

Response to the four anonymous referees on “A Comparative View on the Tax Performance of Developing Countries: Regional Patterns, Non-tax Revenue and Governance”

Christian von Haldenwang / Maksym Ivanyna

Above all, we appreciate the thoughtful and constructive comments provided by four anonymous referees (and by Oliver Morrissey, to whom we will reply in a separate document). We feel encouraged by the positive feedback with regard to the relevance of the subject and the contribution our paper makes (or could make) to the debate. The suggested amendments and corrections will help us to improve the paper.

As a general observation, several comments refer to an expansion of the statistical analysis in order to include additional variables or to provide support for findings from the qualitative analysis. We will try to integrate these observations in a revised version of the paper, providing more statistical analysis. This should help us to smoothen the step from statistical to qualitative analysis and to make the paper more reader-friendly.

Our responses (*in italics*) are inserted directly in the reports.

Referee Report 1

Minor Comments:

- It is not quite clear for the reader the tax data source being used to produce the trend line in Figure 1. Are all sources described in Table 1 used all together or one of them is used in particular?

Data sources are ranked according to priority in Table 1, as mentioned in the main text of the paper (page 6, second paragraph). We will add this information to Table 1.

- The paper describes a number of (supposedly) control variables that explain tax burden and it provides results but it does not report regressions (section “government levels”). This is quite wrong.

True, we do not show regressions in this section. Instead, we compare low tax performing countries with the rest of the world based on a given indicator (governance, non-tax revenue, etc.). Specifically, we indicate which low tax performers score lower than the world's average, and which countries score higher. We will add more graphical material and report regression results in the revised version.

Major comments:

- The subject being addressed is an interesting matter. Nevertheless, a simple regression between tax burden and GDP (whatever the specification is) is not enough to provide strong evidence on it.

We read this comment as a suggestion to introduce additional (control) variables in the original specification, as most other studies of tax effort or tax performance do. This will be done in the revised version of the paper. Screening the literature we have found 46 different indicators used to “explain” tax effort. These indicators account for economic, demographic and political factors. Most models include GDP (different specifications), along with some other factors. In some cases GDP maintains significance, whereas in other cases it loses significance (especially in combination with agriculture and trade). Our decision to use a very basic specification at the start of our analysis is based on two arguments: (i) The available literature does not question the influence of development level on tax revenue (though it does question the adequacy of GDP per capita as a stand-alone explanatory variable for tax revenue). (ii) We do not want to “explain” tax effort. Instead, we want to identify those countries that, given their development level, have a low tax – GDP ratio. Having identified this group, we proceed to discuss other variables (non-tax revenue, governance, region) with regard to the group of “low tax performers”.

As a matter of fact, the evidence of association between GDP per capita and tax ratio is strong and remains robust even excluding outliers, checking for different functional forms, using different definitions of tax revenue, and checking for non-linear relationships between the two variables by including country income dummies. This information is provided in Tables I and Table II in the Appendix. In all specifications we get a strong association between tax ratio and level of income (development).

- An empirical exercise worth doing is to run a panel regression with country fixed effects and then (as a second step), regress fixed effects on all relevant control variables. The linear-non linear specification may be captured by running different versions of the panel, in which GDP may be logged, squared, etc.

Our data would allow to use this method, but our feeling is that this would run counter to the underlying research interest of this paper. We are not trying to identify a causal effect of GDP per capita on tax ratio. Instead, we want to provide policymakers with an alternative benchmark when assessing tax revenue, rather than just using tax ratio. For this purpose, (1) we need cross-country variation based on level effects (which would be wiped out by country fixed effects if we ran panel regression), (2) we want to make our analysis as tractable as possible while keeping it statistically rigorous, and (3) we want to be able to track changes in tax performance with time, which would also be robust to possible changes in the structural relationship between GDP per capita and tax ratio (i.e. different slope coefficients at different time periods).

We are grateful for the suggestion, however, and plan to use the method (panel analysis with fixed effects) when treating other research questions with a similar data set.

Conclusions:

The paper can be only considered for publication after all changes mentioned above (some of them are indeed substantial changes), are included.

Referee Report 2

The paper is a timely effort to understand tax performance, given the renewed global attention to tax matters, not least through recent efforts by the G-20 to call on higher effectiveness for improved tax systems. It is well and thoroughly researched. However, it could be sharper with regards to the conclusions, in particular the way the country clusters are being established and the alternative ways in dealing with underperformance.

1) Assessing progress in tax performance

- The paper could differentiate better the different types of tax performance which can be established and measured: (i) tax revenue collection; (ii) efficiency and growth; (iii) equity. Most research has been researched in area (i) while more international effort is now dedicated to numbers (ii) and (iii), particular in the aftermath of the global crisis and uncertain macro environment. It would be useful to differentiate these goals more clearly when the authors claim that not sufficient analysis on tax matters is available.

This is a valid point. The tax effort / tax performance debate has focused almost exclusively on the revenue objective and much less on growth or distribution. We will discuss that in a revised version with references to Pessino / Fenochietto 2010; Bird / Martínez-Vázquez / Torgler 2004; Timmons 2010, and others.

- Regarding (ii), and in order to have a more differentiated discussion of tax performance, it would be useful to provide insights into different levels of performance in different taxes (VAT versus trade etc). Recent data from the IMF show that there is no substitution effect with aid resources in countries which have raised more taxes from trade.

We do not find evidence of a substitution effect with aid resources and tax revenue either, confirming several recent contributions to the debate on ODA and rent-seeking. With regard to the performance of different taxes, some studies address the issue, especially with regard to direct / indirect taxes and trade taxes – but with much smaller samples. We have decided not to pursue this line of research in this paper, principally because obtaining the necessary data and cleaning up the data set (for instance with respect to different accounting methods) while maintaining a broad (130+) set of countries would clearly exceed our possibilities. Digging more deeply into tax structures and the tax mix of countries is something we plan to do in the future (it is almost a logical next step), but probably with a considerably lower number of countries.

- The paper's critical view about benchmarking should be differentiated in that it should not be the only measure to assess tax performance. Related, it is worth noting that the G-20 intends to start out from benchmarking, but develop "a core set of indicators" to measure progress. The authors are right that simple tax-GDP ratios can provide a distorted picture (it requires similar levels of income and similar economic structure); and that a comparison by per capita GDP can provide further insights (the differences across countries can become larger).

Thank you, we will take a look at G20 and other initiatives in this matter. For instance, the IMF has recently presented its own benchmarking tool.

- The discussion about the measurement problems (page 3 forward) is very substantial. In this respect, the author's argument to measure marginal gains (footnote 8) is absolutely critical. This is one of the key challenges in measuring performance—and for the development community more broadly to have a credible policy dialogue on this matter. It would be highly innovative and a substantial contribution if such marginal gains could be assessed. But this would require an assessment of tax effort. The paper rises expectations that such methodology is being established and measurement being conducted (which it does not).

We do not measure tax effort in this paper. We will try to be more transparent concerning our objectives and messages in the revised version of the paper. However, the results presented in the paper underscore the referee's point in that they draw attention to the fact that an increase of, say, 4% in the tax ratio of a poor country means something different than the same increase in a rich country.

- Customs is part of modern revenue management. Collection targets can be improved by establishing cross-checks and integration among these different agencies. The World Bank is promoting this through integrated assessment tools like IAMTAX, which would allow benchmarking on tax administration. It would be useful to discuss such integrated approaches and its potential to improve revenue.

We will include references to tax administration reform initiatives in the concluding section of the paper.

- Regarding aid, it would be necessary to have a consistent line of argument throughout the paper. On page 14 it is argued that “low tax performers do not receive a great deal of foreign aid”. This is not entirely true in all countries. But on the other hand, there is displacement of domestic tax revenues in countries with weak institutions—hence the importance of the governance aspect.

We will clarify our message by pointing out that the statement on page 14 refers to a statistical relationship. This holds even if we control for GDP per capita (regressions will be included in the revised version of the paper). The argument about low governance and higher foreign aid is valid only within the low tax performers group, it is not about crowding out of tax effort by foreign aid (since all countries in this group have a low tax effort). The argument is used to point out that high-governance low-tax countries are in a position to choose to have low taxes, and do not substitute taxes with other kinds of revenues. We will reformulate this argument in the revised version to make it clearer.

- Regarding the analysis of taxation and governance, the paper could be more explicit that causation can run both ways. Not only is tax performance related to the governance levels; but strengthening domestic resource mobilization also encourages good governance.

Point is well taken. We will mention that.

- Levels of government. The paper usefully distinguishes between tax revenue at central and subnational government. The paper should make it more explicit that GFS not only does have coverage problems which are well documented in this review, but also do not discern

between levels of responsibility with regards to subnational governments, both on the expenditure as well as revenue side. For revenues, there are some cases where subnational governments are subject to earmarking of their own taxes (Colombia is a case in point: only about a third of the department's revenue is freely allocated-the rest is earmarked). These should be mentioned as a separate problem.

That is certainly true. There is a number of issues to be discussed with regard to fiscal relations between national and sub-national jurisdictions, including the distribution of competencies (creating taxes, setting tax rates, spending tax monies) and the problem of perverse incentives linked to transfer schemes. Our goal in this paper is limited to measuring general government tax revenue as precisely as possible, and we are documenting difficulties that arise while we are doing that.

2.) Conclusions of the Paper

- The conclusions of the paper need to be sharpened. The paper starts proposing three different country groups. It is not entirely clear how these groups are being formed. It is clear that several analysis are being done (GDP per capita; governance), but the paper does not explain well how these criteria finally drive the selection of the countries.

We will try to improve our explanations. Basically, the groups are identified by their respective rankings regarding non-tax revenue and governance levels in combination with low tax performance.

- The conclusions start out discussing the country typology but then they fall back into regional patterns—which was precisely considered by the authors not to be a useful unit of analysis in tax performance.

The discussion of regional patterns in section 3 of the paper (pages 12-14) should not be read as arguing against this analytical perspective in general. We highlight some of the methodological challenges (small sample sizes, outliers), but still feel it is possible to suggest some profiles and causality paths, a discussion we take up in the concluding section of the paper. We will clarify the messages in order to avoid misunderstandings.

- Regarding the country patterns the paper could offer more discussion and guidance for a possible way forward in dealing with different levels of performance. To this end, it would be good to have a consistent argument in the paper about the question “should one care about underperformance?”, when it mentions that countries have their own, legitimate right in choosing the tax system they want.

Our argument is that societies characterized by a “good governance” framework (especially with regard to “voice and accountability”) are probably in a position to choose the tax system they want. In countries with low levels of governance, we would not trust the common-interest-orientation of governments. If bad governance goes hand in hand with dependence on ODA or other kinds of financing from abroad, low tax performance becomes an issue to care about. Our goal is to identify these countries.

- In light of the paper's findings, should assistance on tax systems continue to be region-specific? What is then the role of region-specific organizations, for instance the African Tax

Administration Forum; *Centro Interamericano de Administraciones Tributarias*? The findings of the paper at least cast some doubt about continuing with region-specific assistance.

We will reformulate our messages to avoid such an impression. We argue for global and more sophisticated benchmarking and measuring tax performance. Regional organisations such as ATAF and CIAT can be useful to address underperformance – even more so as regional patterns seem to have relevance at least in some regions. We will refer to them in the conclusion.

- How would the author's findings compare to other possible groupings of countries? (i) Fragile states (there, a well functioning tax administration is critical to building up trust); (ii) small states (or islands) which can establish taxes at the border with more ease; (iii) "Successor states" following the break-up of sovereign states which could strive to be as investor-friendly as possible. In addition, it would also be useful to learn more about implications in dealing with natural resource-rich countries (page 16).

This is indeed a tricky question, because the first two country groups mentioned by the referee are precisely those where the data challenge is particularly severe. Most countries missing in our sample are either extremely fragile or small states. Successor states (especially ex-Yugoslavia, ex-Soviet Union) could be an interesting unit of analysis (with, for instance, the Baltic republics pursuing the investor-friendly tax policy envisioned by the referee), as would be, of course, the natural resource-rich countries. We consider this an interesting suggestion for a future deepening of our analysis.

- The paper could provide more guidance on how the international community (and emerging efforts such as the International Tax Dialogue) could deal with these problems, particularly on the three cases which were identified. The G-20 has proposed south-south cooperation, shared knowledge platforms, among others, to strengthening assistance for improved tax systems. Would these be appropriate approaches in light of the findings of the paper?

As mentioned above, the revised version of the paper will refer to the implications of our findings for donors and international organisations.

3. Structure of the paper

- The paper could be much improved with regards to the storyline. In its current style, it provides for quite a difficult reading. Often the reader is guided through several of the alternative measures which were performed. In certain sections, they communicate to the reader several analysis which were performed (for instance paragraph [i] on page 9; or second paragraph on page 11). This gives one the assurance that this is a well-researched document, but it does not help to get across the main messages. Such analysis should be moved either to the footnotes or an Annex.

Thank you, we will check if there are elements of the discussion that can be moved to the footnotes / appendix in order to make the paper more reader-friendly.

Referee Report 3

This paper combines quantitative and qualitative approaches to the comparative analysis of tax systems. It is argued that tax performance should not be assessed against some absolute values but as a function of tax ratio and development level.

I think that the paper is original in the context of the economic literature on tax performance. The paper tackles an important issue and adopts a methodological approach that allows to yield new insights into the impact of some economic characteristics on tax performance. Therefore, I think this paper will be of interest for the readers of the journal and can be published in Economics e-Journal.

Some minor comments:

- The author(s) show in Figure 1 the relationship between the tax ratio and the log GDP per capita for 177 countries. It should be convenient to clean the table or to present several ones, since it is difficult to understand what is going on with so many data.

Perhaps we can just omit labels in Figure 1 to show a general scatter-plot? Low-tax and high-tax performers are already listed in Table 3. On the other hand, readers might be interested in the position of individual countries, even though they are not always easy to identify.

- There are some papers that should be included in the revision of the literature. Firstly, the seminal paper by Richard Bird (1976): - Bird, R.M (1976): "Assessing tax performance in developing countries: a critical review of the literature". FinanzArchiv/Public Finance Analysis, 1976. It would be helpful to include some references to the literature related to tax effort in developing countries, which is very close to the subject of the paper, such as: - Bird, R. M, Martinez-Vazquez, J. and Torgler, B. (2006): "Societal institutions and tax effort in developing countries" in The Challenges of Tax Reform in a Global Economy. Springer.

We will incorporate more literature in the text, including both papers mentioned by the referee. Thank you.

Referee Report 4

This paper tackles the important issue of assessing the relative tax performance of developing countries. In general, I like the authors' idea to use econometric techniques to determine whether, conditional on the development status (as measured by GDP per capita), a country performs well in terms of raising tax revenues or not.

From my reading, the paper is motivated by the need of political actors and development agencies to understand whether a certain country underperforms in terms of raising tax revenues. The authors argue that often the relative tax performance of countries is assessed by comparing them to other economies in the same region or even worldwide whose development level not seldom differs substantially from the country under consideration. This then raises

doubts that they are a good comparison group to evaluate the country's relative tax performance.

The authors' aim is to improve this type of ad hoc comparison in previous work by running simple OLS regressions of the tax to GDP ratio on the country's development level as measured by GDP per capita. They then compare the predicted value of the tax to GDP ratio at a given development level to the actual tax to GDP ratio reported.

- While this is certainly a step forward compared to previous work, a country's tax performance is not only determined by its development level but may be affected by a host of other factors. If one really wants to come up with a list of "poorly performing" countries in terms of taxing powers conditional on the general situation in the country, one probably has to make sure that one also takes other important determinants of tax performance into account. It would for example be interesting to run several regressions which are augmented by additional determinants of the tax-to-GDP ratio, for example:
 - 1) the year under consideration (there may be a general time trend in tax performance)
 - 2) for the general structure of the economy in the country under consideration
 - 3) variables capturing the business cycle
 - 4) political variables and institutional setting

This observation is in line with comments made by referee 1. We will include more statistical analysis in the revised version.

With regard to the year under consideration, we run different regressions for different time periods precisely to be robust not only to the time trend but also to changes of the structural relationship between tax ratio and GDP per capita.

Considering the business cycle, two periods we analyze (2007, 2002) are booms for practically all countries in the sample. 1997-99 is more complicated, but actually many Asian countries (which were at crisis in 1997) have registered lower performance since that time. We will mention that aspect in the paper.

- This type of analysis would allow making statements on whether the country performs well in terms of raising tax revenues conditional on different determinants of taxing power and to identify the particular "problem" of poor tax performers. To illustrate the point, consider the following example: imagine a certain country is identified as a poor tax performer in the authors' baseline analysis. This result is confirmed if one adds year dummies and controls for the business cycle. However, when conditioning on the structure of the economy, the country moves from the poor tax performing group to the average group. This may suggest that the country's poor tax performance conditional on its development level is actually driven by its economic structure which makes it difficult for the country to raise taxes compared to other economies. In turn, if the country would still be in the underperforming group in this type of regression but would move to the normal performing group when political and institutional controls are included, this would suggest that the country's low

taxing power is actually related to the poor institutional and political setting. This type of analysis may provide an interesting and potentially valuable road map for political agents and development agencies in a particular country and help them understand what are the drivers of a country's poor tax performance in their particular case.

The point is well taken, but it is precisely the difficulty of deriving meaningful country-specific information from the large-N regression analyses we found in the literature that has motivated our qualitative approach. Our purpose is to provide a simple benchmarking tool for governments, donors and international organisations. With this in mind, it is important, for instance, to define clearly at which point political variables are integrated into the analysis. Should a low tax performing country be considered an "average" tax performer, because it is a dictatorship (or because it is corrupt), and compared to other dictatorships (corrupt states) it is actually performing well? Some studies present results from regressions with a large number of tax determinants, but which hide many important messages a qualitative analysis is able to uncover. As a further point, every additional variable included in the analysis comes at the cost of smaller sample size or lower data quality.

- I think it might even be interesting to include variables for the actual structure of the tax system in the regression. This would again help to understand whether countries are e.g. in the low-performing group because they rely more on some type of taxes than others (e.g. trade taxes vs. VAT vs. corporate taxes vs. personal income taxes).

As has been said above, this is a relevant observation and we would like to include tax structure in our analysis, but please note that even international organisations dedicating substantial resources to data collection find it extremely difficult to gather the necessary data. We consider this to be a feasible expansion of the analysis in the context of, for instance, case studies or small-N comparisons.

- Taken together, I think this is a nice paper which develops a novel and fruitful approach for assessing the relative tax performance of a country. However, I think the authors could develop their analysis further, taking more tax determinants into account and aiming to identify the particular "problems" of poorly performing countries.

We will try to sharpen our messages concerning this group of countries in the revised version of the paper.