weight matrix being evaluated at the two-step estimate. Comparing the p-values one finds that the J test can be quite sensitive to the user's choice of the estimate where the efficient weight matrix is evaluated. The result of AJRY is based on the popular Stata command *xtabond2* for dynamic panel models. The command calculates the J statistic and the p-value based on the two-step GMM with the one-step weight matrix even when the one-step GMM estimates are reported. This does not seem to be a reliable test but there is no clear guideline on this. We recommend using our J statistic (40) which is not subject to such arbitrariness.

AJRY conclude that income does not have a statistically significant causal effect on democracy. Our results in Table 5 reinforce and extend this conclusion.

	Column (4) five-year data		Column (8) ten-year data	
	One-step	Iterated	One-step	Iterated
	I	II	III	IV
Democracy$_{t-1}$	0.489	0.744	0.227	0.288
Arellano-Bond s.e.	(0.085)	(0.043)	(0.123)	(0.111)
*Windmeijer-corrected s.e.		(0.150)		(0.161)
Misspecification-Robust s.e.	(0.095)	(0.128)	(0.125)	(0.146)
Income$_{t-1}$	-0.129	-0.009	-0.318	-0.280
Arellano-Bond s.e.	(0.076)	(0.040)	(0.180)	(0.170)
*Windmeijer-corrected s.e.		(0.054)		(0.203)
Misspecification-Robust s.e.	(0.088)	(0.039)	(0.183)	(0.202)
Cumulative Income Effect	-0.253	-0.036	-0.411	-0.393
Arellano-Bond s.e.	(0.148)	(0.152)	(0.243)	(0.243)
*Windmeijer-corrected s.e.		(0.206)		(0.305)
Misspecification-Robust s.e.	(0.163)	(0.149)	(0.246)	(0.290)
Hansen J Test	[0.04]	[0.42]	[0.08]	[0.09]
# of Iteration	0	23	0	9
# of Instruments	55		15	
Observations	838		338	
Countries	127		118	

Standard errors clustered by country

Table 5: Extension of Acemoglu, Johnson, Robinson and Yared (2008), Table 2

What are the causes of potential misspecification? One possibility is that the dynamic structure in (45) is incorrect – that lagged values are omitted. If the dynamics are misspecified, then the moment conditions are not satisfied and the Arellano-Bond standard errors will be incorrect. Since the "true" dynamic structure of a panel regression is not known *a priori*, this is a strong reason to generically allow for misspecification.

Another reason for potential misspecification is coefficient heterogeneity. If the coefficients are heterogeneous across countries, then moment conditions will not be satisfied. For example, in model

(45), if the coefficient γ_i (the effect of income on democracy) varies with country i, then the moment conditions will be invalid. To see this, if we set $\gamma = E\gamma_i$ as the mean coefficient, then the effective error in the differenced equation (45) is $\Delta u_{it} + (\gamma_i - \gamma)\Delta x_{i,t-1}$ which will be correlated with the instrument $y_{i,t-2}$. A consequence is that the Arellano-Bond standard errors will be incorrect, but our misspecification-robust standard errors will be appropriate.

There is strong evidence for coefficient heterogeneity in equation (45). Cervellati, Jung, Sunde, and Vischer (2014, CJSV hereinafter) argue that the income effect is heterogeneous between former colonies and non-colonies, and furthermore within colonies based on the quality of political institutions. Bonhomme and Manresa (2015) find evidence of grouped patterns of unobserved heterogeneity in the same dataset. Lu and Su (2017) also find strong evidence of heterogeneity in the income effect across countries. This literature makes a clear case that the coefficients (primarily γ) vary across countries. In this case, model (45) should be viewed as an approximation rather than a tight statistical model. The coefficients should be viewed as projections and the moment conditions acknowledged to be potentially invalid. In this context Arellano-Bond standard errors are incorrect, and our misspecification-robust standard errors appropriate.

	Constraints		Independence		No Late Colonial	
	One-step	Iterated	One-step	Iterated	One-step	Iterated
	I	II	III	IV	V	VI
Democracy_{t-1}	0.289	-0.423	0.343	0.724	0.355	0.666
Arellano-Bond s.e.	(0.123)	(0.039)	(0.110)	(0.044)	(0.101)	(0.040)
*Windmeijer-corrected s.e.		(0.413)		(0.168)		(0.141)
Misspecification-Robust s.e.	(0.142)	(0.380)	(0.127)	(0.152)	(0.115)	(0.125)
Income_{t-1}	-0.417	-0.337	-0.270	-0.011	-0.303	-0.052
Arellano-Bond s.e.	(0.194)	(0.116)	(0.113)	(0.050)	(0.110)	(0.047)
*Windmeijer-corrected s.e.		(0.267)		(0.060)		(0.057)
Misspecification-Robust s.e.	(0.221)	(0.289)	(0.134)	(0.047)	(0.122)	(0.041)
$\text{Income}_{t-1} \times c_i$	0.345	0.296	0.224	0.020	0.318	0.111
Arellano-Bond s.e.	(0.162)	(0.073)	(0.121)	(0.037)	(0.122)	(0.039)
*Windmeijer-corrected s.e.		(0.369)		(0.087)		(0.075)
Misspecification-Robust s.e.	(0.169)	(0.309)	(0.125)	(0.077)	(0.130)	(0.053)
Hansen J Test	[0.03]	[0.02]	[0.04]	[0.38]	[0.15]	[0.37]
# of Iteration	0	297	0	32	0	28
# of Instruments	56		56		56	
Observations	531		628		631	
Countries	79		99		100	

Standard errors clustered by country

Table 6: Extension of Cervellati, Jung, Sunde, and Vischer (2014), Table 4

To highlight this issue further we examine a key table from CJSV (their Table 4) where they