First, the authors would like to thank the referee for their highly detailed review of the paper. Their 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 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.

On issues of Styles:

1.- The referee considers that, although the variables are defined, non-financial current expenditure (NFEXP), financial expenditure (FEXP), revenue from transfers (TRANSFREV), and non-financial revenues need some further explanation.

- Non-financial current expenditure (NFEXP) 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).
- Financial expenditure (FEXP) includes the expenditure needed to pay interest on government debt (Chapter 3 of the expenditure budget) and repay the principal (Chapter 9).
- Revenue from transfers (TRANSFREV) is 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).
- 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)

The first three items (NFEXP, FEXP and TRANSFREV) can be clarified in the definitions of Table A1 of the Appendix, and the non-financial revenues can be explained in a footnote to the central paragraph on page 10, as follows:

"…Given that assigned taxes represent approximately 90% of the non-financial revenues of the regions[1], the explanatory variables of the assigned taxes will also provide a good explanation of the endogenous TAX or total tax revenue…".

2. 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)."

<u>On the Econometric Analysis:</u>

1.- Referee 3 requires more explanation of why GLS are used, and the order-m and order-alpha models.

In fact, we do not use feasible GLS. 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. To avoid any ambiguity, 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, as this corrects the problem of the OLS models, which include tax effort in the residual (Rao, 1993). This is explained in the first paragraph on page 15, although to make it clearer we will change the wording of that paragraph as follows:

"The results of estimating the model above, for which we took variables in logs and used the STATA statistical package, are shown in the second column of Table 2. 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

[1] 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).

> SFA is confirmed as a suitable 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 (2012), 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. …"

And given that, as we have just explained, tax capacity must be estimated through the consideration of a tax frontier, but the stochastic frontier analysis (SFA) may present some econometric problems (e.g., endogeneity), we have also used non-parametric frontier methods (order-alpha and order-m frontier approaches, and the Free Disposal Hull), which enable us to test the robustness of our SFA results, as well as capturing atypical regional behaviours. In fact, this is indicated in the second 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."

2.- We agree with the referee's statement that taxes from the central government can affect tax bases of subnational governments and their ability to collect subnational tax revenue. This is attested by the abundant literature on the subject (e.g., Dahlby and Wilson, 2003; Andersson et al, 2004). However, we feel that in the Spanish case, the vertical tax externalities are very limited, given that although 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[2]), 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 occupied the "fiscal space" vacated by the central government and scarcely exercised their tax autonomy, make the existence of vertical tax externalities very unlikely for the period being studied. For all these reasons, we think this aspect does not need to be taken into account in our estimation. However, we thank the referee for their suggestion, and given its relevance, we will provide a justification in a footnote to the middle paragraph on page 10 for not including it in the analysis, as follows:

> "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. 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)[3]."

Andersson, L.; T. Aronsson and M. Wikström (2004): "Testing for Vertical Fiscal Externalities", *International tax and Public Finance*, 11 (3): 243-263.

Dahlby, B. and L. Wilson (2003): "Vertical fiscal externalities in a federation", *Journal of Public Economics* 87 (5-6): 917-930.

3.- The referee notes possible endogeneity problems in the model. To determine whether the endogeneity problems affect the variables indicated by the referee (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),

[2] Corporate tax is not shared.

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

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.

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 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.

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.