First, the authors would like to thank the referee for their highly detailed review of the paper. His/her useful comments and suggestions have guided us in making several changes we list below, which will improve and clarify the paper considerably. In this document we will also try to answer the referee's questions, and the matters which we think he/she may not have full understood, probably because they are not clear enough in the paper.

As we have been instructed not to upload the new version of the paper to the platform but only the reply to the referee, we will try to ensure this reply covers the changes to be made to the paper as thoroughly as possible.

Major conceptual issues:

1. The referee considers it necessary to define precisely the concepts *tax effort, tax capacity, tax collecting efficiency, tax potential* and *tax behaviour*. Although all these concepts are widely used and known in the field of fiscal federalism and public economics, it would probably be desirable, as the referee suggests, to provide enough explanation of them to clarify the paper. Given that only the concepts *tax effort* and *tax capacity* are specifically defined (in the first paragraph of section 2), because they are estimated in the econometric model, the other concepts could be clarified as indicated below.

Tax potential is the same as *tax capacity*, i.e., the maximum tax revenue a jurisdiction can obtain given its economic, social, institutional, and demographic characteristics. This is indicated in the last paragraph of page 2:

> "In the fourth section, we propose explanatory hypotheses for the tax potential (or tax capacity) and tax effort of those regions."

When we talk about *tax behaviour*, we refer to how the regions use their tax autonomy - i.e., their ability to establish their own taxes and to specify certain elements (tax rates, tax credits, allowances) of the taxes assigned to them. Page 7 would read:

> "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 taxes they are assigned (tax rates, tax credits, allowances), so the regional tax scenario vary widely."

Tax collecting efficiency is efficiency in the collection of the taxes which a jurisdiction has established, so page 5 would read:

"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 (James Alm and Denvil Duncan, 2014)."

Meanwhile, to make the presentation of these concepts less confusing, we have revised and reorganised section 2, explaining first the methods used in the literature to determine tax capacity, and then focusing on reviewing the papers which have calculated tax effort indices based on those methods. We think this will make the section clearer and easier to read.

2. The referee asks "is it worth considering/contrasting efficient tax collection and tax capacity", and whether "they are the same". They also say "it is stated (in our paper) that tax capacity depends on tax rate, tax management and inspection" and that "it would be helpful to define what is the focus of the paper".

To answer the referee's question, efficient tax collection is not the same as tax capacity, although we think the explanation in point 1 above has already clarified this subject.

Moreover, we do not analyse efficient tax collection. This analysis is done, for example, in Esteller (2005). What we do estimate is tax effort and its determinants, as indicated in the second paragraph of the introduction, although to avoid any ambiguity, we could write the sentence as follows:

> "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".

Esteller-Moré, A. (2005): "Is There a Connection Between the Tax Administration and the Political Power?, *International Tax and Public Finance*, 12: 639–63.

Although the text also justifies the interest or need to analyse this matter, in the context of regional funding in a federal country, we can strengthen the justification of the study by reformulating the first paragraph of the introduction thus:

> "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 happened in Spain

during the first two decades of the regional decentralisation model. On the other, the main subject of study in the literature associated with equalisation transfers is usually 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), increasingly large 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, insofar as 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."

Contrary to what the referee says, in this work we never state that tax capacity depends on tax rate, tax management and inspection, but rather the reverse. The text makes clear that tax capacity does not depend on the action of governments, while the numerator of tax effort (i.e., the tax capacity exercised, or in other words, tax revenue) does. Specifically, at the end of the first paragraph of section 2, we say:

"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 (tax capacity) is independent of the action of the government** (Jorge Martínez-Vázquez and Jameson Boex, 1997), and as this variable is unobservable, this figure is difficult to quantify".

Also on page 10, when we specify the factors determining tax potential or tax capacity, we remark:

"To choose the inputs or explanatory variables of the tax potential, we considered the available empirical evidence on sub-central tax behaviour, and performed a series of estimates to select the best indicators of regional tax potential, bearing in mind their explanatory capacity. **We also took into account that tax capacity is independent of government decisions or actions, which excludes the consideration of variables such as tax rate**.…"

3. The referee asks about the theoretical considerations behind the tax frontier. The theoretical model underlying the econometric estimation of tax effort has not really changed since the pioneering works of Lotz and Morss (1970). These authors, while favouring simplicity and minimising the information required, try to overcome the limitations of the pioneering alternatives (Frank, 1959 and ACIR, 1962)[1], estimating tax capacity through regression mechanisms, which take into account the socioeconomic characteristics of the jurisdictions, and then obtaining the tax effort by comparing real tax revenue with the estimated tax capacity. The first econometric papers estimated the tax capacity by OLS, and as can be seen on pages 4 and 5 of the text, there is enough international literature to validate this methodology, although it is only beginning to be used in the field of sub-central governments. However, there is a basic criticism of approximation using OLS regression, as pointed out by Rao (1993): it considers the random component of the residual as tax effort. This is the drawback to be overcome with SFA, by breaking down the residual into two components, u and v. The error term, v, represents the usual statistical noise, i.e., everything beyond the control of the region. The second error term, u, represents the error in obtaining the maximum amount of revenue for

---

[1] An exhaustive analysis of the drawbacks of the different alternatives can be found in Mikesell (2007) and Costa (2008).

given inputs or tax bases, and would be the function of variables z_{it}, which may vary over time and would include observed heterogeneity.

The theoretical framework has therefore been the same since the 60s, and what has changed is the quantitative approximation. At first this was done with average OLS estimates and their different variants, but since the 90s the SFA approach has gained ground, as it is considered to better approximate the tax capacity concept. Our paper also refers to the international literature which made the first efforts to implement the use of frontier techniques in the measurement of tax effort and tax capacity.

Given that the theoretical model underlying our paper is therefore sufficiently well-known and accepted by the literature on fiscal federalism and public economics, and our paper is basically an empirical work adapting these works to the sub-central context, we feel there is no need to go into greater depth on the theoretical explanations underlying the SFA. In any case, the text (first paragraph of page 4) explains the idea underlying this methodology:

> " SFA… provides 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 SFA 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 ….

and these questions are again discussed in section 4.

However, we can introduce theoretical references on the measurement of tax capacity and tax effort so that readers who are not specialists in these matters can further explore them. Specifically, we could add a footnote to the first paragraph of page 2, as we indicate below:

> "…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)[2]"

Advisory Commission on Intergovernmental Relations-ACIR (1988): *State Fiscal Capacity and Effort*. Information Report M-170, Washington, DC: ACIR.

Bird, R., and E. Slack (1990): "Equalization: the representative tax system revisited" *Canadian Tax Journal* 38, 913-27

Costa, A. (2008): *Un análisis de las disparidades fiscales generadas por la actividad turística: Aplicación al caso de los municipios españoles*, Doctoral Thesis. Palma.

Cyan, M., Martínez-Vazquez, J. and Vulovic, V.: (2014): "New approaches to measuring tax effort", in R. Bird and J. Martinez-Vazquez: *Taxation and Development: The Weakest Link?; Essays in Honor of Roy Bahl*, Edward Elgar: 27-68.

Dahlby B. and S. Wilson (1994): "Fiscal Capacity, Tax Effort, and Optimal Equalization Grants" *Canadian Journal of Economics* 27: 657–672.

Frank, H. J. (1959): "Measuring state tax burdens", *National Tax Journal,* 12, 179-185.

Gold, S. (1986): "Measuring fiscal effort and fiscal capacity: sorting out some of the controversies". In *Measuring Fiscal Capacity* ed. H.C. Reeves (Boston: Oelgeschlager, Gunn and Hain)

Kincaid, J. (1989): "Fiscal capacity and tax effort of the American states: trends and issues", *Public Budgeting and Finance* 9, 4-26.

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.

Rao, H. (1993). *Taxable Capacity Tax-efforts and forecasts of tax-yield of Indian States*. ISEC. Available at: http://203.200.22.249:8080/jspui/bitstream/2014/9760/1/Taxable_capacity_for_tax_efforts_and_forecast-Content.pdf. Accessed November 2013.

4. The referee says "it is worth differentiating between corporate and private taxes". However, as can be seen in Table 1, the taxes assigned to the Spanish regions are taxes on individuals, not on companies. Corporate Tax is not assigned to the regions. And although businesses are liable for Value Added Tax, as a tax on the value added in the production process, it is ultimately borne by the end consumer. Excise taxes (tobacco, alcohol, fuel, certain forms of transport) are also borne by the consumer. In the case of fees, there is no disaggregated information available to separate them as the referee

---

[2] The theoretical aspects of the tax capacity and tax effort concepts can be reviewed in Frank (1959), ACIR (1962), Gold (1986), Kincaid (1989), Bird and Slack (1990), Dahlby and Wilson (1994), and Cyan, Martínez-Vázquez and Vulovic (2014).

proposes, and most of them are essentially payments by the taxpayers for the public services provided by the regional governments. For this reason, we think that at the regional level and for the Spanish case, the differentiation the referee proposes is not appropriate.

Major empirical issues

1. The determinants of the frontier and tax effort, which the referee asks about, are detailed on pages 10-13 under the headings "Factors determining tax potential" and "Explanatory hypotheses of tax effort", respectively.

Specifically, to estimate tax capacity, we use population, income, proxies of certain tax bases (those which a regression analysis considers to be best), and several factors which are intended to reflect the institutional context of Spain. All of this is summarised in the last paragraph on page 11, as follows:

> "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 on the frontier. 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 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. 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."

To explain tax effort, we have considered political, budget, demographic, management and collection efficiency, and economic cycle variables. All of these can be seen in detail on pages 11 - 13.

Also, in response to the referee's question, we would say that many of them are variables commonly used in the literature, and which we reference in some way in the text, as can be seen in the second paragraph on page 10:

"…to choose the inputs or explanatory variables of the tax potential, we considered the available empirical evidence on sub-central tax behaviour",

and in the last paragraph of page 14:

"The results are consistent with what we would expect from a theoretical point of view, and with the available but scanty empirical evidence (Pessino and Fenochietto, 2010, and Garg et al, 2017)."

However, it must be taken into account that, as the referee points out, we have also used a regression analysis to select the variables which best measure the tax capacity of the Spanish regions, and other variables which have already been used for countries, but whose validity must be checked in the sub-central context, given the relative novelty of this literature at the sub-central level. We also add new explanatory factors which are not used in the literature as possible causes of tax potential, some of which comes from a more general fiscal federalism literature on sub-central tax behaviour, which was also used to improve the model explaining tax effort.

Following the suggestion of the referee, we could introduce more references to the papers which have used these variables (Cyan et al, 2014; Jha et al, 1999; Ramírez and Erquicio, 2011) in the explanation of the hypothesis or the results. In any case, we have noted here, as we do in the text, that there are still very few papers using SFA.

2. As the referee indicates, regional governments can be inefficient in their tax management, monitoring and inspection. However, there is no regional-level information for Spain which would let us measure the efficiency of these tasks in a disaggregated way. For this reason, we have opted to approximate this inefficiency more generally with the variable QMANAGE, which we have constructed as the quotient between the non-financial current revenues the region really receives and the revenues it budgets for. However, this general approximation makes it impossible to specify where the inefficiency

comes from, although it does let us identify poor management skills among the economic management personnel of regional governments as a possible explanatory factor.

3. We agree with the referee's statement that "if XTSCC is correct, SFA is wrong", and although we have tried to express this in the text (first paragraph on page 15), it is true that this may not be very clear. We also thank to the referee for spotting the erratum related to the SFA and XTSCC coefficients. As he/she points out, their values are not similar, but rather their signs and significance, which would reinforce the validity of the sign theoretically attributed to the coefficient of the chosen variables, an aspect which cannot be ratified by non-parametric techniques. To correct both issues, we have revised the last paragraph on page 15, which would now read:

> "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 (XTSCC), adapted by Daniel Hoechle (2007), which can be seen in the third column of Table 2, under the heading XTSCC. 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."

The referee also asks why we use feasible GLS, order-alpha and order-m frontier approaches. In fact, we do not use feasible GLS, which as the referee points out, do not measure inefficiency. We do mention this method on page 11 when we discuss which would be the best indicators of regional tax potential, but we do not use it. In order to there is no doubt, we will write the following sentence in conditional

> "Other methods which **would** let us simultaneously eliminate the problems mentioned are Parks-Kmenta feasible generalized least squares (FGLS), and Beck and Katz´s panel corrected standard errors (PCSE), although the former cannot be used when T<N, as in our case, and the latter perform better with smaller samples."

We actually analyse the tax effort of the Spanish regions using SFA. We use the Driscoll-Kraay robust errors method (XTSCC), the order-alpha and order-m frontier approaches, and the Free Disposal Hull as complements, in order to measure the robustness of the tax effort results obtained with SFA[3]. In fact, this is indicated in the second

---

[3] As we explain above, we also use XTSCC to test the validity of the hypotheses explaining tax capacity.

paragraph on page 17, although to avoid confusion we will change the wording, which will now read:

> "To check the robustness of our 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 the section 2 (i.e., Order-m and Order-α partial frontier methods and the Free Disposal Hull). The 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."

4. It is true that, as the referee indicates, 15 regions may be too few to guarantee a consistent estimation of the proposal of Greene (2005), implemented in the *sfpanel* command (Belotti et al., 2013). For this reason we also use non-parametric estimations to test whether the econometric problems which might affect the SFA are relevant. However, the results show a fair amount of uniformity between the tax efforts calculated with the different frontier techniques used, in relation to the jurisdictions making the most tax effort (this can be seen in Table 3). There is slightly lower concordance for the regions with laxer tax behaviour, a result which we attribute, as the evaluator says, to the inability of non-parametric frontier models to incorporate heterogeneity and panel structure, especially in the case of single-province regions (Asturias, Balearic Islands, Cantabria, La Rioja).

Meanwhile, thanks to the comments of reviewer 3, we want to mention that there may be endogeneity problems in the model. To determine whether the endogeneity problems affect the variables indicated by the referee 3 (gross domestic product-INCOME, population, political variables, grants received, and non financial current spending-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 1 below. In all cases we rejected the endogeneity of the variables.

Table 1: Analysis of potential endogeneity (Wu-Hausman and Durbin)

Variable	Wu-Hausman F (1.1151)	Prob>F	Durbin $X_i^2(1)$	Prob>X_i^2
INCOME	0.0666	0.7966	0.0711	0.7897
POP	2.5042	0.1156	2.6233	0.1053
dPOLITCOLOUR	0.0091	0.9239	0.0099	0.9204
dSINT	0.0755	0.7838	0.0826	0.7738
TRANSFREV	1.9394	0.1658	1.8146	0.178
NFEXP	0.0.5799	0.4475	0.6020	0.4378

However, taking advantage of the fact that Karakaplan and Kutlu (2017, 2018) recently developed an estimation procedure for taking endogeneity into account in frontier models, together with a new test for detecting endogeneity in stochastic frontiers, we have implemented this procedure with the command *xtsfkk* in Stata. This command can handle endogenous variables in the frontier. Although this estimation (which can be seen in Table 2) shows that the variable INCOME can present some endogeneity (eta1=-2.102*), as the technique proposed by Karakaplan and Kutlu (2017) is robust against this problem, it generates unbiased results. Table 2 shows that the same variables are found to be significant as in our initial model, and with the same sign, except for population, which is now not significant.

Table 2: Results of the estimates of endogenous panel stochastic frontier models 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)
IP0911	-0.247*** (0.031)	-0.244*** (0.031)
CAN	-0.372*** (0.058)	-0.368*** (0.054)
DPROV	-0.014 (0.039)	-0.024 (0.037)
STOCKP	0.178 (0.241)	0.029 (0.239)
GAMBLINGEXP	0.034 (0.090)	0.046 (0.087)
TEND	0.050*** (0.008)	0.054*** (0.008)
Dep.var: $\ln(\sigma^2\_u)$		
Constant	-6.011*** (1.441)	-6.515*** (1.887)
Dep.var: $\ln(\sigma^2\_v)$		
Constant	-3.897*** (0.117)	
Dep.var: $\ln(\sigma^2\_w)$		
Constant		-3.951*** (0.117)
eta1 (INCOME)		-2.102* (0.830)
eta2 (POP)		4.824 (3.516)
eta Endogeneity Test		X^2=9.5 p=0.009
Observations	165	165
Log Likelihood	84.7	1.257.54
Mean Tech Efficiency	0.6444	0.6942
Median Tech Efficiency	0.6399	0.6860

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

As the proposal of Karakaplan and Kutlu (2017) does not let us simultaneously estimate the tax effort and tax frontier equations, we have also tried instrumentalising the

variable INCOME, which is the one which can present endogeneity problems, based on the end consumption expenditure declared by households, maintaining the initial estimation approach implemented in the *sfpanel* command (Belotti et al., 2013). We must also point out that based on the comments and suggestions of referee 1, we have redefined the variable CRISIS in our estimation, now assigning the value 1 to the years 2010 - 2012, given that regional governments did not suffer from the decrease in resources until 2010, when transfers were negatively adjusted by the central government; and we have used 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 results obtained are very similar to those obtained initially, as can be seen in Table 3. The final version of the article (if accepted for publication) will clarify and include all these questions and considerations.

Table 3: Results of the estimates of tax potential with instrumental variables

	Coef	z	P>|z|
Tax frontier			
INCOME	.5913903	4.24	0.000
POP	.2685138	2.84	0.004
IP0911	-.1525747	-10.58	0.000
CAN	-.3218095	-5.45	0.000
DPROV	.0262852	1.09	0.274
STOCKP	.1839746	1.58	0.114
GAMBLINGEXP	-.0735474	-1.48	0.138
TEND	.0284289	7.76	0.000
CONS	7.321896	15.74	0.000
Fiscal gap			
DENSITY	.0011281	3.71	0.000
POPGROWTH	-.0037158	-1.15	0.251
QMANAG	.0023321	1.37	0.172
TRANSFREV	.0005645	6.21	0.000
PATREV	-1.175633	-0.49	0.622
ACTIVISM1	-.2225374	-2.51	0.012
ACTIVISM2	-2.27e-06	-2.37	0.018
dPOLITCOLOUR	.1003027	1.81	0.070
dSINT	.0799681	1.82	0.069
NFEXP	-.0005173	-4.58	0.000
RATE(INCOME)	-.0238122	-2.57	0.010
CRISIS1012	.3785368	4.52	0.000
FEXP	-.0006928	-1.80	0.072
CONS	.2631531	0.81	0.419
θ	-.0696303	-5.11	0.000
σ_u^2	.1376102	7.73	0.000
σ_v^2	.0230962	3.35	0.001
λ (Ho: $\gamma = \sigma_u^2/\sigma_v^2 = 0$)	5.958146	284.07	0.000

Durbin, J. (1954): "Errors in variables". *Review of the International Statistical Institute* 22: 23-32.

Hausman, J. A. (1978): " Specification tests in econometrics". Econometrica 46: 1251-1271

Karakaplan, Mustafa 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

Karakaplan, Mustafa U. and Kutlu, Levent (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, Mustafa U. (2017) "Fitting Endogenous Stochastic Frontier Models in Stata", *The Stata Journal* 17(1): 39-55. https://www.stata-journal.com/article.html?article=st0466

Wu, D.-M. (1974): "Alternative tests of independence between stochastic regressors and disturbances: Finite sample results". *Econometrica* 42: 529-546

Minor empirical issues

1. Zones in figure 1 are not arbitrary. In fact, the zones are bounded by the dotted lines which are the averages of the concepts shown on the axes. The vertical line is the average tax effort calculated with SFA, and the horizontal line shows average per capita tax revenue. The referee's comment is reasonable because this was not indicated in the paper, so we have introduced this clarification in the last paragraph on page 18:

> "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 low per capita tax revenue (**below average**) …"

At the same time, we have corrected an error we have found in the translation of the Y axis in Figure 1. Where it says "per capita income" this should read "per capita tax revenue".

2. We have revised the wording of the first paragraph on page 3 to avoid confusion, and following the referee's suggestion, we have specified that tax effort is in fact a ratio. It now reads:

"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). 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 (Jorge Martínez-Vázquez and Jameson 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."

3. Medina (2012) also calculates the tax effort of the Spanish regions, but approximating tax capacity with little precision and in a too simply way. For Medina (2012), the socio-economic structure of the regions (the participation rate of industry and the service sector, current transfers and the tax revenue of local governments) determines the maximum tax revenue frontier, while we use a more sophisticated approximation, using different specific indicators for the tax bases of the regions and other general tax capacity indicators, and taking into account a series of variables which capture the institutional context of the regions (we identify the single-province regions; the Canary Islands region, which has unique features in its funding system; and other relevant aspects). Also, Medina does not analyse the determinants of tax effort, and uses an econometric technique which does not include heterogeneity and panel structure.

Medina (2012) also excludes the Canary Islands, when this region, although with unique aspects, is within the financing system of the "common regime" and therefore we think should be included in the analysis. Our work captures the uniqueness of the Canary Islands with a dummy in the tax capacity equation [1].

The case of the Basque Country and Navarre is different, with a totally distinct funding system. Their funding system is characterised by their governments' power to establish, maintain and regulate their tax regime. This means that tax demands, management, settlement, collection and inspection of most state taxes (currently all, except those related to imports) correspond to each of the three territories (provinces) of the Basque Country and to the Foral Community of Navarre. This is the main reason why we say these systems to not be comparable.

4. We would like to thank the referee for warning us that the estimates in Table A.3 of the Appendix are in levels and the models in Table 2 in logarithms. All the estimates are in logarithms, but this was not specified in the Appendix. The final version (if the discussion paper is accepted as journal article) will make this question clear, as suggested by the referee.

5. We explained above, in point 3 of the *Major empirical issues*, that the XTSCC estimation of Table 3 is there to test for the robustness of the hypotheses explaining tax capacity. Meanwhile, we chose to estimate tax capacity with Driscoll and Kraay standard errors because it provides robust estimates in the presence of cross-sectional dependence (Pesaran, 2004), so ignoring this problem in the estimation of panel models can lead to severely biased statistical results (Hoechle, 2007):

Driscoll, J., and A. C. Kraay (1998): "Consistent covariance matrix estimation with spatially dependent data", *Review of Economics and Statistics* 80: 549-560.

Pesaran, M. H. (2004): "General diagnostic tests for cross section dependence in panels". *Cambridge Working Papers in Economics* No. 0435. University of Cambridge, Faculty of Economics.

Hoechle, D. (2007): "Robust standard errors for panel regressions with cross-sectional dependence", *Stata Journal*, 7: 281-312.

6. The expression "homoscedastic sample" is a translation error. We meant homoscedastic residuals, so that the first paragraph on page 11 should read:

"…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 (XTSCC estimates)."

7. We explained the variable QMANAGE in point 2 of the *Major empirical issues*.