Supplementary Material

"Functional distribution of income as a determinant of importing behavior: An empirical analysis"

Vinicius Curti Cícero
vinicius.cicero@colostate.edu

Gilberto Tadeu Lima
giltadeu@usp.br

Online Appendices

A Binary Variable for Developed/Developing Countries

Due to the nature of our equation of interest (Equation (15) in the paper), which is a dynamic specification, we cannot account for the heterogeneity between the groups of developed and developing countries by including a dummy variable in most of our estimations, even though, we would keep the number of observations of our full sample of countries. The reason is quite straightforward: in the case of the difference GMM, we are estimating a first-difference equation and, given the constancy of the dummy variable for all countries in the sample over the analysis period, the dummy is simply disregarded in the estimation. Besides, the inclusion of such a binary variable would imply that the two types of estimators would be little comparable in the presentation of the results. Still, it seems interesting to try to follow the suggestion made by an anonymous reviewer with a view to at least presenting an additional robustness test of our system GMM results for the two groups of countries.

Nevertheless, even for GMM system estimators, the inclusion of such a dummy variable to capture the heterogeneity between developing and developed countries proved problematic. Roodman (2009a) argues that one can include time-invariant regressors in system GMM estimations, but including a binary variable that has value 0 (or 1) for the majority of individuals - in our case, countries - might cause bias in the estimators particularly when the number of periods (T) is relatively small. In addition, there is an even more delicate

issue. The validity of the instruments in a system GMM estimation depends directly on the assumption that changes in the instrumenting variables are uncorrelated with the unobserved country-specific characteristic or fixed effects (Roodman, 2009a, p. 128). However, one can quite directly speculate that the dummy variable that captures the classification of a country as developed or developing in our sample (according to the intersection of World Bank and International Monetary Fund classifications) is related to time invariant country-specific factors.

Therefore, even if we do not incur in biased estimators, it is possible (and likely) that the instruments of our system GMM regressions are no longer valid with the inclusion of the binary variable. Although the statistical robustness is quite limited, due to the factors discussed above, we estimate the suggested regressions by system GMM with the inclusion of such a dummy variable, and the results are reported in Table 1 below.

	Baseline Specification	Intermediate Specification	Complete Specification
	System GMM	System GMM	System GMM
Log of import volume, lag 1	0.990***	0.912***	1.007***
	(0.06)	(0.05)	(0.08)
Log of import volume, lag 2			-0.110
			(0.07)
Log of import price	-0.055	-0.098	-0.157**
	(0.07)	(0.09)	(0.07)
Log of real GDP	-0.004	0.029	0.096
	(0.06)	(0.08)	(0.08)
Log of wage share	-0.380**	-0.427**	-0.322
	(0.16)	(0.21)	(0.23)
Developed dummy	0.056	0.005	0.007
	(0.10)	(0.10)	(0.07)
Constant	0.229	0.294	1.009
	(0.72)	(1.01)	(0.83)
Time-specific effects	Yes	Yes	Yes
Control variables	No	Yes	Yes
Number of lags (instruments)	8	5	7
AR(2) test - p value	0.505	0.587	0.121
Hansen "J" test - p value	0.009	0.124	0.033
Instruments	42	58	73
Observations	992	992	868
Groups	124	124	124

Notes: Standard errors, are reported in parentheses. Two-step standard errors are robust to the Windmeijer (2005) heteroscedasticity correction. In the Intermediate and Complete specifications, as in the original manuscript, we included the terms of trade, exchange rate, capital stock at constant prices, share of gross capital formation and share of government consumption (percentage of real GDP) as control variables to sharpen our analysis. In all estimations, only period dummies and the exchange rate are exogenous variables. All variables were transformed into natural logarithms. Following Roodman (2009b), we have collapsed the instruments in order to restrict its number. The null hypothesis of the Hansen test is that the instruments are not correlated with the residuals, and the null hypothesis of the AR(2) test is that the errors in the first difference regression have no second order serial correlation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1: Entire sample - Including binary variable for developed countries

B Robustness Checks for Aggregate Income

Another important point brought to our attention by an anonymous reviewer is the potential discrepancy between GDP and GNDI in some developing countries that compose our

sample. Indeed, it would be important for our analysis to consider such a discrepancy as an additional control in a robustness check of our empirical results for the subsample of developing countries.

In fact, income remittances alone account for a significant portion of the GNDI in some countries of our sample over the analyzed period. Using WDI data (again, from the World Bank's open database), of the 98 developing countries in our sample, 10 of them have personal remittances received as a percentage of GDP which are greater than 15% as an average for the 2001-2017 period. In order of magnitude, we have the following: Lesotho, Tajisktan, Madagascar, Kyrgyzstan, Lebanon, Bermuda, Jordan, Honduras, Armenia, and Jamaica. Table 2 below presents the code of such countries in our database and the average personal remittances received as a share of GDP from 2001 to 2017.

Country	Country Code	Personal Remittances Received (% of GDP)
Lesotho	**LSO**	31.27%
Tajikstan	**TJK**	30.46%
Madagascar	**MDA**	24.21%
Kyrgyzstan	**KGZ**	19.96%
Lebanon	**LBN**	18.98%
Bermuda	**BMU**	18.52%
Jordan	**JOR**	16.44%
Honduras	**HND**	16.25%
Armenia	**ARM**	16.16%
Jamaica	**JAM**	15.34%

Table 2: List of countries, codes and share

In view of the preceding evidence on the magnitude of the personal remittances received as a share of GDP, and the importance of investigating the impact of such a discrepancy between GDP and GNDI in countries in our sample, we have structured the needed robustness check as follows.

First, looking only at the subsample of developing countries, we have verified the stability of our results by repeating our econometric estimates by first removing from the sample the top five personal remittances recipients according to Table 2 (with that income representing approximately 20% of GDP on average in the period of analysis) and, subsequently, all the ten countries described above. Note that, in this second case, we are reducing our sample of developing countries by slightly more than 10%. Nevertheless, the results are qualitatively similar to those presented in the paper - which points to the robustness of our empirical analysis to this characteristic of some countries in our sample. Next, to leave no doubt regarding "fallacy of composition" problems and also as a sensitivity analysis, we have estimated the preferred (and statistically most robust, as discussed in detail in the paper) specification by separately disregarding each one (that is, disregarding one at a time) of the countries listed in Table 2 in each regression. We graphically show that the estimated coefficients of interest, associated with the functional distribution of income, are qualitatively and quantitatively very similar to the ones presented in the paper. Finally, we have followed the same steps to structure the robustness check for the results concerning the entire sample of countries.

B.1 Developed Countries

We began our robustness check by estimating the import function described in Equation (15) of the paper for the sample of developing countries initially disregarding the top five countries in terms of personal remittances received as a share of GDP presented in Table 2. Table 3 presents the results. Similar to the paper, and throughout this document, the first two columns show the results of both GMM estimators for the baseline specification (considering only one lag of the dependent variable and without control variables), the third and fourth columns present the results of the GMM estimators for the intermediate specification (that considers the set of control variables and one lag of the explained variable), and, the fifth and sixth columns show the estimation results of the complete specification (considering two lags of the import volume and the set of control variables).

Note that, compared to Table 2 of the paper, the results presented in Table 3 here are quite similar. In particular, looking at the variable of interest - the log of wage share -, even though the magnitude of the coefficient increases after constraining our sample of developed countries in columns one, two, and six, there is a slight reduction in the magnitude of the estimated coefficients for the Difference GMM estimation in both the intermediate and

	Baseline specification		Intermediate specification		Complete specification	
	Diff GMM	System GMM	Diff GMM	System GMM	Diff GMM	System GMM
Log of import volume, lag 1	0.233**	0.978***	0.088	0.826***	0.341***	0.910***
	(0.09)	(0.07)	(0.06)	(0.06)	(0.04)	(0.09)
Log of import volume, lag 2					-0.020	-0.092
					(0.02)	(0.08)
Log of import price	-0.553***	-0.082	-0.798***	-0.166	-0.470***	-0.188
	(0.14)	(0.07)	(0.14)	(0.13)	(0.12)	(0.12)
Log of real GDP	1.132***	-0.006	1.851***	0.025	1.458***	0.087
	(0.30)	(0.05)	(0.23)	(0.12)	(0.18)	(0.12)
Log of wage share	-0.142	-0.394**	-0.734***	-0.586**	-0.105	-0.472*
	(0.31)	(0.19)	(0.25)	(0.25)	(0.15)	(0.27)
Constant		0.259		0.799		0.973
		(0.59)		(1.26)		(1.23)
Time-specific effects	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	No	Yes	Yes	Yes	Yes
Number of lags (instruments)	8	8	5	5	7	7
AR(2) test - p value	0.558	0.457	0.311	0.498	0.334	0.099
Hansen "J" test - p value	0.100	0.143	0.111	0.444	0.296	0.162
Instruments	36	41	48	57	63	72
Observations	651	744	651	744	558	651
Groups	93	93	93	93	93	93

Notes: Standard errors, are reported in parentheses. Two-step standard errors are robust to the Windmeijer (2005) heteroscedasticity correction. In the Intermediate and Complete specifications, as in the original manuscript, we included the terms of trade, exchange rate, capital stock at constant prices, share of gross capital formation and share of government consumption (percentage of real GDP) as control variables to sharpen our analysis. In all estimations, only period dummies and the exchange rate are exogenous variables. All variables were transformed into natural logarithms. Following Roodman (2009b), we have collapsed the instruments in order to restrict its number. The null hypothesis of the Hansen test is that the instruments are not correlated with the residuals, and the null hypothesis of the AR(2) test is that the errors in the first difference regression have no second order serial correlation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Robustness check for developing countries - Dropping top five countries

complete specification. The only qualitative difference between the results presented in Table 3 is in the complete specification, as the coefficient associated with the log of wage share is no longer statistically significant at the 5% significance level. Even with this caveat, it is worth noting that, in the aggregate, the results are quite similar to those presented in the paper - which then indicates the robustness of our estimates to the fact that some developing countries in our sample are characterized by an important distinction between GDP and GNDI due to personal remittances (of income).

Furthermore, in the next step of this robustness check, we removed all ten countries listed

in Table 2 from the sample and re-estimated the regressions. The results are reported in Table 4 , in which we observe similar variations in the coefficients of interest compared to

	Baseline specification		Intermediate specification		Complete specification	
	Diff GMM	System GMM	Diff GMM	System GMM	Diff GMM	System GMM
Log of import volume, lag 1	0.242**	0.994***	0.184***	0.853***	0.331***	0.928***
	(0.10)	(0.06)	(0.06)	(0.06)	(0.04)	(0.09)
Log of import volume, lag 2					0.000	-0.086
					(0.02)	(0.08)
Log of import price	-0.502***	-0.062	-0.710***	-0.117	-0.445***	-0.134
	(0.17)	(0.07)	(0.16)	(0.14)	(0.14)	(0.14)
Log of real GDP	1.188***	-0.010	1.948***	0.007	1.593***	0.077
	(0.33)	(0.05)	(0.24)	(0.12)	(0.16)	(0.12)
Log of wage share	-0.114	-0.354*	-0.486**	-0.547**	-0.098	-0.476*
	(0.32)	(0.19)	(0.25)	(0.24)	(0.13)	(0.27)
Constant		0.141		0.477		0.689
		(0.69)		(1.34)		(1.34)
Time-specific effects	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	No	Yes	Yes	Yes	Yes
Number of lags (instruments)	8	8	5	5	7	7
AR(2) test - p value	0.663	0.532	0.440	0.576	0.421	0.145
Hansen "J" test - p value	0.138	0.212	0.098	0.225	0.337	0.161
Instruments	36	41	48	57	63	72
Observations	616	704	616	704	528	616
Groups	88	88	88	88	88	88

Notes: Standard errors, are reported in parentheses. Two-step standard errors are robust to the Windmeijer (2005) heteroscedasticity correction. In the Intermediate and Complete specifications, as in the original manuscript, we included the terms of trade, exchange rate, capital stock at constant prices, share of gross capital formation and share of government consumption (percentage of real GDP) as control variables to sharpen our analysis. In all estimations, only period dummies and the exchange rate are exogenous variables. All variables were transformed into natural logarithms. Following Roodman (2009b), we have collapsed the instruments in order to restrict its number. The null hypothesis of the Hansen test is that the instruments are not correlated with the residuals, and the null hypothesis of the AR(2) test is that the errors in the first difference regression have no second order serial correlation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Robustness check for developing countries - Dropping all ten countries

those shown in Table 3. In summary, the results are qualitatively quite similar to those of the paper, with a small reduction in the statistical significance of the coefficient associated with the functional distribution of income for one of the estimations (potentially, in this case, due to the exclusion of 10% of the countries in the sub-sample).

Finally, as a sensitivity analysis and to deal separately with each of the countries presented in Table 2, we estimated the intermediate specification by removing, in each regression, one of

the analyzed countries. Figure 1 graphically presents the results, emphasizing the coefficient of interest and the 95% confidence interval for each one of the regressions, named by the country dropped (their country code) in each estimation. The dotted black line describes the coefficient obtained for the complete sample of developing countries (as presented in Table 2 of the paper). Notice that both qualitatively and quantitatively the additional results are very close to the estimation presented in the paper.

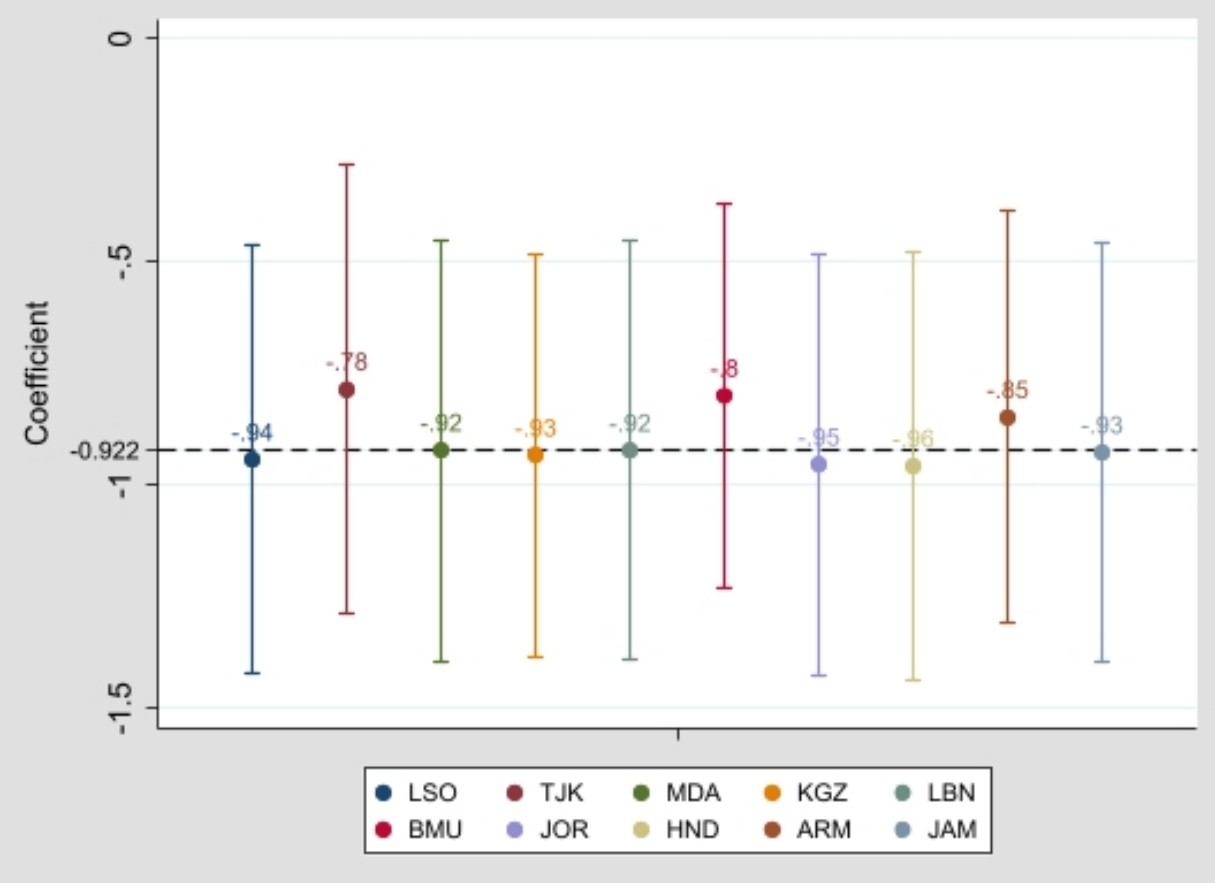

Figure 1: Intermediate specification - Diff GMM - Sensitivity analysis (developing countries)

In summary, the empirical-econometric exercise developed in this section confirmed that the results presented in the paper for the subsample of developed countries are robust to an important characteristic of some of these countries: the discrepancy between GDP and GNDI. To complete our analysis, we look in the next section at the entire sample of countries.

B.2 Entire Sample

We also performed the same robustness check for the entire sample of countries. Initially, as in the previous section, we dropped the top five countries in terms of personal remittances received as a share of GDP presented in Table 2. The results are reported in Table 5.

	Baseline specification		Intermediate specification		Complete specification	
	Diff GMM	System GMM	Diff GMM	System GMM	Diff GMM	System GMM
Log of import volume, lag 1	0.198**	0.955***	0.067	0.899***	0.256***	1.011***
	(0.09)	(0.04)	(0.06)	(0.05)	(0.04)	(0.09)
Log of import volume, lag 2					-0.006	-0.115
					(0.02)	(0.08)
Log of import price	-0.521***	-0.138	-0.815***	-0.106	-0.703***	-0.160**
	(0.11)	(0.10)	(0.13)	(0.10)	(0.09)	(0.08)
Log of real GDP	0.909***	-0.014	2.061***	0.013	1.721***	0.066
	(0.21)	(0.05)	(0.20)	(0.11)	(0.18)	(0.12)
Log of wage share	-0.216	-0.465***	-0.710***	-0.491**	-0.295*	-0.379*
	(0.27)	(0.16)	(0.23)	(0.20)	(0.17)	(0.23)
Constant		0.896		0.250		0.977
		(0.99)		(1.04)		(0.95)
Time-specific effects	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	No	Yes	Yes	Yes	Yes
Number of lags (instruments)	8	8	5	5	7	7
AR(2) test - p value	0.571	0.411	0.384	0.494	0.352	0.090
Hansen "J" test - p value	0.035	0.008	0.144	0.171	0.158	0.045
Instruments	36	41	48	57	63	72
Observations	833	952	833	952	714	833
Groups	119	119	119	119	119	119

Notes: Standard errors, are reported in parentheses. Two-step standard errors are robust to the Windmeijer (2005) heteroscedasticity correction. In the Intermediate and Complete specifications, as in the original manuscript, we included the terms of trade, exchange rate, capital stock at constant prices, share of gross capital formation and share of government consumption (percentage of real GDP) as control variables to sharpen our analysis. In all estimations, only period dummies and the exchange rate are exogenous variables. All variables were transformed into natural logarithms. Following Roodman (2009b), we have collapsed the instruments in order to restrict its number. The null hypothesis of the Hansen test is that the instruments are not correlated with the residuals, and the null hypothesis of the AR(2) test is that the errors in the first difference regression have no second order serial correlation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Robustness check for entire sample - Dropping top five countries

Compared to Table 3 of the paper, the additional results presented in Table 5 below are very similar - both quantitatively and qualitatively. The only change to be highlighted for the coefficients associated with the log of wage share is that the standard errors increased

marginally, as expected due to the reduction of observations, even though the coefficients maintained statistical significance. Again, as in the case of the results for developing countries presented earlier, it is interesting to note that the exclusion of the considered countries, also as expected, increased the magnitude of the coefficients associated with real GDP (the estimated income elasticity of imports).

Next, we removed all ten countries listed in Table 2 from the sample and re-estimate the regressions. The results are presented in Table 6. Notice that we observe similar variations

	Baseline specification		Intermediate specification		Complete specification	
	Diff GMM	System GMM	Diff GMM	System GMM	Diff GMM	System GMM
Log of import volume, lag 1	0.194**	0.964***	0.206***	0.914***	0.270***	1.018***
	(0.09)	(0.04)	(0.06)	(0.05)	(0.04)	(0.09)
Log of import volume, lag 2					0.001	-0.110
					(0.02)	(0.08)
Log of import price	-0.476***	-0.110	-0.674***	-0.049	-0.719***	-0.114
	(0.12)	(0.12)	(0.13)	(0.11)	(0.10)	(0.10)
Log of real GDP	0.860***	-0.009	1.925***	0.003	1.842***	0.066
	(0.21)	(0.05)	(0.19)	(0.11)	(0.15)	(0.12)
Log of wage share	-0.197	-0.426***	-0.683***	-0.491**	-0.383***	-0.406*
	(0.28)	(0.16)	(0.22)	(0.20)	(0.14)	(0.24)
Constant		0.495		-0.257		0.296
		(1.15)		(1.15)		(1.07)
Time-specific effects	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	No	Yes	Yes	Yes	Yes
Number of lags (instruments)	8	8	5	5	7	7
AR(2) test - p value	0.615	0.484	0.466	0.567	0.384	0.133
Hansen "J" test - p value	0.053	0.013	0.084	0.024	0.149	0.031
Instruments	36	41	48	57	63	72
Observations	798	912	798	912	684	798
Groups	114	114	114	114	114	114

Notes: Standard errors, are reported in parentheses. Two-step standard errors are robust to the Windmeijer (2005) heteroscedasticity correction. In the Intermediate and Complete specifications, as in the original manuscript, we included the terms of trade, exchange rate, capital stock at constant prices, share of gross capital formation and share of government consumption (percentage of real GDP) as control variables to sharpen our analysis. In all estimations, only period dummies and the exchange rate are exogenous variables. All variables were transformed into natural logarithms. Following Roodman (2009b), we have collapsed the instruments in order to restrict its number. The null hypothesis of the Hansen test is that the instruments are not correlated with the residuals, and the null hypothesis of the AR(2) test is that the errors in the first difference regression have no second order serial correlation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Robustness check for entire sample - Dropping all ten countries

in the coefficients of interest compared to those presented in Table 5. Once again, the results

presented here are quantitatively and qualitatively quite similar to those of the paper.

Lastly, we performed the sensitivity analysis for the preferred specification for the entire sample of countries. Figure 2 graphically represents the results, emphasizing the coefficient of interest and the 95% confidence interval for each one of the regressions, named by the country dropped (their code) in each estimation. Similar to Figure 1, the dotted black line describes the coefficient obtained for the complete sample of countries (as presented in Table 3 of the paper).

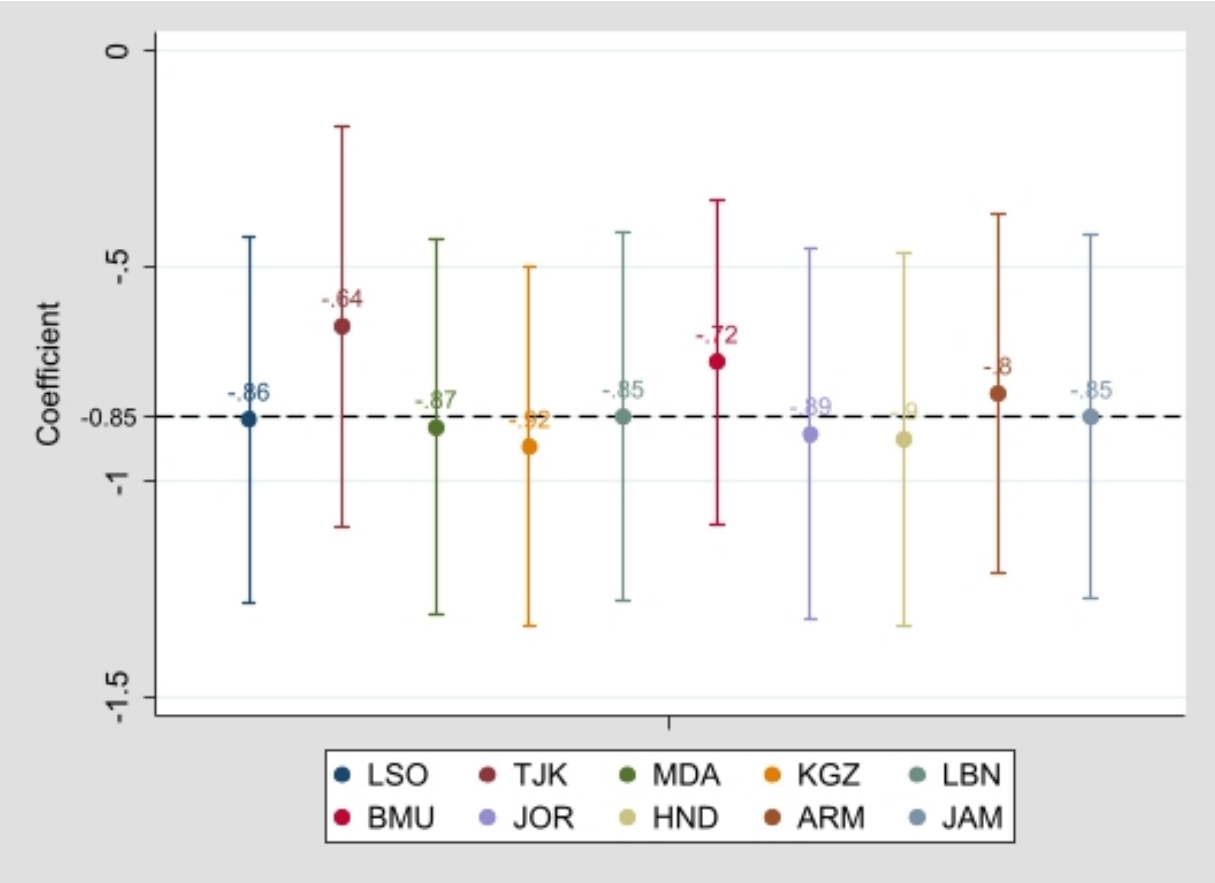

Figure 2: Intermediate specification - Diff GMM - Sensitivity analysis (full sample)

By way of conclusion, it is worth presenting a brief summary of the robustness tests presented in this section of this appendix. In short, the exclusion of the main countries receiving personal remittances as a share of GDP, both the top five and the top ten in our sample, did not qualitatively affect the results presented in the paper, whether looking only at the group of developing countries or looking at the entire sample of countries. We showed that our empirical analysis is not altered by the inclusion or exclusion of countries in which

remittances account for a significant portion of the GNDI. In addition to the robustness checks, we also performed a sensitivity analysis that highlighted the stability of the results presented in the paper when considering countries where the discrepancy between GDP and GNDI is economically significant. As a whole, our original results proved to be robust - both quantitatively and qualitatively - to specific characteristics of some countries in the sample.

References

Roodman, D. (2009a). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, *9*(1), 86–136.

Roodman, D. (2009b). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, *71*(1), 135–158.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, *126*(1), 25–51.