergovernmental Relations-ACIR (1988). *State Fiscal Capacity and Effort*. Information Report M-170, Washington, D.C.: ACIR. https://library.unt.edu/gpo/acir/Reports/information/m-170.pdf

Aigner, D., K. Lovell, and P. Schmidt (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, 6: 21–37. https://pdfs.semanticscholar.org/3d6c/a6a699e579463e0b01526d9842d41f26b314.pdf

Alm, J., and D. Duncan (2014). An International Comparison of Tax Agency Efficiency. *Public Budgeting and Finance*, 34(3): 90–110. https://doi.org/10.1111/pbaf.12043

Andersson, L.; T. Aronsson and M. Wikström (2004). Testing for Vertical Fiscal Externalities, *International tax and Public Finance*, 11 (3): 243–263. https://link.springer.com/article/10.1023/B:ITAX.0000021970.40929.9a

Aragon, Y., A. Daouia, and C. Thomas-Agnan (2005). Nonparametric Frontier Estimation: A Conditional Quantile-based Approach. *Econometric Theory*, 21: 358–389. https://doi.org/10.1017/S0266466605050206

Arellano, M., and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo. Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58: 277–297. http://people.stern.nyu.edu/wgreene/Econometrics/Arellano-Bond.pdf

Badunenko, O., D.J. Henderson, and S.C. Kumbhakar (2012). When, Where and How to perform Efficiency Estimation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(4): 863–892. https://doi.org/10.1111/j.1467-985X.2011.01023.x

Baer, P., A. Berger, G. Ferrier, and D. Humphrey (1998). Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods. *Journal of Economic and Business*, 50: 85–114. https://econpapers.repec.org/article/eeejebusi/v_3a50_3ay_3a1998_3ai_3a2_3ap_3a85-114.htm

Bahl, R.W. (1972). A Representative Tax System Approach to Measuring Tax Effort in Developing Countries. I*MF Economic Review,* 19(1): 87–124

Bahl, R.W. (2004). Reaching the Hardest to Tax: Consequences and Possibilities. *Contributions to Economic Analysis*, 268: 337–354.

Barro, S.M (1986). State Fiscal Capacity Measures: A Theoretical Critique. In H. Clyde Reeves (ed.), *Measuring fiscal capacity*. Cambridge, Massachusetts, Oelgeschlager, Gunn & Hain.

Belotti F., S. Daidone, G. Ilardi, and V Atella (2013). Stochastic Frontier Analysis using Stata. *The Stata Journal*, 10(2): 1–39 https://econpapers.repec.org/article/tsjstataj/v_3a13_3ay_3a2013_3ai_3a4_3ap_3a718-758.htm

Bird, R., and E. Slack (1990). Equalization: The Representative Tax System Revisited. *Canadian Tax Journal,* 38: 913–927.

Bird, R.M., Martinez-Vazquez J. and Torgler, B. (2008). Tax Effort in Developing Countries and High Income Countries: The Impact of Corruption, Voice and Accountability. *Economic Analysis and Policy,* 38(1): 55–71. https://econpapers.repec.org/article/eeeecanpo/v_3a38_3ay_3a2008_3ai_3a1_3ap_3a55-71.htm

Cazals, C., J-P. Florens, and L. Simar (2002). Nonparametric Frontier Estimation: A Robust Approach. *Journal of Econometrics*, 106: 1–25. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1007.4581&rep=rep1&type=pdf

Comisión de Expertos para la Revisión del Modelo de Financiación Autonómica (2017). *Informe de la Comisión de Expertos para la Revisión del Modelo de Financiación Autonómica, Ministerio de Hacienda y Función Pública*. http://www.hacienda.gob.es/Documentacion/Publico/CDI/Sist%20Financiacion%20y%20Deuda/Informaci%C3%B3nCCAA/Informe_final_Comisi%C3%B3n_Reforma_SFA.pdf

Cordero, J.M., R. Fernández, C. Navarro, F. Pedraja, J. Suárez, and A. Utrilla (2010). *Esfuerzo Fiscal de los Municipios Españoles.* Instituto de Estudios Fiscales, Madrid.

Costa, A. (2008). *Un Análisis de las Disparidades Fiscales generadas por la Actividad Turística: Aplicación al Caso de los Municipios Españoles.* Doctoral Thesis, Palma. https://www.uib.es/es/recerca/estructures/grups/grup/AEIT/produccio/tesis_tesines_treballs/2008.html

Cyan, M., J. Martínez-Vázquez, and V. Vulovic (2014). New Approaches to Measuring Tax Effort. In Richard M. Bird and Jorge Martínez-Vázquez (eds.), *Taxation and Development: The Weakest Link? Essays in Honor of Roy Bahl*. Edward Elgar. https://www.e-elgar.com/shop/taxation-and-development-the-weakest-link?___website=uk_warehouse

Dahlby B., and S. Wilson (1994). Fiscal Capacity, Tax Effort, and Optimal Equalization Grants. *Canadian Journal of Economics,* 27: 657–672. https://www.jstor.org/stable/135789?seq=1#metadata_info_tab_contents

De la Fuente, A. (2013). La Financiación de las Comunidades Autónomas de Régimen Común: En 2011. Fundación SEPI, Madrid. http://www.fundacionsepi.es/investigacion/publicaciones/otrasPublicaciones/Financiacion%20CCAA%20en%202011%20(Angel%20de%20la%20Fuente).pdf

Departamento Nacional de Planeación (2005). *Capacidad fFscal de los Gobiernos Territoriales Colombianos*. Departamento Nacional de Planeación-Dirección de Desarrollo Territorial Sostenible, Republic of Colombia.

Driscoll, J.C., and A.C. Kraay (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics* 80(4): 549–560. https://doi.org/10.1162/003465398557825

Durbin, J. (1954). Errors in Variables. *Review of the International Statistical Institute,* 22: 23–32. https://www.jstor.org/stable/i261037

Frank, H.J. (1959). Measuring State Tax Burdens. *National Tax Journal,* 12: 179–185.

Garg, S., A. Goyal, and R. Pal (2017). Why Tax Effort falls Short of Capacity in Indian States: A Stochastic Frontier Approach. *Public Finance Review,* 45(2): 232–259. https://journals.sagepub.com/doi/full/10.1177/1091142115623855

Gold, S. (1986). Measuring Fiscal Effort and Fiscal Capacity: Sorting out some of the Controversies. In H.C. Reeves (ed.), *Measuring Fiscal Capacity*. Boston: Oelgeschlager, Gunn and Hain.

Greene, W. (2005). Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model, *Journal of Econometrics,* 126: 269–303. https://doi:10.1016/j.jeconom.2004.05.003

Hausman, J.A. (1978). Specification Tests in Econometrics. Econometrica, 46: 1251–1271. https://www.jstor.org/stable/1913827?seq=1#metadata_info_tab_contents

Herrero, A., and J. Martínez-Vázquez (2007). La Nivelación en el Marco de la Financiación de las Comunidades Autónomas. Papeles de Trabajo, Instituto de Estudios Fiscales, no. 13/2007.

Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-sectional Dependence. *Stata Journal,* 7(3): 281–312. http://serviciostelematicosext.minhap.gob.es/SGCAL/publicacionliquidacion/aspx/menu Inicio.aspx.

Hy, R., C. Boland, R. Hopper, and R. Sims (1993). Measuring Revenue Capacity and Effort of County Governments: A Case of Arkansas. *Public Administration Review,* 53: 330–337.

Jha, R., M.S. Mohanty, S. Chatterjee, and P. Chitkar (1999). Tax Efficiency in Selected Indian States. *Empirical Economics*, 24: 641–654. https://www.researchgate.net/publication/225177801_Tax_efficiency_in_selected_Indian_states

Karakaplan, M. U. and L. Kutlu (2017). Handling Endogeneity in Stochastic Frontier Analysis", *Economics Bulletin* 37(2): 889–901. http://www.accessecon.com/Pubs/EB/2017/Volume37/EB-17-V37-I2-P79.pdf

Karakaplan, M.U. (2018). XTSFKK: Stata Module to Estimate Endogenous Panel Stochastic Frontier Models in the Style of Karakaplan and Kutlu. Statistical Software Components S458445, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458445.html

Kincaid, J. (1989). Fiscal Capacity and Tax Effort of the American States: Trends and Issues. *Public Budgeting and Finance,* 9: 4–26. https://www.researchgate.net/publication/229694601_Fiscal_Capacity_and_Tax_Effort_of_the_American_States_Trends_and_Issues

Kumbhakar, S.C., and K. Lovell (2000). *Stochastic Frontier Analysis,* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139174411

Le, T.M.; B. Moreno-Dodson, and J. Rojchaichaninthorn (2008). Expanding Taxable Capacity and Reaching Revenue Potential: Cross-Country Analysis. World Bank Policy Research Working Paper Series, no. 4559. https://www.researchgate.net/publication/23550556_Expanding_taxable_capacity_and_reaching_revenue_potential_cross-country_analysis

Le, T.M., B. Moreno-Dodson, and N. Bayraktar (2012). Tax Capacity and Tax Effort: Extended Cross-country Analysis from 1994–2009. Policy Research Working Paper 6252, World Bank, Washington, DC. *http://documents.worldbank.org/curated/en/534031468332056098/Tax-capacity-and-tax-effort-extended-cross-country-analysis-from-1994-to-2009*

López-Casasnovas, G., and A. Castellanos (2002). Aproximación al Concepto de Esfuerzo Fiscal. Mimeograph, Centre de Recerca en Economia i Salut, Barcelona.

López-Laborda, J. (2015). La Medición de la Capacidad Fiscal de las Comunidades Autónomas: Algunas Alternativas. FEDEA Policy Papers, 2016-1, Madrid. http://documentos.fedea.net/pubs/fpp/2016/01/FPP2016-01.pdf

Lotz, J., and E. Morrs (1967). Measuring Tax Effort in Developing Countries. International Monetary Fund Staff Papers 14: 478–499, Washington, DC.

Martínez-Vázquez, J., and J. Boex (1997). Fiscal Capacity: An Overview of Concepts and Measurement Issues and their Applicability in the Russian Federation. International Studies Program Working Paper Series 3, Andrew Young School of Policy Studies, Georgia State University.

Mattos, E., F. Rocha, and P. Arvate (2011). Flypaper Effect Revisited: Evidence for Tax Collection Efficiency in Brazilian Municipalities. *Estudos Econômicos,* 41(2): 239–267. http://dx.doi.org/10.1590/S0101-41612011000200002

Medina, A.F. (2012). Medición de la Capacidad Fiscal y Esfuerzo Fiscal de las Comunidades Autonómicas Españolas 2002–2009. Final masters project. University of Valladolid. https://uvadoc.uva.es/bitstream/10324/1742/1/TFM-E-17.pdf

Mikesell, J. (2007). Changing State Fiscal Capacity and Tax Effort in an Era of Devolving Government, 1981–2003. *Publius: The Journal of Federalism*, 37(4), 532–550. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1455597

Ministerio de Hacienda y Función Pública (2002–2016). *Tributación Autonómica. Medidas.* Madrid. http://www.minhafp.gob.es/en-GB/Areas%20Tematicas/Financiacion%20Autonomica/Paginas/libro%20electronico%20tributacion.aspx.

Mkandawire, T. (2010). On Tax Efforts and Colonial Heritage in Africa. *Journal of Development Studies* 46(10): 1647–1669. https://www.tandfonline.com/doi/abs/10.1080/00220388.2010.500660

Pessino, C., and R. Fenochietto (2010). Determining Countries' Tax Effort. *Hacienda Pública Española/Revista de Economía Pública*, 195(4): 65–87. https://www.researchgate.net/publication/227439705_Determining_Countries'_Tax_Effort

Rajaraman, I., and R. Goyal (2012). Tax Effort of Indian States, 2002–2007. In Srivastava, D.K., and Sanka, U. (eds.). *Development and Public Finance, Essays in Honour of Raja J. Chelliah*. SAGE Publications.

Ramírez, R., and A. Erquizio (2011). Capacidad y Esfuerzo Fiscal en las Entidades Federativas en México: Medición y Determinantes. *Paradigma Económico*, 3(1): 37–70. https://www.researchgate.net/publication/331385157_Capacidad_y_esfuerzo_fiscal_en_las_entidades_federativas_de_Mexico

Rao, H. (1993). Taxable Capacity Tax-efforts and Forecasts of Tax-yield of Indian States. *ISEC*. http://203.200.22.249:8080/jspui/bitstream/123456789/9760/1/Taxable_capacity_for_tax_efforts_and_forecast-Content.pdf

Secretaria General de Financiación Autonómica y Local (2002–2016). *Liquidación de los Presupuestos de las Comunidades Autónomas*. Ministerio de Hacienda y Función Pública.

Thirtle, C., B. Shankar, P. Chitkara, S. Chatterjee, and M.S. Mohanty (2000). Size Does Matter: Technical and Scale Efficiency in Indian State Tax Jurisdiction. *Review of Development Economics*, 4(3): 340–352. https://doi.org/10.1111/1467-9361.00099

Vallés-Giménez, J. and A. Zárate-Marco (2017). The Tax Effort of Local Governments and its Determinants: The Spanish Case. *Annals of Economics and Finance*, 18(2): 323–348. http://aeconf.com/Articles/Nov2017/aef180205.pdf

Wooldridge, J. (2013). *Introductory Econometrics. A Modern Approach*. South-Western, Cengage Learning. USA.

Wu, D.-M. (1974). Alternative Tests of Independence between Stochastic Regressors and Disturbances: Finite Sample Results. *Econometrica*, 42: 529–546. https://www.jstor.org/stable/1911789?seq=1#metadata_info_tab_contents

Yilmaz, Y., S. Hoo, M. Nagowski, K. Rueben, and R. Tannenwald (2006). *Measuring Fiscal Disparities Across the U.S. States: A Representative Revenue System/ Representative Expenditure System Approach, Fiscal Year 2002*, Washington, DC: Urban–Brookings Tax Policy Center and New England Public Policy Center. http://ssrn.com/abstract=1957339 or http://dx.doi.org/10.2139/ssrn.1957339

Zabalza, A. and M.A. Lasheras (2000). Índices Territoriales de Esfuerzo Fiscal en el IRPF. *Revista de Economía Aplicada*, 22 (vol.VIII): 95–129.

APPENDIX

Table A.1: Definition of the variables and their sources

	Description of the variable	Source of the information	Mean	Std. Dev.	Min	Max
DEATHS	Number of deaths in the region	National Statistics Institute (INE)	23,512.13	18,779.18	2,615.90	67,675.72
12w65	Total salaries from €12,000 to €65,000	By the authors, based on Job Market and Pensions in the Tributary Sources, Spanish Tax Agency (AEAT)	1.41E+10	1.29E+10	1.48E+09	4.80E+10
HOUSCON	Final household consumption of each region (in €)	Regional database of the Spanish economy BD.MORES (Ministry of Finance and Civil Service)	3.72E+10	3.43E+10	3.68E+09	1.16E+11
RICH	Total salaries over €150,000	By the authors, based on Job Market and Pensions in the Tributary Sources, Spanish Tax Agency (AEAT)	2.70E+09	4.11E+09	1.26E+08	1.76E+10
STOCKP	Stock of private capital	BD.MORES	1.82E+08	1.66E+08	1.66E+07	6.01E+08
HOUSEP	The average regional housing price per square metre	INE	1,664,413.00	489,678.30	739,037.40	3,271,030.00
IP0911	=1 in 2009-2011, when IP is not collected; = 0 otherwise		0.27	0.45	0.00	1.00
GAMBLINGEXP	Regional expenditure on gambling	INE	1.93E+09	1.68E+09	1.95E+08	5.53E+09
MORTGAGES	Number of mortgages issued	INE	74,882.32	84,220.36	1,811.00	404,875.00
FINACT	Financial assets	BD.MORES	1,626,469	2,189,477	75,176	1.51e+07
TAX	Tax revenue of the region	General Secretary of Autonomic and Local Financing (Secretaria General de Financiación Autonómica y Local - SGFAL).	5.28E+09	5.19E+09	2.77E+08	2.06E+10
INCOME	Gross Domestic Product	BD.MORES	5,138,756	4,884,390	536,084	1.70E+07
POP	Number of inhabitants	INE	2,768,462	2,388,630	279,359	8,299,100
DPROV	=1 if the region comprises a single province: = 0 otherwise		0.33	0.47	0.00	1.00
CAN	=1 if Canary Islands, =0 otherwise		0.07	0.25	0.00	1.00
TEND	=1 in 2002 =2 in 2003		6	3.17	1.00	11

	… =11 in 2012					
CRISIS	= 1 in 2010-12, =0 otherwise	INE	0.27	0.45	0.00	1.00
POLITCOLOUR	=1 if the region is governed by a left-wing party, =0 otherwise	Ministry of the Interior	0.39	0.49	0.00	1.00
DSINT	=1 if the central and regional governments are of the same party, =0 otherwise		0.48	0.50	0.00	1.00
ACTIVISM1	= Order-α tax effort for the regions that Order-α places beyond the frontier (Catalonia, Madrid, Andalusia, and Valencia) = 0 otherwise	By the authors	0.40	0.73	0.00	3.39
ACTIVISM2	Per capita revenue from own taxes	SGFAL	19,330.78	32,795.91	0.00	160,925.70
PATREV	(Income from assets + income from disposal of real investments)*1000/population	SGFAL	0.02	0.01	0.00	0.09
TRANSFREV	Income from current transfers and capital transfers which the Autonomous Regions receive from the State or from higher levels of government, such as the EU (Chapters 4 and 7 of the revenue budget)	SGFAL	1,463.28	835.79	-804.15	3,244.38
NFEXP	Per capita current spending represents the resources consumed over the year and necessary for the government's activity. More specifically, it covers the costs of personnel, running costs for services, and any current transfers (Chapters 1, 2 and 4 of the expenditure budget).	SGFAL	2,831.23	447.81	1,537.60	4,401.69
FEXP	Per capita current financial spending includes the expenditure needed to pay interest on government debt (Chapter 3 of the expenditure budget) and repay the principal (Chapter 9).	SGFAL	133,601	147.3708	0	1,412.031
DENSITY	Population /surface area	INE	154.64	178.04	22.34	798.02
POPGROWTH	(Final population - initial population)/initial population	INE	10.84	10.73	-6.48	33.17
QMANAG	Non-financial current income received/budgeted	SGFAL	101.43	12.05	80.33	148.57
GDPGROWTH	(Regional GDPt-Regional GDPt-1)/Regional GDPt-1	INE	1.12	3.78	-6.02	6.05

Table A.2: Correlation matrix of variables

	INCOME	POP	IP0911	STOCKP	GAMBLINGEXP	TEND	DENSITY	POPGROWTH	TRANSFREV	QMANAG	PATREV	ACTIVISM1	ACTIVISM2	NFEXP	GDPGROWTH	FEXP
INCOME	1,000															
POP	0,950	1,000														
IP0911	0,000	0,028	1,000													
STOCKP	0,993	0,956	0,047	1,000												
GAMBLINGEXP	0,954	0,934	-0,057	0,945	1,000											
TEND	0,010	0,045	0,581	0,078	-0,107	1,000										
DENSITY	0,620	0,456	0,024	0,594	0,564	0,039	1,000									
POPGROWTH	0,182	0,113	-0,421	0,157	0,284	-0,502	0,277	1,000								
TRANSFREV	-0,327	-0,221	0,015	-0,328	-0,288	-0,269	-0,597	-0,209	1,000							
QMANAG	-0,027	-0,058	-0,303	-0,061	0,014	-0,613	-0,013	0,287	0,043	1,000						
PATREV	0,191	0,153	-0,263	0,162	0,149	-0,164	0,145	-0,090	0,238	0,013	1,000					
ACTIVISM1	0,888	0,864	0,013	0,894	0,828	0,080	0,597	0,152	-0,427	-0,042	0,146	1,000				
ACTIVISM2	-0,276	-0,272	0,062	-0,267	-0,258	0,139	0,050	0,093	0,067	-0,105	-0,143	-0,293	1,000			
NFEXP	-0,082	-0,065	0,273	-0,060	-0,158	0,274	-0,406	-0,332	0,636	-0,305	0,187	-0,166	-0,004	1,000		
GDPGROWTH	-0,001	-0,046	-0,652	-0,063	0,080	-0,827	-0,031	0,513	0,267	0,444	0,204	-0,080	-0,088	-0,191	1,000	
FEXP	0,286	0,302	0,067	0,311	0,150	0,409	0,094	-0,318	-0,301	-0,186	0,056	0,371	0,009	-0,038	-0,369	1,000

Table A.3: General and specific indexes of tax capacity
(Regressions using Driscoll-Kraay standard errors for the 2002-2012 period)

	Individual estimates (by taxes)							Aggregate estimate
	IP	**ISD**	**ITPAJD**	**IRPF (income tax)**	**VAT**	**IICCEE**	**TJ**	**Total tax revenue**
Estimates of tax revenue using specific indicators of tax capacity								
DEATHS		5920.36 9.65						968.2755 0.02
12w60				0.00014 10.14				-.1202643 -0.91
HOMEXP					0.0368 14.43	0.02056 25.52		.0765143 2.04
RICH	2.52e-06 2.00							.0549367 0.92
STOCKP	0.3830 2.74							29.51556 2.86
HOUSEP	82.57 8.69	152.3004 16.28	337.6023 4.19					-178.8065 -0.61
ANUAL0911	-9.77 e+07 -7.97							-1.00e+09 -2.93
PLAYEXP							0.05295 7.63	-1.023991 -2.92
MORTGAGE			6976.109 6.21					1393.307 0.47
FINACT			0.08607 2.94					.1653496 1.21
CONS	-1.09+08 -3.20	-2.46e+08 -2.36	-5.65e+08 -3.63	-2.81e+08 -11.91	1.89e+07 1.34	4.99e+07 2.93	-3494809 -0.74	5.31e+08 0.85
R^2	0.6723	0.7607	0.8809	0.8253	0.8611	0.9128	0.6608	0.972
Estimates of tax revenue using specific indicators of tax capacity								
INCOME	.0034 4.11	.0042322 7.76	.0140434 4.85	.0409073 16.42	.0109878 5.23	.0028435 2.44	.0005747 0.753	0.07660 16.29
POPr	-54.21 -4.19	-47.69523 -4.61	-30.81779 -0.82	-275.2731 -3.64	251.4532 3.11	226.7382 7.54	29.94585 2.60	100.1408 0.59
CONS	1313950 -0.92	5607628 1.20	-5.19e+07 -1.81	-1.27e+08 -5.78	-1.68e+07 -0.71	4428533 0.49	3912222 1.50	-1.80+08 -4.16
R^2	0.60	0.8332	0.7123	0.8531	0.8597	0.9246	0.5659	0.955

Source: By the authors.

Table A.4: Explanatory variables used in the SFA literature

Variable	Papers using each variable
Tax frontier	
INCOME	Jha et al. (1999) Ramírez and Erquizio (2011) Pessino and Fenochietto (2010) Cyan et al. (2014) Garg et al. (2017)
POP	Ramírez and Erquizio (2011)
STOCKP	None
GAMBLINGEXP	None
CAN	None
DPROV	None
IP0911	None
TEND	None
Unexploited tax potential	
POLITCOLOUR	None
DSINT	Garg et al. (2017)
ACTIVISM1	None
ACTIVISM2	None
TRANSFREV	Jha et al. (1999) Garg et al. (2017)
PATREV	None
NFEXP	Garg et al. (2014)
FEXP	Ramírez and Erquizio (2011) Cyan et al. (2014) Garg et a.l (2017)
DENSITY	None
POPGROWTH	Cyan et al. (2014)
QMANAG	Pessino and Fenochietto (2010) use inefficiencies in collection with a corruption perception index. Cyan et al. (2014) use an index of assessment of corruption in the political system. Garg et a.l (2017) use crime rate, stolen property recovered by police, pendency rate of crimes and police personal /total population.
CRISIS	None
GDPGROWTH	None

Source: By the authors.

Economics
The Open-Access, Open-Assessment E-Journal

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

http://dx.doi.org/10.5018/economics-ejournal.ja.2019-31

The Editor