We submit a reply to the anonymous referee 1, who points out the weak points of ours paper: i) the lack of empirical robustness checks with alternate data/estimation, ii) model variables and collinearity, and iii) GMM and other alternative methods. Thus, in response to him/her, below we explain our answers.

1) The lack of empirical robustness checks with alternate data/estimation. For instance, "hard" R&D data from World Bank.

In the paper we have used the R&D indicators from the GCI since they were readily available in the original dataset. The referee is right in pointing that these indicators, as most of the rest of indicators in the GCI dataset, are not "hard" data on R&D since they come from experts' survey rather than from direct observation.

In order to do a quick robustness check, we have obtained the data from WB databank on two R&D variables for the period and countries of interest.

Following is a correlation matrix of the World Bank data on R&D retrieved from the WB databank, with R&D indicators from the GCI dataset used in our study. The correlation is higher for the R&D expenditure than for the number of researchers but is high enough to presume that results would be robust to changes.

variable name	storage type	display format		variable label
rdexp	float	%9.0g	(WorldBank)	R&D expenditure as GDP pc.
rdnumr	float	%9.0g	(WorldBank)	R&D number of researchers
cci	float	%9.0g		Capacity for innovation
qsr	float	%9.0g		Quality R&D institutions
csrd	float	%9.0g		Company spending on R&D
uicol	float	%9.0g		Univ.-indust coop. in R&D

. correlate rdexp rdnumr cci qsr csrd uicol
(obs=463)

	rdexp	rdnumr	cci	qsr	csrd	uicol
rdexp	1.0000					
rdnumr	0.8951	1.0000				
cci	0.8724	0.7894	1.0000			
qsr	0.8071	0.7664	0.8593	1.0000		
csrd	0.8616	0.7703	0.9436	0.8644	1.0000	
uicol	0.7896	0.7515	0.8283	0.9094	0.8675	1.0000

2) Model variables and collinearity.

When choosing the variables our main guide has been the previous literature. The number of indicators in the GCI dataset is high, and we have complemented with some other variables from other sources but it is indeed true that one could think of other indicators from alternative sources. Nevertheless, we believe that the pool of variables we include is comprehensive. Regarding the case of collinearity between R&D and education, one of the things we did in a previous analysis was to compute the variance inflation factor of the correlation matrix for the variables of the pool. These results are not included in the paper but are available upon request. The VIFs for the R&D and education variables do not exceed 4 and the ones for the education enrollment rates are particularly low.

Thus, we don't think that including both R&D and education is causing a serious problem of collinearity and eliminating one or the other would not modify the results in essence. Moreover, what we propose is precisely a methodology, the Incremental Forward Stagewise regression, which checks for strong and robust specifications. Precisely, in table 3, our methodology drops the R&D variables and keeps the TEED.

3) GMM and other alternative methods.

Regarding the GMM suggestion we find it interesting but in the context of a dynamic setting, where Arellano and Bond methods should play a role. Ours is, on the contrary, a non-dynamic model, one of the reasons being that the piracy index data is discontinued.