View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Repositori Institucional de la Universitat Jaume I

UNIVERSITAT JAUME·I

Título artículo / Títol article:	Comment on Lof, Mekasha, and Tarp (2014)
Autores / Autors	Herzer, Dierk ; Nowak Lehmann, Felicitas ; Dreher, Axel ; Klasen, Stephan ; Martínez Zarzoso, Inmaculada
Revista:	World Development Volume 70, June 2015
Versión / Versió:	Preprint de l'autor
Cita bibliográfica / Cita bibliogràfica (ISO 690):	HERZER, Dierk, et al. Comment on Lof, Mekasha, and Tarp (2014). World Development, 2015, 70: 389-396.
url Repositori UJI:	http://hdl.handle.net/10234/125244

Comment on Lof, Mekasha, and Tarp (2014)

DIERK HERZER[a], FELICITAS NOWAK-LEHMANN[b], AXEL DREHER[c,b], STEPHAN KLASEN[b] and INMACULADA MARTINEZ-ZARZOSO[b,d]
[a] Helmut Schmidt University Hamburg, Germany
[b] University of Göttingen, Germany
[c] Heidelberg University, Germany
[d] University Jaume I, Spain

(preprint version, World Development Vol. 70, pp. 389–396, 2015)

1. Introduction

The question of whether "foreign aid works" is fiercely debated. Proponents of a big push suggest to substantially increase the amount of aid (Sachs 2006), while some skeptics go as far to propose ending aid altogether (Moyo 2009). One reason for why this important question is still unsettled is the failure of the academic literature to find a robust positive effect of aid on growth, using a large variety of econometric approaches (Easterly et al. 2004, Rajan and Subramanian 2008, Werker et al. 2009). In a recent contribution to this debate, we proposed taking account of the underlying variables' time series properties to minimize the risk of spurious regressions, among other approaches relying on panel dynamic feasible generalized least-squares (DFGLS) to investigate the effect of aid on per-capita income (Nowak-Lehman et al. 2012). In line with much of the literature, our results show no significant positive effect of aid.

Lof, Mekasha, and Tarp (2014) (henceforth LMT) criticize one of our approaches in Nowak-Lehman et al. 2012) as unsuitable and for omitting a large number of observations from our sample. This occurred as a result of taking logs of covariates that sometimes take on negative values. We welcome the LMT study, as it allows us to show that our results do not depend on the particular choice of countries and observations included in the sample. We also welcome the opportunity to clarify some of the methodological choices we had to deal with. While we thus appreciate the LMT study, we are surprised by their own empirical approach and conclusions. They fail to test the key assumptions underlying their model and measure aid in per capita terms rather than relative to GDP, which introduces an upward bias to their estimates. Even LMT's own results show the aid-income relationship to be fragile.

In the reminder of this comment, we will first test for the robustness of our results in Nowak-Lehman et al. (2012) to different treatments of missing observations and provide some reflections on the methods we used. We then turn to the approach in LMT of estimating the aid-income relationship using different time series methods and a different aid variable. To anticipate our key findings, the results reported in Nowak-Lehman et al. (2012) are robust to alternative ways of dealing with missing data. Regarding the analysis of the aid-income relationship in LMT, we show that the way they measure aid as well as the econometric methods they propose are inappropriate. When we modify their approach to take account of some of these problems, we find no evidence that aid has a positive long-run effect on income. In the conclusion, we point to some agreements between us and LMT but also emphasize the major differences.

2. Tests for Robustness, Nowak-Lehmann et al. (2012)

In Nowak-Lehmann et al. (2012) we examine the impact of aid on per capita GDP in a Solow-type framework. In the underlying Cobb-Douglas production function, the covariates enter the income equation multiplicatively, so we log transform them to linearize the model. [1] We investigate the impact of aid as a share of GDP rather than using the absolute value of aid or aid/capita, in line with much of the literature on aid effectiveness (and being aware that this variable varies with aid in the numinator and GDP in the denominator alike).

In our econometric approach, we made use of time series techniques with the purpose to minimize the risk of spurious regressions (which occur when non-stationary variables are regressed on each other), controlling for endogeneity of all regressors and eliminating autocorrelation which again can lead to biased results. As discussed in detail in Nowak-Lehmann et al. (2012), after first

[1] When we estimate the model in absolute levels using an additive framework, our results are qualitatively unchanged. See below.

providing some evidence of a potentially cointegrating relationship and Ganger causality tests suggesting that aid causes income in the short and long-run (while acknowledging that, in the long-run, the aid-income relationship is bi-directional), we argue that DFGLS is well-suited to address these econometric issues in a time series context.[2]

In deriving some of our specifications, particularly those including (non-aid) external savings as a covariate, we lose large parts of the potential sample mostly due to taking logs of negative external savings values, as correctly pointed out in Table 1 of LMT which is the exact replication of our results (Nowak-Lehmann et al. 2012, Table 1). Our reference to the use of a "balanced sample" was thus unfortunate and misleading. However, our results do not depend on this choice. As shown in Table 1 of LMT, the coefficient of aid is scarcely affected by whether or not observations are dropped. In the first column, where hardly any observations are dropped, the coefficient of aid is negative (and significant). In the subsequent columns, where the number of dropped observations becomes larger, the coefficient is negative and of essentially identical in magnitude but turns insignificant. The result thus seems hardly affected by the inclusion of further covariates. Given that we find some evidence of cointegration between aid and per capita income and Granger causality tests suggesting a relationship running from aid to income in the short and long-run, we estimate the relationship between aid per capita and income and are able to reject the h. of a positive relationship and find some non-robust evidence of a small *negative* one.

To investigate how the dropped observations might affect our results, we provide a series of robustness tests in Table 1, treating dropped observations in four different ways.[3] Column 1 reproduces our DFGLS regressions dropping zero and negative values for any of the included variables, as in Nowak-Lehmann et al. (2012). In column 2, we exclude external savings from the

[2] We use a model with a model that includes two lags in the estimation based on both model specification tests (using the Bayesian Information Criteria) and tests for autocorrelation (using Durban-Watson Statistics).
[3] LMT also criticize that our theoretical model is mis-specified as the multiplicative framework in the Solow model cannot sensibly deal with negative values of growth determinants. Dropping these observations (as done in Nowak-Lehmann et al, 2012) or lifting them into positive territory as done here arguably addresses this issue as well so that the potential problem of mis-specification has been addressed in these robustness checks.

model, substantially increasing the number of observations. Column 3 adds a positive constant to the negative values before taking logs (a commonly used technique to address this issue in the literature),[4] while column 4 replaces missing values after taking logs with zeroes and adding a binary indicator variable that is one in this case (and zero otherwise). Thus the sample is then set in a way that all observations are included but the missing values do not affect the computation of the coefficients of the respective covariates. Lastly, column 5 includes the covariates in levels and thus estimates an additive model. There, the problem of dropped observations due to the log transformation does not arise.[5]

Table1 Does a different handling of the log transformation of net external savings change the results?

	Dependent variable: Per-capita income (multiplicative model)[6] Estimation via DFGLS				Additive Model
	Full model	Model without net external savings	Model with transformation of net external savings	Model controlling for missing values of net external savings	Model with unlogged variables
	(1)	(2)	(3)	(4)	(5)
Population growth	-0.003 (-0.02)	0.10 (1.25)	0.13 (1.62)	-0.01 (-0.20)	-1532.77 (-0.95)
Domestic savings	0.07*** (5.56)	0.08*** (7.27)	0.09*** (8.17)	0.06*** (6.02)	1170.52*** (3.60)
Net external savings	0.05*** (4.79)	------	0.09*** (5.57)	0.03*** (4.14)	1534.95*** (2.94)
Net aid transfer (aid-to-GDP)	-0.01 (-1.47)	-0.01* (-1.89)	-0.02*** (-2.85)	-0.01 (-1.61)	-547.85** (-2.14)
External savings missing				-0.01*** (-2.63)	
Fixed effects	yes	yes	Yes	yes	Yes

[4] We add 15 to external savings to ensure that all observations are positive.

[5] We include this specification as a further robustness check. As we derive our model from the Solow framework with a Cobb-Douglas production function, we prefer the models in logs that are derived from this framework.

[6] All variables in columns 1-4 are in logs. In column 5 the model has been estimated with unlogged variables as a robustness check.

2 leads and 2 lags	yes	yes	yes	yes	Yes
Cross sections included	50	56	56	58	57
Periods included	41	41	41	42	41
Number of observations	755	1642	1543	2389	2014
R-squared adj.	0.99	0.99	0.99	0.99	0.99
Durbin-Watson stat.	2.02	1.72	1.80	1.83	1.61

t-values are in parentheses. *** (**): significant at the one (five) percent level. All variables are in the firstfour columns are in logs, while they are in levels in the fifth column.

As can be seen from Table 1, these different specifications all lead to the same conclusion, which is a negative coefficient of the aid variable that is either significant or marginally insignificant. We conclude that our approach to drop these observations does not affect the results.[7]

3. LMT – An alternative time-series approach?

Rather than testing for the robustness of the results in Nowak-Lehmann et al. (2012), LMT apply a totally different time series approach to show that aid increases income. They relate their analysis to ours by pointing to the use of the same data. However, they choose to omit the earlier years of the data from their sample, in order to increase the number of countries in their panel. What is more, LMT omit our control variables from most of their regressions. This makes their analysis hardly comparable to ours in any dimension, other than relying on the same source for the GDP per capita and aid variables. Specifically, LMT estimate a panel vector autoregression (VAR)/vector error-correction (VEC) model. As we did, they aim to examine the causal effect of aid on GDP per capita.

[7]Interestingly, the coefficient on the missing external savings dummy is negative and significant, suggesting that negative external savings are associated with lower growth. We also include an interaction effect between the missing external savings dummy and the aid variable which is insignificant and does not affect the coefficient of the main effect of the aid variable. Results are available on request.

Instead of using aid as a share of GDP, as we do, they prefer to use aid per capita of the recipient population. Relying on impulse response analysis, they conclude that aid per capita has a positive and statistically significant long-run effect on income. We are not convinced by their approach, for a number of reasons.

First, aid as a share of GDP can directly be linked to a Solow-type growth model, while the estimates in LMT have no theoretical foundation, and most of them are based on bivariate estimates only. What is more, expressing aid and income in per capita terms introduces an upward bias to the regressions. To the extent that shocks to population affects aid per capita and GDP/capita in the same way through the same denominator, this will cause aid per capita to be positively related to income per capita even in the absence of any effect of aid on income.

Second, we are not convinced by LMT's analysis for a number of additional reasons. Their choice of which observations should be included in the estimations is as ad hoc as ours, and can thus hardly be considered as an improvement. Specifically, for five countries in their 59-country sample,[8] complete time-series on GDP per capita and net aid transfers (relative to GDP) are not available for the 1970-2006 period in the original data set.[9] LMT log their data, dropping all observations on net aid transfers with negative values, as we did in Nowak-Lehman et al. (2012). There are eight countries in their sample with negative values for net aid transfers and thus with incomplete time-series data after taking logs.[10] What is more, LMT choose to omit observations for the 1960-69 period, in order to increase the number of countries with complete time-series. Their results can thus hardly be considered to be an alternative time-series analysis in the spirit of Nowak-Lehman et al. (2012). To the contrary, it is a totally different approach, incompatible with any of the choices we made, other than relying on the same source for per capita GDP and aid.

[8] According to their list of countries included in the analysis (provided in Table 8), the sample for the period 1970-2006 consists of 58, rather than 59, countries. Thus, one country is missing. We believe that the missing country is Guinea-Bissau, given that this country has complete time series for the period 1970-2006.

[9] These countries are Fiji, Jordan, Uganda, Suriname, and Zambia.

[10] These are Argentina, Republic of Congo, El Salvador, Honduras, Indonesia, Saudi Arabia, Syria, and Turkey.

Third, the estimated VECM is based on the assumption that aid per capita and GDP per capita (in logs) are cointegrated, i.e., that there is a (non-spurious) long-run relationship between the two variables. However, LMT do not report results of cointegration tests, which is again rather surprising. They simply *assume* the existence of co-integration (p.7).

Fourth, the estimated VECM incorporates the previously estimated long-run relationship between aid and income. That is, the long-run relationship between aid and income must be estimated before the parameters of this relationship can be used to construct the error correction term in the panel VECM. Surprisingly, LMT do not report the estimated cointegrating relationship, even though this is a key prerequisite for their analysis.

Fifth, the standard approach in testing for long-run causality in cointegrated systems is to test for weak exogeneity in a VECM specification. Weak exogeneity of one variable (in our case aid) in a system of two cointegrated variables implies that the aid variable has a long-run (causal) effect on income, while income has no long-run effect on *aid*, although income might have a short-run effect on aid. While such causality tests are performed as routine procedures in similar analyses, they are ignored in LMT.

Finally, LMT's conclusion of "a positive and statistically significant long-run effect of aid on income" is based exclusively on impulse responses—although the problems of constructing confidence intervals for impulse responses are well documented in the literature (see, e.g., Benkwitz et al. 2000, 2001; Lütkepohl, 2014).

We proceed to assess whether the statistical basis of the VAR model, in particular co-integration and weak exogeneity holds. Following LMT, we base our analysis on a panel VAR model that includes the log of GDP per capita, gdp_{it}, and the log of net aid transfers per capita, aid_{it}. Specifically, the model is as follows:

$$y_{it} = \mu_i + \sum_{j=1}^{p} A_j y_{it-j} + \varepsilon_{it}, \qquad (1)$$

where i and t are country and time subscripts, $y_{it} = (gdp_{it},\ aid_{it})'$ is a vector of endogenous variables, and μ_i is a vector of country-specific intercepts.

Contrary to LMT, we tested whether the variables included in y_{it} are integrated of order one, or $I(1)$, which is not rejected by panel unit root tests at conventional levels of significance (available on request). Assuming for the moment that that aid_{it} and gdp_{it} are cointegrated (see below), the panel VAR model in equation (1) can be transformed into a panel VECM, as given in equation (2):

$$\Delta y_{it} = \mu_i + \alpha\beta' y_{it-1} + \sum_{j=1}^{p-1} B_j \Delta y_{it-j} + \varepsilon_{it} \tag{2}$$

where Δy_{it} is the vector of log first differences $(\Delta gdp_{it},\ \Delta aid_{it})'$, $\alpha = (\alpha_1,\ \alpha_2)'$ is the vector of adjustment (or loading) coefficients, B_j is a matrix of short-run coefficients, and $\beta = (1,\ -\beta_{aid})'$ is the cointegrating vector for $y_{it} = (gdp_{it},\ aid_{it})'$ normalized on gdp_{it}. This normalization suggests the long-run relationship $gdp_{it} = \beta_{aid}\, aid_{it} + \varepsilon_{it}$.

The term $\beta' y_{it-1} = ec_{it-1}$ is the (lagged) error correction term, or cointegrating residual. It represents the error in, or deviation from, the equilibrium, and the adjustment coefficients (α_1 and α_2) capture how gdp_{it} and aid_{it} respond to deviations from the equilibrium. The Granger Representation Theorem (Engle and Granger, 1987) states that, if a vector y_{it} is cointegrated, at least one of the adjustment coefficients must be nonzero in the VEC representation (2).

If the adjustment coefficient (α_2) in the Δaid_{it} equation is zero, then aid_{it} is weakly exogenous for β. Hall and Milne (1994) show that weak exogeneity in a cointegrated system is equivalent to the notion of long-run non-causality. Thus, non-rejection of the null hypothesis of weak exogeneity of aid, $H_0: \alpha_2 = 0$, implies that GDP per capita has no causal effect in the long run on aid per capita; the long-run causality will run from aid to income if there is cointegration. If the null hypothesis $\alpha_1 = 0$ is not rejected, then GDP per capita is weakly exogenous with respect to the long-run parameters, implying that aid per capita has no long-run effect on GDP per capita; the long-run causality will run from income to aid if there is cointegration. If both null hypotheses are rejected, then the variables are jointly determined and long-run causality runs in both directions.

From the above discussion, it should be clear that aid_{it} and gdp_{it} must be cointegrated (such that ec_{it-1} is stationary). Otherwise, there is no long-run relationship, and the panel VECM is misspecified. Another important point is that the cointegrating parameters must be estimated before they can be used to construct the error correction term in the panel VECM.

We use the same data set as LMT, those taken in our study in Nowak-Lehmann et al. (2012). This data set is unbalanced in the sense that the number of time-series observations per country varies, while the methods used here, like that of LMT, require balanced panel data. To construct a balanced panel, we select all countries for which, after taking logs, complete time-series data on real GDP per capita and net aid transfers (relative to GDP) are available over the 1970-2006 period – the same sample period as in LMT. This yields a sample of 46 countries (contrary to LMT, dropping those with missing and/or negative net aid observations, as listed in the Appendix) and 37 time-series observations per country.[11]

We proceed as follows: First, we test whether there is a cointegrating or long-run relationship between the log of aid per capita and the log of GDP per capita of the form $gdp_{it} = \mu_i + \beta_{aid} aid_{it} + \varepsilon_{it}$. Then we estimate the long-run coefficient β_{aid} and use the estimated cointegrating coefficient $\hat{\beta}_{aid}$ to construct the error-correction term for the panel VECM. Finally, we use the panel VECM to examine the direction of long-run causality between aid and income and test for the robustness of our results.

[11] Since our original data set does not include real aid per capita, we construct this variable by multiplying "net aid transfer/GDP" by real GDP per capita.

The most commonly employed tests for panel cointegration are the residual-based cointegration tests of Kao (1999), Pedroni (1999, 2004), and Holly et al. (2010), the ECM-based cointegration tests of Westerlund (2007), and the maximum-likelihood-based panel cointegration approach developed by Larsson et al. (2001).

As shown in Table 2, the overwhelming majority of these tests are unable to reject the hypothesis that there is no long-run relationship of the form $gdp_{it} = \mu_i + \beta_{aid} aid_{it} + \varepsilon_{it}$ – despite the fact that these tests are known to have high power in panels of this dimensionality.[12] The one test that rejects the null hypothesis that r = 0 (no cointegration) is those by Larsson et al. (2001). However, this test also rejects the hypothesis of one cointegrating vector (r = 1) and thus suggests that there are two cointegrating vectors. This, would however imply that gdp_{it} and aid_{it} are stationary (which is not the case according to our unit root test results) and indicates that the results of this test are less reliable in our case than the results of the other panel cointegration tests. In fact, it is well known that likelihood-based panel cointegration methods suffer from substantial size distortions when the time dimension is comparably small, as is the case here (Wagner and Hlouskova, 2010). Overall, the results in Table 2 provide little evidence of a long-run relationship between aid per capita and income, contradicting the conclusion of LMT that aid per capita has a positive and statistically significant long-run effect on income.

Table 2 *Panel cointegration tests*

Kao (1999)		
DF *rho*-statistic	2.80	
DF *t*-statistic	1.51	
ADF *t*-statistic	-0.94	
DF^* *rho*-statistic	0.71	
DF^* *t*-statistic	-0.28	
Pedroni (1999, 2004)	Panel statistics	Group mean statistics
Variance ratio	-3.42	
PP *rho*-statistics	2.55	2.30
PP *t*-statistics	1.08	0.95

[12] We also run the same tests using the aid/GDP variable where we find evidence for cointegration; this provides further support that the use of the aid/capita variable used by LMT is not only conceptually but also econometrically problematic, compared to the usually used aid/GDP variable.

ADF *t*-statistics	0.69	0.28
Larsson et al. (2001)	Cointegration rank	
	$r = 0$	$r = 1$
Standardized panel trace statistics	6.63***	7.46***
Westerlund (2007)	Panel statistics	Group mean statistics
τ-statistics (P_τ, G_τ)	3.54	5.20
	(0.84)	(1.00)
α-statistics (P_α, G_α)	3.82	4.41
	(0.91)	(1.00)
Holly et al. (2010)		
CIPS statistic	-1.49	

Notes: For the Kao (1999) and the Pedroni (1999) tests, the number of lags was determined by the Schwarz criterion with a maximum of four lags. To avoid overparametrization and the resulting loss of power, only one lag was included in the Larsson et al. (2001), Westerlund (2007), and Holly et al. (2010) tests. The relevant 5% critical value for the CIPS statistic is -2.11. All other test statistics are asymptotically normally distributed. The variance ratio and the panel trace tests test are right-sided, while the other tests are left-sided. To account for potential cross-sectional dependence, we used the bootstrap approach of Westerlund (2007). Bootstrap *p*-values (based on 400 replications) in parentheses. We used the mean and variance of the asymptotic trace statistic tabulated by Breitung (2005) to construct the standardized panel trace statistics for our model. *** indicate a rejection of the null hypothesis of *r* cointegrating relations at the 1% level.

In order to further investigate the methods used by LMT, we nevertheless proceed below under the assumption that aid and income are cointegrated by estimating the long-run relationship between these variables, using the panel DOLS estimator suggested by Kao and Chiang (2000). The DOLS estimator is superconsistent, asymptotically unbiased, and normally distributed, even in the presence of endogenous regressors (provided that the variables are cointegrated). What is more, recent Monte Carlo evidence by Wagner and Hlouskova (2010) suggests that Kao and Chiang’s panel DOLS estimator outperforms other cointegration estimators such as Phillips and Moon’s (1999) FMOLS estimator and Breitung’s (2005) two-step estimator. We estimate $\hat{\beta}_{aid} = -0.0738$; the *t*-statistic for this estimate being –3.68. Again, this result is not supportive of the claim that aid has a positive and statistically significant long-run effect on income.

Of course, we must be cautious in concluding from this result that aid has a negative long-run effect on income. Although the existence of cointegration implies long-run Granger causality in at least one direction (Granger, 1988), cointegration says nothing about the direction of the causal relationship between the variables. A statistically significant cointegrating relationship between gdp_{it} and aid_{it} does therefore not necessarily imply that, in the long run, changes in aid cause changes in income. Causality may well run in the opposite direction, from gdp_{it} to aid_{it}, or in both directions.[13]

To test the direction of long-run causality, we follow the common practice in the applied panel cointegration literature and employ a two-step procedure, as described above. In the first step, the (DOLS) estimate of the long-run coefficient is used to construct the error correction term $ec_{it} = gdp_{it} - [\hat{\mu}_i - 0.0738\ aid_{it}]$. In the second step, the lagged error correction term is included into the VECM to test for weak exogeneity. Given that all variables in the model, including ec_{it-1}, are stationary (because the level variables are cointegrated by assumption), a conventional likelihood ratio chi-square test can be used to test the null hypothesis of weak exogeneity $H_0 : \alpha_{1,2} = 0$.

We determine the number of lags in the model by using the Schwarz information criterion and the Hannan-Quinn criterion with a maximum of four lags. Given that both criteria suggest a lag length of one, we estimate the VECM with one lag of the differenced endogenous variables.

Table 3 reports the results. The null hypothesis of weak exogeneity is rejected for both variables, implying that long-run causality is bidirectional. Given that the long-run coefficient on aid is negative, it is reasonable to conclude that, in the long-run, an increase in aid reduces income and that, in turn, more growth leads to less aid (provided that the variables are cointegrated).

Table 3. *Weak exogeneity tests / long-run causality tests*

Variable (Adjustment coefficient)	gdp_{it} (α_1)	aid_{it} (α_2)
$\chi^2(1)$ (*p*-values)	41.95 (0.000)	15.89 (0.000)

Notes: The number of degrees of freedom υ in the standard $\chi^2(\upsilon)$ tests corresponds to the number of zero restrictions. The number of lags was set to 1, based on the Schwarz information criterion and the Hannan-Quinn criterion.

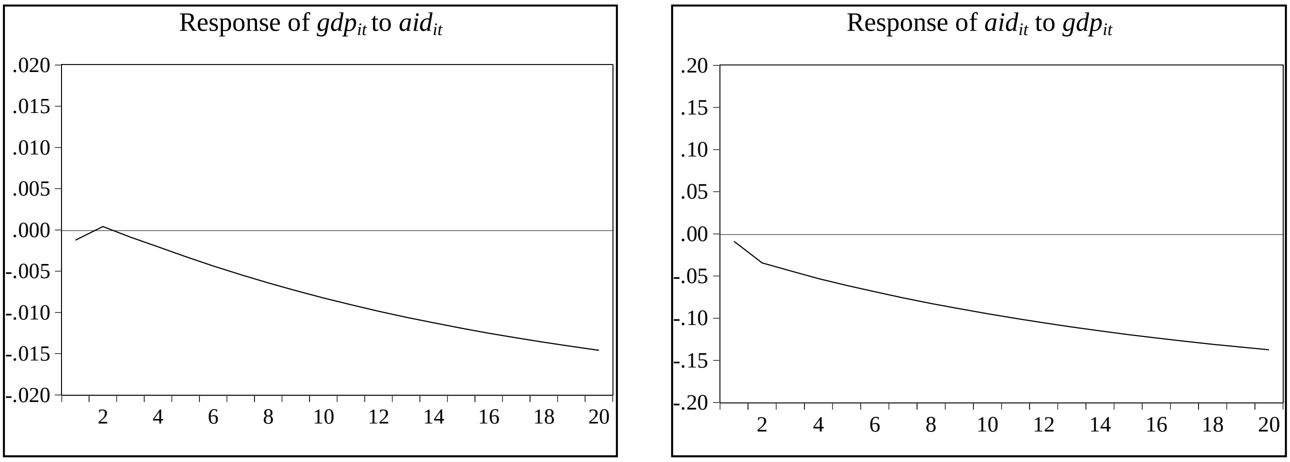
To check the robustness of this conclusion, we present generalized impulse response functions based on the one-lag panel VECM over a 20-year horizon.[14] As can be seen in the left

[13] However, recall that cointegration could not be established (see Table 2).

[14] We do not give confidence bands for our impulse responses because there are a number of problems related to the standard confidence bands which are often reported in the literature, as mentioned above.

panel of Figure 1, income first slightly decreases, then slightly increases, and finally decreases in the long-run due to a one standard deviation shock in aid. The right panel shows that aid gradually and permanently decreases in response to a one-standard-deviation shock in income. Thus, the impulse response functions are consistent with the weak exogeneity tests reported above.

Figure 1. *Impulse responses*



Finally, we perform a standard panel Granger-causality test based on a levels VAR regression with fixed effects, as given by equation (1). While it is well known from the time-series literature that in general the asymptotic distributions of the Wald (or likelihood ratio) test for Granger causality in levels VARs with integrated variables are nonstandard (see, e.g., Toda and Phillips 1993), Lütkepohl and Reimers (1992) show that the conventional Wald test for a bivariate cointegrated VAR model is asymptotically distributed as chi-square and therefore valid as a test for Granger causality.

We report the *p*-values of the Granger causality chi-square statistics using two and four lags (based on the Schwarz criterion and the Hannan-Quinn criterion, respectively) in Table 4. As can be seen from the first row, the null hypothesis of no Granger causality from aid_{it} to gdp_{it} is not rejected for two lags. Accordingly, this result implies that, although the sum of the coefficients on lagged aid is positive, changes in aid do not significantly cause changes in income. Row (2), in contrast, shows that the Granger causality test rejects the null hypothesis that four lags of aid_{it} do not help predict *gdp*$_{it}$ at the 5% level and that the sum of the coefficients on lagged aid is negative in the income equation. According to this result, an increase in aid leads to a decrease in income, consistent with the results above.

Finally, the Granger causality tests in rows (3) and (4) reject the null hypothesis that two/four lags of gdp_{it} do not help predict aid_{it} at the 1% level with a negative sum of the coefficients on GDP per capita. Accordingly, more income leads to less aid, as expected.

In summary, we find no evidence that aid per capita has a positive and statistically significant long-run effect on income. LMT's finding that "[w]hen the same dataset […] is evaluated using a Panel VAR model, […] a positive and statistically significant long-run effect of aid on growth emerges," is not robust to more rigorous statistical analysis.

Table 4. *Panel VAR causality tests*

Null hypothesis	Lags	*p*-value of the Granger causality chi-square statistic	Sum of the coefficients of the causal variable
(1) aid_{it} does not cause gdp_{it} Dependent variable: gdp_{it}	2	0.155	0.002
(2) aid_{it} does not cause gdp_{it} Dependent variable: gdp_{it}	4	0.038	-0.001
(3) gdp_{it} does not cause aid_{it} Dependent variable: aid_{it}	2	0.000	-0.157
(4) gdp_{it} does not cause aid_{it} Dependent variable: aid_{it}	4	0.001	-0.173

Notes: This table reports the *p*-values of Granger-causality VAR tests. The null hypothesis is that two/four lags of the series of aid_{it} (gdp_{it}) not help predict the series of gdp_{it} (aid_{it}).

4. Conclusions

In this comment we have shown that our results in Nowak-Lehmann et al. (2012) are robust to how we deal with missing data. Second, we have clarified why we think that our time series econometric approach is substantively justified, based on the required assessments of co-integration and Granger-causality tests. Third, we have shown that the econometric approach used by LMT lacks theoretical and econometric justification. Their analysis suffers from missing data problems, uses variables for aid and income where there is no clear evidence of cointegration, and where causality assessments do not support the assertions they made. We also explain why their choice of relating income per capita to aid per capita produces an upward bias to the relation between aid and income. This is because shocks to population size would lead to a positive effect of aid per capita on income per capita even if aid is unrelated to income. Fourth, LMT selectively emphasize those of their results that show a significant positive relation between aid and income, even though significance is only achieved by dropping the earlier years from the sample (see their figure 5a). Finally, LMT selectively cite those previous papers that find a positive impact of aid on growth, neglecting others, even though methodological problems apply to both groups of papers equally.[15] Overall, we therefore conclude that it will always be possible to show a positive impact of aid on income if a sufficient variety of data and models are tried; such evidence fails to be robust however. Overall, we conclude that appropriate time series approaches to the aid-income relationship suggest that there is no robust evidence of a positive impact of aid on income.

We would like to end by pointing to where we agree with LMT, and to reiterate where we disagree. LMT conclude their study with observing that a positive effect of aid on growth is consistent with "results from the aid-growth literature more generally." This conclusion represents are rather selective reading of the literature. While we are aware that a number of studies reports positive effects of aid on growth and at least equally large number finds insignificant effects (see Doucouliagos and Paldam 2013 for a summary). While LMT are critical with respect to studies that report no positive and significant effect on growth, they seem to lack such skepticism towards studies finding the effect to be positive. Clearly, the identifying assumptions in recent analyses can as easily be dismissed as LMT seem to be willing to dismiss ours.[16] We prefer to be skeptical with

[15] LMT also lack a balanced way of discussing the papers they do cite. For example LMT refer to Gillanders' (2011) VAR model as being in favor of a positive relation between aid and income, while they refer to Kang et al. (2012) as "add exchange rates to their VAR" (p. 7), without referring to the results, that show aid to increase growth in half of their sample, but to reduce growth in the other half.

[16] Clemens et al. (2012) use no instruments, but rely on first-differencing and lagging aid. To the extent that donors give more aid to reform-oriented recipients, and such reforms increase growth, the positive effect of aid on growth that they

respect to both types of studies, those showing positive and insignificant effects, also stressing the limitations of our own approach. This leads us to areas of agreement with LMT.

First, we agree that the absence of a significant effect of aid on growth in a regression framework does not prove the ineffectiveness of aid.[17] Statistically, the inability to reject a hypothesis does not imply the truth of the alternative hypothesis. What is more, we are aware of no study (including ours) that was able to address potential simultaneity of the aid-growth relationship in a bullet-proof way. In addition to methodological difficulties, low data quality might prevent us from finding significantly positive effects. Maybe the effect of aid on growth is too indirect to be measurable with the methods at hand, and it is well known that few variables turn out to be robustly significant in growth regressions across studies.[18]

Donors pursue a multitude of objectives when granting aid, with growth being just one of them. Political objectives like marshalling support in international organizations or fighting terrorism, commercial objectives, as well as more specific developmental outcomes (such as education or health) might be more important than growth for the donors' decision to grant aid, and there is some evidence that aid can be effective in achieving these goals (Dreher et al. 2008a, 2008b). To the extent that these varied donor motives affect the impact of aid on one of these motives, economic growth, it might not be surprising that we are unable to find a robust positive impact in a growth regression (see Dreher et al. 2014). In summary, while we believe that aid can be effective, we do not believe there is robust empirical evidence for that in a growth-regression framework.

---

report is not causal. Brückner (2013) relies on rainfall and international commodity price shocks as instruments, where the exclusion restriction could easily be violated.

[17] We made this more explicit in earlier versions of Nowak-Lehmann et al. (2012), where our last paragraph starts with "This is not to say, however, that aid cannot have important effects on GDP levels and growth if focused exclusively on maximizing these effects." However, we were advised by the editor of the Canadian Journal of Economics to remove this paragraph ("There is no need to be apologetic about the result that foreign aid does not have the hoped-for result of lifting per-capita income. If foreign aid in its current form isn't working, don't be afraid of saying so." Decision letter from the CJE, November 26, 2010).

[18] However, Sala-i-Martin (1997) reports that some variables indeed are robust correlates of growth.

Appendix: Countries in the Sample, 1970-2006

Algeria, Belize, Benin, Bolivia, Botswana, Burundi, Cameroon, Central African Republic, Chad, Colombia, Democratic Republic of Congo, Cote d'Ivoire, Ecuador, Egypt, Gambia, Ghana, Guatemala, Guinea-Bissau, Guyana, Haiti, India, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Morocco, Nepal, Nicaragua, Niger, Nigeria, Pakistan, Panama, Peru, Philippines, Rwanda, Senegal, Seychelles, Sierra Leone, Sri Lanka, Sudan, Togo, Tunisia, and Uruguay

References

Benkwitz, A., Neumann, M.H.. Lütkepohl, H., 2000. Problems related to confidence intervals for impulse responses of autoregressive processes. *Econometric Reviews*, 19(1), 69-103.

Benkwitz, A., Lütkepohl, H., Wolters, J., 2001. Comparison of bootstrap confidence intervals for impulse responses of German monetary systems. *Macroeconomic Dynamics*, 5(1), 81-100.

Breitung, J., 2005. A parametric approach to the estimation of cointegrating vectors in panel data. *Econometric Reviews*, 24(2), 151-173.

Brückner, Markus, 2011, On the simultaneity problem in the aid and growth debate, Journal of Applied Econometrics 28: 126-150.

Doucouliagos; H. and Paldam; M. (2013) The robust result in meta-analysis of aid effectiveness: A response to Mekasha and Tarp. *Journal of Development Studies,* 49(4): 584-587.

Dreher, Axel, Rainer Thiele and Peter Nunnenkamp, 2008a, Does Aid for Education Educate Children? Evidence from Panel Data, *World Bank Economic Review* 22, 2: 291-314.

Dreher, Axel, Rainer Thiele and Peter Nunnenkamp, 2008b, Does US Aid Buy UN General Assembly Votes? A Disaggregated Analysis, *Public Choice* 136, 1: 139-164.

Dreher, Axel, Vera Eichenauer and Kai Gehring, 2014, Geopolitics, Aid and Growth, CEPR Discussion Paper No. 9904.

Easterly, William, R. Levine and D. Roodman. 2004. New data, new doubts: A Comment on Burnside and Dollar’s 'Aid, Policies, and Growth' (2000), *American Economic Review* 94, 3: 774-780.

Engle, R.E., Granger, C.W.J., 1987. Cointegration and error-correction: representation, estimation, and testing. *Econometrica*, 55(2), 251-276.

Gillanders, R. (2011). The effects of foreign aid in Sub-Saharan Africa (No. 11/16). Working Paper Series, UCD Centre for Economic Research.

Goldbach, C. (2013). The link between development aid and economic development in Africa. Evidence from panel data analysis. MA thesis in "International Economics". University of Göttingen.Mimeo. Available upon request.

Granger, C.W.J., 1988. Some recent developments in a concept of causality. *Journal of Econometrics*, 39(1-2), 199-211.

Haavelmo, T. (1944). The Probability Approach in Econometrics. *Econometrica*, 12, supplement: 1-115.

Hall, S.G., Milne, A., 1994. The relevance of P-star analysis to UK monetary policy. *Economic Journal*, 104(424), 597-604.

Holly, S., Pesaran, M.H., Yamagata, T., 2010. A spatio-temporal model of house prices in the US. *Journal of Econometrics*, 158(1), 160-173.

Kang, J.S., Prati, A. and Rebucci, A. (2012). Aid, Exports, and Growth: a Time-Series Perspective on the Dutch Disease Hypothesis. *Review of Economics & Institutions* 3, 2.

Kao, C., 1999. Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, 90(1), 1-44.

Kao, C., Chiang., M., (2000). On the estimation and inference of a cointegrated regression in panel data. *Advances in Econometrics*, 15, 179-222.

Larsson, R., Lyhagen, J., Löthegren, M., 2001. Likelihood-based cointegration tests in heterogeneous panels. *Econometrics Journal*, 4(1), 109-142.

Lof, M., Mekasha, T.J. and Tarp, F., 2014. Aid and Income. Another Time-series Perspective. *World Development*, forthcoming.

Lütkepohl, H., Reimers, H.–E., (1992). Granger-causality in cointegrated VAR processes. The case of the term structure. *Economics Letters*, 40(3), 263-268.

Lütkepohl, H., Staszewska-Bystrova, A., Winker, P., 2014. Comparison of methods for constructing joint confidence bands for impulse response functions. *International Journal of Forecasting*, forthcoming.

Moyo, Dambisa. (2009). Dead aid: Why aid is not working and how there is a better way for Africa. Macmillan.

Nowak-Lehmann, F., Dreher, A., Herzer, D., Klasen, S., Martínez-Zarzoso, I., 2012. Does aid really raise per capita income? A time series perspective. *Canadian Journal of Economics*, 45(1), 288-313.

Pedroni, P., 1999. Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* 61, 653-670.

Pedroni, P., 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20, 597-625.

Phillips, P.C.B., Moon, H.R., 1999. Linear regression limit theory for nonstationary panel data, *Econometrica*, 67(5), 1057–1111.

Sachs, Jeffrey. 2006. The end of poverty: economic possibilities for our time. Penguin.

Sala-i-Martin, Xavier, 1997, I just ran two million regressions, *American Economic Review* 87, 2: 178-183.

Toda, H.Y., Phillips, P.C.B., 1993. Vector autoregressions and causality. *Econometrica*, 61(6), 1367-1393.

Wagner, M., Hlouskova, J., 2010. The performance of panel cointegration methods: results from a large scale simulation study. *Econometric Reviews*, 29(2), 182-223.

Werker, Eric, Faisal Z. Ahmed, and Charles Cohen. 2009. How is foreign aid spent? Evidence from a natural experiment. *American Economic Journal: Macroeconomics* 225-244.

Westerlund, J., 2007. Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, 69(6), 709-748.