

De Haas, Ralph; Popov, Alexander

Working Paper

A reply to Listo, Saberian and Thivierge (2023)

I4R Discussion Paper Series, No. 96

Provided in Cooperation with:
The Institute for Replication (I4R)

Suggested Citation: De Haas, Ralph; Popov, Alexander (2023) : A reply to Listo, Saberian and Thivierge (2023), I4R Discussion Paper Series, No. 96, Institute for Replication (I4R), s.l.

This Version is available at:
<https://hdl.handle.net/10419/280693>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



No. 96
DISCUSSION PAPER SERIES

A reply to Listo, Saberian and Thivierge (2023)

Ralph De Haas
Alexander Popov

This paper responds to:

Listo, Ariel, Soodeh Saberian, and Vincent Thivierge. 2023. Finance and green growth: A comment on De Haas and Popov (2023). *I4R Discussion Paper Series* No. 95. Institute for Replication.

December 2023

I4R DISCUSSION PAPER SERIES

I4R DP No. 96

A reply to Listo, Saberian and Thivierge (2023)

Ralph De Haas¹, Alexander Popov²

¹European Bank for Reconstruction and Development (CBRD), London/Great Britain, KU Leuven/Belgium, and Center for Economic Policy Research (CEPR), London/Great Britain

²European Central Bank (ECB), Frankfurt/Germany, and Center for Economic Policy Research (CEPR), London/Great Britain

DECEMBER 2023

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](https://www.zbw.eu/), and [RWI – Leibniz Institute for Economic Research](https://www.rwi-essen.de/), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

A reply to Listo, Saberian and Thivierge (2023)

Ralph De Haas¹ and Alexander Popov²

¹EBRD, KU Leuven and CEPR

²ECB and CEPR

November 2023

Abstract

Listo, Saberian and Thivierge (2023) conduct a careful replication of De Haas and Popov (2023) using the data, code, and instructions we made available at the time of publication. They highlight an inconsistency between the table notes and the main text in how the clustering of standard errors is described; uncover a coding mistake in the GMM regressions; and show that weak instruments are unlikely to bias our 2SLS coefficients. In this reply, we show that our results remain economically meaningful and statistically precise (though smaller in magnitude) when we (i) cluster standard errors by country in the country panel; (ii) correct the GMM code; and (iii) include or exclude China in the country and industry samples.

We thank Ariel Listo, Soodeh Saberian and Vincent Thivierge for their careful replication of [De Haas and Popov \(2023\)](#) – using the replication data, code, and instructions that we made available at the time of publication and after the usual checks by the EJ data editor. Here we provide our response to the issues that [Listo, Saberian, and Thivierge \(2023\)](#) bring up.

1. Standard errors

Claim: *"For their country panel, the authors mention in the text that they cluster their standard errors at the country level. However, in their script they either only adjusted their standard errors to account for heteroskedasticity or did not make standard error adjustments"*

Response: This refers to p. 642 of the published article where, when introducing the exploratory country-level regression framework, we write: *"We cluster the standard errors by country"*. At the same time, we write in the notes to Table 2 (p. 648 of the published article) that the reported standard errors in this table are heteroscedasticity robust (as far as the non-GMM results are concerned). While the latter is accurate, the text on p. 642 is not. This is an unfortunate and inadvertent oversight, which results from tables being updated during the review process without us fully bringing the related text in line. We should have checked the consistency between the text and the table notes more carefully to prevent leftover text like this.

Having said that, while definitely regrettable, we believe that the implications of this inconsistency are limited for two reasons. First, the main takeaway from Table 2 is robust to clustering the standard errors at the country level. Also when we include China (see point 5 below), the point estimate on *'Equity Share'* has a p -value of 0.10 in the OLS specification and of 0.06 in the IV specification. When correctly specifying the GMM model (see point 2 below), the p -value in this GMM specification is 0.11.¹ Second, we always presented the evidence in Table 2 as suggestive and as a first exploratory step in the data analysis. As we write on page 650, *"[...] we consider these country-level regressions mainly as a first exploratory step in our analysis [...]"*.

Regarding the GMM case, the answer lies in the claim in the next paragraph in Listo et al. (2023):

"[...] in the case of the GMM first stage estimation, the author (sic) did not adjust their standard errors, and therefore are assuming homoskedasticity. [...] we adjust the GMM standard errors to account for heteroskedasticity. As shown Stata is not able to produce robust standard errors."

Because Stata is indeed unable to produce robust standard errors in the tests using the country-industry-year panel, we opted for homoscedastic errors (that is, `vce(gmm)` instead of `vce(robust)`). To be consistent across specifications, we did the same in the country panel (Table 2, column 3). The text in the Notes to e.g. Table 3 should have read *"Standard errors clustered at the country-sector level [in the OLS and 2SLS case] are included in parentheses [underlining added]."*

¹These values are very similar when we drop China: 0.09 in the OLS; 0.08 in the IV; and 0.09 in the GMM.

2. GMM

Claim: "When reproducing their GMM estimators for their country and the country-industry panels, we uncovered that the authors improperly specified the GMM program in Stata. Indeed, they failed to recognize the endogenous variables."

Response: We believed that the way we had coded these regressions, namely

```
xi: xtabond cotwo_total_per_gdp l(0/0).(fin_dev1_l1 fin_str2_l1 log_gdp_per_cap_l1
    log_gdp_per_cap_sq_l1 pop_mil_l1 recession_l1 dom_st_lawpol_l1) i.year, lags(1)
    vce(gmm)
```

and

```
xi: xtabond cotwo_per_cap l(0/0).(dirty_fin_dev1_l1 dirty_fin_str2_l1 share_l1) i.ci
    i.it i.ct, lags(1) vce(gmm)
```

was the correct way to code an Arellano-Bond panel estimation accounting for endogeneity, because of the inclusion of "lags(1)" at the end, in both our country-year and country-industry-year panel specifications:

$$(1) \text{CO}_2\text{emissions}_{c,t} = \theta\text{CO}_2\text{emissions}_{c,t-1} + \beta_1\text{FD}_{c,t-1} + \beta_2\text{ES}_{c,t-1} + \beta_3X_{c,t-1} + \psi_c + \phi_t + \epsilon_{c,t}$$

and

$$(2) \text{CO}_2\text{emissions}_{c,s,t} = \theta\text{CO}_2\text{emissions}_{c,s,t-1} + \beta_1\text{FD}_{c,t-1} \times \text{CO}_2\text{intensity}_s \\ + \beta_2\text{ES}_{c,t-1} \times \text{CO}_2\text{intensity}_s + \psi_{c,s} + \gamma_{s,t} + \phi_{c,t} + \epsilon_{c,s,t}$$

Instead, we should have included "endogenous (fin_dev_l1 fin_str2_l1)" in the country-year panel regressions, and "endogenous (cotwo_fin_dev1_l1 cotwo_fin_str2_l1)" in the country-sector-year panel regressions. In this fashion, the variables 'Financial development' and 'Equity share' are treated as endogenous and their lags are included as instruments.

It is important to note that the GMM specifications estimated by the replication team do not fully align with those outlined in the original paper. Specifically, the replicators include one extra lag for the endogenous variables, deviating from the specification on p. 642 of the published article. As a result, the coefficient on the equity share in Table 1 is affected, with the point estimate being four times lower relative to the correct specification. This alteration not only decreases the coefficient's magnitude but also reduces its precision.

The code we now use to produce columns 1, 2 and 3 of Table 1 below is (using one-period lags):

```
xi: xtabond cotwo_total_per_gdp l(0/0).(log_gdp_per_cap_l1 log_gdp_per_cap_sq_l1
    pop_mil_l1 recession_l1 dom_st_lawpol_l1) i.year, lags(1) endogenous(fin_dev1_l1
    fin_str2_l1) vce(gmm)
```

and

```
xi: xtabond cotwo_total_per_gdp l(0/0).(log_gdp_per_cap_l1 log_gdp_per_cap_sq_l1
pop_mil_l1 recession_l1 dom_st_lawpol_l1) i.year, lags(1) endogenous(fin_dev1_l1
fin_str2_l1) vce(robust)
```

and

```
xi: xtabond cotwo_per_gdp l(0/0).(share_l1) i.it i.ct, lags(1) endogenous(
cotwo_fin_dev1_l1 cotwo_fin_str2_l1) vce(gmm)
```

We note here that the data is "xtset" using the panel and time variables as part of the cleaning code. This has two implications for how the "xtabond" code is written. First, there is no need to include country or country-industry fixed effects explicitly in the regression code since "xtabond" runs the specification on the differenced data, and therefore the group unit fixed effect is removed. Second, unlike in the case of "reg" or "reghdfe" because the data is "xtset", "vce(robust)" as part of "xtabond" invokes robust standard errors that are adjusted for clustering at the level of the panel variable set via "xtset", in this case, country.

Table 1 below reports GMM regressions in De Haas and Popov (2023), this time properly accounting for the endogeneity of the main explanatory variables (and using conventional instead of robust standard errors for the reason described above). In columns (1)–(3) we recover negative and mostly statistically significant point estimates on the main variable of interest in line with those in the original paper, though smaller in magnitude. Columns (4)–(9) replicate GMM regressions from Tables 4, 5, and 6 in De Haas and Popov (2023) accounting for endogeneity. Again, the GMM results are similar to those in the original paper, though smaller in magnitude.

TABLE 1. Robustness of the GMM estimations in De Haas and Popov (2023)

	CO ₂ emissions/GDP			Growth in value added	CO ₂ emissions/ value added	Green patents/ GDP (all)	Green patents/ GDP (excl. trans. & waste)	Green patents/ GDP (industrial production)	Green patents/ GDP (industrial production), OECD
	T2 (5)	T2 (5)	T3 (3)	T4 (3)	T5 (5)	T6 (1)	T6 (2)	T6 (3)	T6 (4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Financial development	0.0161*	0.0161*							
	(0.0090)	(0.0092)							
	[0.0728]	[0.0796]							
Equity share	-0.0435**	-0.0435							
	(0.0213)	(0.0276)							
	[0.0416]	[0.1147]							
Financial development × CO ₂ intensity			0.0000	-0.0012	-0.0541	-0.0003	0.0002	-0.0006	0.0000
			(0.0001)	(0.0060)	(0.0372)	(0.0013)	(0.0010)	(0.0005)	(0.0002)
			[0.6892]	[0.8465]	[0.1465]	[0.8381]	[0.8386]	[0.2539]	[0.8287]
Equity share × CO ₂ intensity			-0.0015***	-0.0093	-0.6829***	0.0011	0.0048	0.0040***	0.0016***
			(0.0003)	(0.0145)	(0.0878)	(0.0040)	(0.0031)	(0.0014)	(0.0005)
			[0.0000]	[0.5191]	[0.0000]	[0.7838]	[0.1254]	[0.0054]	[0.0017]
Sector share			0.0051*	-5.6575***	5.2725***	-0.0020	0.0051	0.0238**	-0.0012
			(0.0028)	(0.1260)	(0.9083)	(0.0265)	(0.0193)	(0.0093)	(0.0047)
			[0.0649]	[0.0000]	[0.0000]	[0.9399]	[0.7902]	[0.0105]	[0.8052]
Country FE	Yes	Yes							
Year FE	Yes	Yes							
Country × Sector FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector × Year FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	978	978	7,057	7,979	6,945	2,990	2,648	2,188	5,090

Notes: This table reports GMM regressions from De Haas and Popov (2023). Column headings refer to the dependent variable and reference the table and column number in De Haas and Popov (2023) (i.e. T2 (5) is Table 2 column 5 on p. 648 of De Haas and Popov (2023)). Conventional standard errors are reported in parentheses in columns (1) and (3)–(9) while in column (2) standard errors are clustered at the country level. Significance levels are indicated by * < .1, ** < .05, *** < .01.

3. First-stage

Claim: *"Since the authors have two first-stages for each model, it is more sensible to report F-statistics for the strength of their instruments for each first stage. Instead, the authors either fail to report their first-stage F-statistics or they only report one F-statistic. In Tables 1 and 2, we report the F-statistics for each of the first stages for the country and country-industry panels. In the case of Table 1, each F-statistic is below 10, which is an indication of weak instruments. For Table 2, only the instrument for the financial structure is above 10, which also suggests testing for the effects of weak instrument bias on the coefficients."*

Response: In our view, there is no clear consensus as to what to report as first-stage statistics. We have seen published papers reporting either separate F-statistics from the first stages or the F-statistics that Stata reports in the second stage. We opted for the latter. Having said that, we have always acknowledged and been fully transparent about the “weak instrument” problem. For example, on page 649 in the printed version of the paper, we write: *“While the Kleibergen-Paap LM statistic is reasonably high, the F-statistics are quite low, pointing to relatively weak instruments.”*

4. Weak instrument test

Claim: *"We view this as evidence that the potential weak instrument problem in this paper is not biasing the 2SLS estimator, especially for the structure of the financial sector coefficient."*

Response: We appreciate this additional analysis, which seems to confirm that the main coefficient of interest—the one on ‘Equity Share’—is not biased by weak instruments.

5. Consistent samples

Claim: *"Indeed, in their country sample, the authors note in their code that they drop China since there are not sufficient data on ‘No. environmental laws and policies’. However, they do not drop China in their country-industry panel models. The authors do not further discuss these choices in the code or in the main text. Therefore, as a replication exercise, we drop China from the country-industry sample in order to keep a consistent country sample across both panels. Table 4 presents the coefficients of interest for the interacted financial market size and structure when dropping China. Relative to the results presented in Table 2, which corrects for the coding errors, the results are qualitatively similar."*

Response: During the early stages of this project, the environmental data for China (which we need in the country-year but not in the country-sector-year analysis) were only available for two points in time. Our approach was deliberately to run each part of our analysis on the largest country sample possible. Dropping China from the sector-level analysis in order to be consistent with the country-level analysis would (in our view) be too high a price to pay for sample consistency across tables, especially given that we view the country-level regressions mainly as setting the scene for the later sector-level analysis. We thus wrote in the original code:

“Do not have sufficient data on "No. environmental laws and policies" alongside the line that drops China from the country-level sample.

Whilst preparing the replication package, we re-downloaded all data from scratch and at this point the full time-series data for China had become available. At that stage, we should have deleted the line of code that dropped China from the country-level analysis. When we include China in the country-year analysis, the main results go through, as shown in Table 1 above. In addition, it is reassuring that the country-sector analysis is robust to including or excluding China, too, and we appreciate you having made that explicit.

Conclusion

Thanks to the diligent work of the replicators, we have realized that there was an inconsistency between how we described the standard errors in the text and in the tables, and that we made a coding mistake by not properly accounting for endogeneity in the GMM part of the analysis. In this Reply, we have shown that our results remain economically meaningful and statistically precise (though smaller in magnitude) when we (i) cluster standard errors by country in the country panel; (ii) correct the GMM code; and (iii) include or exclude China in the country and industry samples.

References

- De Haas, Ralph, and Alexander Popov. 2023. “Finance and Green Growth.” *The Economic Journal* 133 (650): 637–668.
- Listo, A., S. Saberian, and V. Thivierge. 2023. “A Comment on De Haas and Popov (2023).”