Centre International pour la Fiscalité et le Développement
International Centre for Tax and Development

ICTD

Working Paper 20

Foreign Aid And Domestic Taxation: Multiple Sources, One Conclusion

Paul Clist

September 2014

www.ictd.ac

UKaid from the Department for International Development

Norad

ICTD Working Paper 20

Foreign Aid And Domestic Taxation: Multiple Sources, One Conclusion

Paul Clist

September 2014

Foreign Aid and Domestic Taxation: Multiple Sources, One Conclusion
Paul Clist
ICTD Working Paper 20
First published by the Institute of Development Studies in September 2014
© Institute of Development Studies 2014
ISBN: 978-1-78118-190-4

A catalogue record for this publication is available from the British Library.
All rights reserved. Reproduction, copy, transmission, or translation of any part of this publication may be made only under the following conditions:

– with the prior permission of the publisher; or
- with a licence from the Copyright Licensing Agency Ltd., 90 Tottenham Court Road, London W1P 9HE, UK, or from another national licensing agency; or
- under the terms set out below.

This publication is copyright, but may be reproduced by any method without fee for teaching or nonprofit purposes, but not for resale. Formal permission is required for all such uses, but normally will be granted immediately. For copying in any other circumstances, or for reuse in other publications, or for translation or adaptation, prior written permission must be obtained from the publisher and a fee may be payable.

Available from:
The International Centre for Tax and Development
at the Institute of Development Studies,
Brighton BN1 9RE, UK
Tel: +44 (0) 1273 606261 Fax: +44 (0) 1273 621202
E-mail: info@ictd.ac.uk
Web: www.ictd/en/publications

Foreign Aid And Domestic Taxation: Multiple Sources, One Conclusion

Paul Clist

Summary

There are genuine concerns that foreign aid may crowd out domestic tax revenue. In the short run this would have negative consequences for the recipient government's revenue, and over a longer period could corrode governance through breaking the social contract. In recent years, two papers have presented empirical results that suggest while aid loans are free from such concerns, aid grants do crowd out tax revenue. Previous research showed that the results from the first paper, Gupta et al. (2004), did not survive the inclusion of more recent data or a minimal lag on aid variables (a simple way of reducing concerns of endogeneity). This article deals with the second contribution, Benedek et al. (2012), and finds that the results cannot be replicated. Furthermore, they suffer from serious problems resulting from a dependent variable comprised of several incompatible data sources and definitions. A variety of econometric techniques are used, including new data, with the weight of evidence pointing to a modest but positive effect from foreign aid on domestic tax revenue. Fears over a negative effect for aid grants appear unwarranted, and are accounted for by the inappropriate use of data or endogeneity concerns.

Keywords: development aid; tax; fiscal response; replication; MIMIC model.

Paul Clist is a lecturer in Development Economics in the University of East Anglia’s School of International Development. He has published on the topics of aid allocation, aid modalities and payment by results in international development.

Contents

	Summary	3
	Acknowledgements	5
	Acronyms	5
	Introduction	**6**
1	**Replication attempt**	**6**
	1.1 Replicating Benedek et al. (2012)	6
	1.2 A negative effect from aid?	8
2	**Multiple sources**	**9**
3	**Four options**	**10**
	3.1 Option 1 Treat variables as interchangeable	11
	3.2 Option 2 Different regressions	12
	3.3 Option 3 MIMIC	13
	3.4 Option 4 New dataset	15
4	**Discussion and conclusion**	**16**
	Appendices	
	Appendix 1 System break	18
	Appendix 2 Data sources	19
	References	**20**

Tables

Table 1	Determinants of tax revenue 1980-2009	7
Table 2	Turning points	8
Table 3	Estimating the effects on tax revenue by source, 1980-2009	10
Table 4	P statistics to test for agreement between sources	12
Table 5	Preferred specification 1980-2011, by data source	13
Table 6	MIMIC model, 1980-2011	14
Table 7	ICTD dataset, 1980-2009	16
Table A1	System break	18

Figures

Figure 1	The MIMIC model	14

Acknowledgements

I am pleased to acknowledge the financial support of the International Centre for Taxation and Development (ICTD). The paper has been improved by the willingness of Dora Benedek and colleagues at the IMF to share data and answer some questions regarding data. I am also grateful to Oliver Morrissey, Paddy Carter and an anonymous reviewer for their comments. Any errors remain my own.

Acronyms

GFS	Government Finance Statistics
GMM	Generalised Method of Moments
GRD	Government Revenue Dataset
ICTD	International Centre for Tax and Development
IMF	International Monetary Fund
MIMIC	Multiple indicators multiple cause
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Squares
WEO	World Economic Outlook

Introduction

The debate over whether, and in what circumstances, foreign aid displaces domestic tax revenue has a long history. The debate is clearly important: the immediate effect would be to reduce the effective value of aid flows, as they would not be additional resources for the recipient government but rather crowd out tax revenue. A recent quote from the House of Commons International Development Committee (2013: 4) illustrates the point well: '[w]e cannot expect the people in the UK to pay taxes to improve education and health in Pakistan if the Pakistan elite is not paying income tax'. An even greater fear relates to a potential pernicious effect of aid over a longer time horizon, where it could conceivably undermine governance through fracturing the social contract. Deaton (2013: 295) argues that '[o]ne of the strongest arguments against large aid flows is that they undermine these constraints, removing the need to raise money with consent and in the limit turning what should be beneficial political institutions into toxic ones'.

This article focuses specifically on the link between aid and tax revenue, where it is argued that aid has different effects depending on whether it is given as grants or loans. Schmidt (1964) contains the majority of the theoretical reasons regarding whether there should be a differential effect from aid grants and loans, and since then the question has predominately been recognised as empirical. Gupta et al. (2004) entered the debate with a strong claim: aid grants depress domestic tax revenue but aid loans do not. The empirical basis for this claim was disputed by Clist and Morrissey (2011) and Carter (2010, 2013). The former extended the dataset and found that the negative relationship is not present if: a) the period 1985-2005 is examined, or b) aid is included in specifications with a more reasonable lag. The latter found the results to be fragile to sample, and suggested that endogeneity may be driving the negative relationship between aid grants and tax revenue. Endogeneity is a concern, as the composition and volume of aid is not exogenous to factors such as per capita income or aid. Benedek et al. (2012) can be seen as a follow-up to Gupta et al. (2004), and responds to criticism with new methods and data but ultimately presents very similar findings.[1] Here, I examine the evidence base for these claims, which influence important decisions about the correct composition of aid, and feed into the wider debate regarding the effect of aid on governance.

The article proceeds as follows. Section 1 attempts, and fails, to replicate Benedek et al.'s (2012) results using the provided and described data. An advanced discussion of the suitability of Generalised Method of Moments (GMM) techniques is found elsewhere (Carter 2010, 2013) but the data is found to have more basic weaknesses. These weaknesses, stemming mainly from the use of multiple different sources, are explored in Section 2. A variety of solutions are examined in Section 3, and Section 4 concludes.

1 Replication attempt

1.1 Replicating Benedek et al. (2012)

I have attempted to replicate Benedek et al.'s (2012) results using the provided dependent variable and listed sources. There are several discrepancies. Limiting the sample to the countries in Appendix 1 of Benedek et al. (2012), there are six countries which are apparently included in the regression results but for whom no data exists for the dependent

1 Both are written by IMF-based researchers, and they share one author.

variable (four countries are not in the dataset, two exist but have no data). Clearly, both the provided dataset and the article cannot both be fully accurate, stymieing attempts to replicate the results. Comparisons between summary statistics in Benedek et al. (2012) and those from the reported sources reveal several other discrepancies. For example, the paper reports a mean value of 18.75 for trade openness, (exports+imports)/GDP, but a newly constructed dataset using the reported data sources gives a mean of 54.

Table 1 Determinants of tax revenue 1980-2009

	Model 1		**Model 2**		**Model 1**	**Model 2**
With source dummies:	**No**	**Yes**	**No**	**Yes**	**From Benedek et al.**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Aid	0.005 (1.19)	0.006 (1.49)			-0.007* (0.004)	
Aid squared	-0.000 (-1.27)	-0.000 (-1.49)			0.000 (0.000)	
Loans			0.011** (2.33)	0.010** (2.16)		0.000 (0.004)
Loans squared			-0.000* (-1.72)	-0.000 (-1.64)		-0.000 (0.000)
Grants			0.003 (0.50)	0.005 (0.82)		-0.006* (0.003)
Grants squared			-0.000 (-1.03)	-0.000 (-1.28)		0.000 (0.000)
Agriculture	-0.006* (-1.67)	-0.006 (-1.63)	-0.007* (-1.89)	-0.007* (-1.88)	-0.008*** (0.002)	-0.003 (0.003)
Industry	-0.007 (-1.19)	-0.006 (-1.29)	-0.002 (-0.64)	-0.003 (-0.76)	0.000 (0.003)	0.000 (0.000)
GDP pc (Logged)	0.248** (2.09)	0.230* (1.88)	0.209 (1.62)	0.145 (1.23)	0.304*** (0.122)	0.305*** (0.122)
Trade openness	0.003*** (2.90)	0.003*** (3.11)	0.003** (2.30)	0.003** (2.61)	-0.002** (0.000)	-0.002 (0.001)
WEO Gen		-0.350 (-1.26)		0.081 (0.71)		
GFS 2001 Cen		. .		0.155 (1.60)		
GFS 2001 Cen		-0.275 (-1.32)		0.088 (1.07)		
GFS 1986 Cen		-0.377* (-1.71)		. .		
GFS 2001 Bud		-0.561** (-2.27)		-0.192 (-1.58)		
African R M		-0.265 (-0.88)		0.183 (1.28)		
Overall R squared	0.28	0.29	0.32	0.36	Not reported	
Between R squared	0.28	0.27	0.39	0.41	Not reported	
Observations	2174	2174	1968	1968	2589	2589
Countries	99	99	97	97	118	118

Note: The dependent variable is logged total tax revenue divided by GDP. Year dummies and a constant are included but not reported. Estimation is by Ordinary Least Squares (OLS) using country fixed effects, with clustered standard errors. ***, ** and * denote 1%, 5% and 10% significance respectively. The source dummies are relative to OECD central government data, where Cen stands for Central, Gen for General and Bud for Budgetary. 'African R M' stands for *African Revenue Mobilisation*. GFS denotes the IMF *Government Finance Statistics*, and WEO the *World Economic Outlook*. 2001 and 1986 refer to different vintages of IMF data. Columns (5) and (6) are taken from Benedek et al. (2012: 13) Table 1, columns (1) and (4), where the estimator is also OLS with country fixed effects.

While an accurate replication has proved impossible, the regression results from the attempted replication are found in columns (1) and (3) of Table 1. For ease of reference, Benedek et al. (2012) results are reproduced in columns (5) and (6). These new results concur that tax revenue has a negative association with agriculture, and a positive association with income per capita. However, the coefficients reported in (1) and (3) differ in their sign on several key variables: Aid, Aid Squared, Industry and Trade Openness. The least important of these differences relates to industry: the coefficient is insignificant in all specifications. The coefficient for trade openness is positive in Table 1, but was found to be negative by Benedek et al. (2012). Despite apparently using the same sources, there seems to be quite a large difference in the two variables, as shown by comparing summary statistics. Logically, we would expect higher trade openness to mean higher tax revenue as a percentage of GDP as import taxes are relatively easy to collect, and so the results here are more intuitively plausible. In later analysis I follow Clist and Morrissey (2011), who disaggregate 'trade' into imports and exports as they logically have different effects.

1.2 A negative effect from aid?

Turning to the coefficients of interest, there is a large disparity between the coefficients found using the new dataset and those reported by Benedek et al. (2012). They find that aid has a negative effect on tax revenue but has *increasing* returns. They find a similar pattern for aid grants, but a positive effect for loans with *decreasing* returns. By contrast, I find the pattern of a positive effect with diminishing returns for all types of aid: total, grants and loans. Benedek et al. (2012) conclude that aid (especially in the form of grants) has a negative association with tax revenue. They acknowledge but do not emphasise that their results imply that a higher level of grants actually has a positive effect on domestic tax revenue collection.

The results here are more in line with expectation (at least regarding non-linearity): aid's positive effect on tax revenue diminishes as it becomes increasingly large relative to GDP. Table 2 reports the implied turning points, taken from Table 1. The turning point for total aid is calculated to be 65 per cent of GDP, meaning that aid can be said to have a negative effect on 0.3 per cent of the sample and a positive effect for the other 99.7 per cent. Grants are found to be more negative than loans when using this metric, but can still only be said to have a negative effect on domestic tax revenue in 1.8 per cent of the sample.

Table 2 Turning points

Variable	Coefficient	Coefficient squared	Turning point	Mean	% of sample above turning point
Aid/GDP	0.005	-0.00008	65%	8.2%	0.3%
Loans/GDP	0.011	-0.0003	39%	2.6%	0.1%
Grants/GDP	0.003	-0.0001	26%	5.6%	1.8%

Note: Turning points are calculated using the coefficients in columns 1 and 3 in *Table 1*. The final two columns use the sample years and recipients.

The results from these regressions do not support the conclusion that aid, in whatever form, is negatively associated with domestic tax revenue. It is not clear how the sample or data used here varies from Benedek et al. (2012), as I use their reported sources for independent variables and their provided dataset for the dependent variable, but the difference is meaningful. Perhaps the most important point to notice from the reported results is that the estimated coefficients for total aid and aid grants are not significant. Of the six relevant coefficients in columns (5) and (6) only two are significant, and only then at the 10 per cent level. Indeed, it appears that aid is a relatively small factor in tax revenue. This is consistent with other results such as Gupta (2007); he focuses on determinants of tax revenue and finds aid has a small but significant positive effect.

2 Multiple sources

Clist and Morrissey (2011) reported a system break when extending the dataset of Gupta et al. (2004) from 1970-2000 to 1970-2005. This time effect implied that while Gupta et al. (2004) were right to report a negative relationship, this effect was only relevant in the period 1970-1985, with a positive relationship being a better characterisation of the period 1986-2005. Using the reported data sources I cannot replicate the weak negative association found by Benedek et al. (2012), nor the system break of Clist and Morrissey (2011), with the new data and time period. Regressions that augment the independent variables with an interaction term (where aid variables are multiplied by a dummy for post-1984) lack extra explanatory power as the augmented variables are insignificant, as shown in Table A1 in Appendix 1. A key element of the time trend may actually have been that using two different data sources as pre-1990 data (termed 'historical') was not comparable with post-1990 data,[2] but was treated as such. While there appears to be no systematic time trend effect in the data, there are several different sources used.

While Benedek et al. (2012) treat the dependent variable as a single variable, they acknowledge that there are actually several sources. Each of these has their own data definition and methodology, and so the data is not strictly comparable. The largest difference between sources is perhaps that they explicitly measure different things: IMF GFS data measures tax revenue, while the *World Economic Outlook* data (the largest single source, Table 3) relates to *all government revenue*.[3] The definition of the WEO variable implies that it includes aid, making it especially problematic. In addition, three different types of data are used: general, central and budgetary definitions are treated as the same. The difference in source is compounded by differences in coverage between datasets. The OECD data covers higher-income countries, and so the two samples overlap for only three countries: Chile, Mexico and Turkey. In the majority of cases, sources are inconsistent by recipient: on average countries are represented by about 1.8 sources over the 30-year period.

Does the difference in sources matter? Columns (2) and (4) of Table 1 report the regression results from the two models proposed by Benedek et al. (2012) augmented by source dummies. There is clearly some effect of the source: the GFS 2001 central government figures appear to be relatively high, and GFS 2001 budgetary figures are relatively low. These are significant effects, with the latter coefficient much larger than the coefficient on aid loans, for example. However, this table does not represent a particularly elaborate test. If different sources were consistent in their differences, then the problem is easily rectified using source dummies, and the coefficients would be stable.

Table 3 presents results from a more elaborate test, which allows the estimated relationship between independent variables and dependent variables to differ by source. If the relationship is robust and the source does not differ systematically we would expect stable parameter estimates. Of the seven sources, we might expect variation for those with few

2 The IMF (2001: 3-4) states that: 'The methodology for compiling government finance statistics described in this [2001] manual differs substantially from the methodology of the 1986 GFS Manual'.

3 Personal correspondence and the associated dataset state that OECD data was used, but the paper itself mentions only GFS data and consultation reports. It is not clear which was actually used given that even the cumulative total of sources does not reach the number of observations reported by Benedek et al. (2012).

observations or countries, but the results are quite different across the sources. Benedek et al. (2012) report that aid has a negative effect with increasing returns to scale. This pattern is only found when using GFS 2001 data; indeed the two largest datasets find opposite coefficient signs. Most variables are found to be significant in at least one regression or other, but this is not a sign of consistency. Trade Openness stands out as the only variable not to be found significant with both positive and negative signs. To summarise, Table 3 shows that coefficient estimates are not robust to the various sources used, and questions the robustness of results in Benedek et al. (2012).

Table 3 Estimating the effects on tax revenue by source, 1980-2009

Source of the dependent variable:	**OECD Cen**	**WEO Gen**	**GFS 2001 Gen**	**GFS 2001 Cen**	**GFS 1986 Cen**	**GFS 2001 Bud**	**African R M**
Aid	0.095 (0.28)	0.001 (0.30)	-0.032** (-3.71)	-0.007 (-1.65)	0.009 (0.97)	-0.010 (-0.94)	0.018* (2.69)
Aid squared	-0.012 (-0.06)	-0.000 (-0.22)	0.001** (3.11)	0.000 (1.60)	-0.000 (-1.13)	0.000 (0.97)	-0.000* (-2.71)
Agriculture	-0.038 (-3.19)	-0.000 (-0.11)	0.021* (2.16)	-0.004 (-0.56)	-0.018* (-3.13)	0.007 (1.04)	-0.011** (-2.92)
Industry	0.010 (2.32)	0.001 (0.19)	-0.008 (-1.40)	-0.010* (-2.41)	-0.002 (-0.43)	0.000 (0.01)	0.010* (2.56)
GDP pc (Logged)	-0.091 (-0.28)	0.242* (2.65)	-0.922*** (-7.68)	0.228 (1.08)	0.087 (0.39)	0.631 (1.89)	0.320* (2.28)
Trade openness	-0.006 (-0.74)	0.003** (3.45)	0.004 (1.87)	0.001 (0.28)	0.001 (0.49)	0.004** (3.46)	0.003 (1.22)
Overall R squared	0.45	0.33	0.09	0.12	0.02	0.08	0.59
Between R squared	0.21	0.15	0.01	0.09	0.03	0.05	0.57
Observations	80	783	151	349	124	150	537
Countries	3	55	19	32	11	24	31

Note: The dependent variable is 'government tax revenue', though definitions vary. Year dummies are included, as are country fixed effects. In the source row, Cen stands for Central, Gen for General and Bud for Budgetary. 'African R M' stands for African Revenue Mobilisation. GFS 2001 and 1986 refer to two different vintages of the IMF GFS data, with substantial differences in data definitions. Other details as in Table 1.

More elaborate techniques, such as GMM, would seem to compound the problem rather than solve it. For example, in the 2,458 cases where the same source is available for two consecutive years for a country, there is an average difference of 1.3 per cent in tax revenue as a percentage of GDP. In the ninety-six cases where different sources are used for consecutive years, the mean difference is 4.9 per cent. Difference GMM estimations would see large random fluctuations that result from using different sources. These are not controlled for, and represent outliers that may erroneously drive the results. While Carter (2013) deals with many of the more technical aspects of empirical research in this area (concluding that there is little evidence of a negative effect from aid on tax), it appears there are much more fundamental problems with the data that are only compounded by the use of methods such as GMM.

3 Four options

Given the variety of dependent variables available there are four approaches: to treat the candidate dependent variables as interchangeable, to treat them completely separately, to model their relationship, or to use completely new data. In order to explore the validity of the first three options, I have constructed a new datset which includes all of the independent *and* *dependent* variables described by Benedek et al. (2012) for 1980-2011 for their chosen sample, from the original sources.

3.1 Option 1 Treat variables as interchangeable

Benedek et al. (2012) treat the different variables as essentially interchangeable, and replace any missing observations from one dataset with those of any other. The benefit of this approach is that it minimises selection issues by maximising the sample size. However, it is potentially very problematic and at the very least introduces abnormally high measurement error. The preceding section showed that sample-specific dummies are significant when added to the main specification used in Benedek et al. (2012), and that running seven separate regressions leads to very different coefficient estimates for the variables of interest. A further complication that is noted here, but not explored further, is that the order in which missing data was replaced may have large effects. This may seem trivial, but with seven sources there are 5,040 possible orderings[4], which is problematic unless the different sources are completely compatible.

As a further test of whether the various sources are measuring the same thing, I use the newly-constructed dataset to investigate cases where more than one source exists for any given observation. If the variables are to be treated as interchangeable they must be more than correlated: two variables could be perfectly correlated if one is always double the other, but they would not be interchangeable. The requirement for treating the variables as exchangeable is that they must be consistent with each other. As such, I run twenty regressions of the form

$$y_i = \alpha_{ij} + \beta_{ij} y_j + \varepsilon$$

where y_i is a candidate dependent variable from source i, and y_j is a candidate dependent variable from source j being used as an independent variable. The tests are F-tests of two assumptions, where the null hypotheses represent variables which could be substituted without introducing bias or noise.

$$H_0 : \alpha_{ij} \neq 0$$

$$H_0 : \beta_{ij} \neq 1$$

The results of these tests are displayed as p statistics in Table 4. Columns in Table 4 refer to the dependent variable (i.e. i) and rows to the independent variable (i.e. j). Of the forty comparisons shown, Table 4 shows that the null hypothesis is rejected in thirty-three cases at the 1 per cent level: the different sources do not measure the same thing in the same way.

4 This is due to 7!=5040 capturing all possible permutations.

Table 4 P statistics to test for agreement between sources

		Budgetary GFS	Central GFS	General GFS	WEO	OECD
Budgetary GFS	Alpha	.	0.00	0.00	0.00	0.00
	Beta	.	0.14	0.91	0.75	0.00
	N	1691	1018	497	1385	23
Central GFS	Alpha	0.00	.	0.00	0.00	0.00
	Beta	0.01	.	0.34	0.00	0.28
	N	1018	1797	843	1161	52
General GFS	Alpha	0.00	0.00	.	0.00	0.00
	Beta	0.00	0.00	.	0.00	0.01
	N	497	843	868	676	27
WEO	Alpha	0.00	0.00	0.00	.	0.64
	Beta	0.00	0.00	0.00	.	0.52
	N	1385	1161	676	3602	57
OECD	Alpha	0.00	0.00	0.00	0.00	.
	Beta	0.00	0.00	0.00	0.00	.
	N	23	52	27	57	89

Note: These regressions do not limit the sample to that of Benedek et al. (2012), and so may state a larger number of sources per observation with regard to their sample.

The differences between sources have been shown in Section 2 to make a difference when there is only one source per observation. Table 4 reports tests regarding whether the various sources measure the same variable, and demonstrates very strongly that they do not meet simple tests of their agreement. In only one case do the tests fail to reject the null hypothesis: when the OECD variable is used as the dependent variable and the WEO variable is used as the independent variable. In the regression where those roles are reversed, the null hypothesis is rejected.

3.2 Option 2 Different regressions

The second option, if treating the various dependent variables as interchangeable is not sensible, is to treat all variables as completely independent. This avoids needlessly introducing noise into the regressions, though it does potentially introduce problems of sample selection bias that Benedek et al. (2012) avoid. The newly-created dataset means that the maximum number of sources for the maximum number of variables was created. Unlike the results in Table 3, any individual country can appear multiple times if the relevant data exists. I make three other changes to the preferred specification of Benedek et al. (2012). First, the specification introduces a one-year lag for aid variables. This is more realistic, as it is difficult to imagine aid having a contemporary effect on tax revenue. It also reduces fears of endogeneity (while not removing them), as tax shortfalls could easily lead to short-term contemporary increases in aid, such as in the case of a natural disaster. Second, I disaggregate trade into imports and exports, as they are likely to have different effects (see Clist and Morrissey 2011). Specifically, it is to be expected that imports increase tax revenue, as they provide a relatively easy avenue for tax collection. By contrast, exports are usually taxed less than domestic consumption, and so we would expect a negative coefficient given the opportunity cost of high exports includes foregone tax revenue. Third, I move away from estimating nonlinearities in the aid-tax relationship. Given data constraints and quality, it does not seem prudent to attempt to recover precise estimates of turning points. As can be seen from Table 2 estimates of turning points are very unstable, and so the preferred estimation is less ambitious.

Table 5 Preferred specification 1980-2011, by data source

Source of dependent variable	Budgetary GFS	Central GFS	General GFS	WEO	OECD
Loans/GDP, Lagged	-0.000 (-0.06)	0.004 (0.47)	-0.020*** (-3.95)	-0.000 (-0.21)	0.123 (1.07)
Grants/GDP, Lagged	0.006* (1.76)	-0.005 (-1.25)	0.006 (0.75)	0.004** (2.12)	-0.017 (-0.07)
Agriculture	-0.006 (-1.49)	-0.010 (-1.38)	-0.013 (-1.35)	-0.009** (-2.52)	-0.054* (-3.28)
Industry	-0.004 (-1.53)	-0.002 (-0.36)	0.000 (0.03)	0.000 (0.14)	0.014** (6.04)
Ln(GDP pc)	0.349 (1.57)	0.237 (1.33)	0.173 (0.89)	-0.134 (-1.26)	-0.268 (-0.89)
Imports	0.008*** (4.16)	0.004 (1.48)	0.003 (1.20)	0.004*** (3.17)	0.000 (0.03)
Exports	-0.003 (-1.41)	-0.003 (-0.94)	-0.005 (-1.42)	0.002 (1.34)	-0.013 (-2.76)
Overall R squared	0.00	0.11	0.00	0.13	0.55
Between R squared	0.03	0.11	0.03	0.06	0.40
Observations	908	756	254	1540	77
Countries	81	63	35	99	3

Note: See Table 1 notes for details. Here, the dependent variable is given in the first row, using all available data from the reconstructed dataset. I do not include IMF data using the 1986 definition or data from the *African Revenue Mobilisation* dataset due to constraints on space/availability. Loans and Grants are both measured as a percentage of GDP, logged and lagged.

What can be learnt from the results reported in Table 5? It is heartening that despite problems with the dependent variable, coefficients on the variables agriculture, industry, exports and imports tend to agree. The most consistent coefficient estimate is that for agriculture, which is always found to be negative. The last column, relating to the OECD source, has fewer observations and is the least similar to others. As in Clist and Morrissey (2011), the decision to split trade into imports and exports is justified by the coefficients on the two variables being of opposite signs. The signs are in line with theoretical predictions: imports are relatively easily to tax and exports lead to foregone domestic consumption, which would have been taxed. The log level of per capita income is insignificant in every regression.

Turning to the variables of interest, aid grants and loans are significant in only three of a possible twelve cases. Both are found to be positive and negative in difference specifications. However, it is remarkable that in no specification is the coefficient on aid grants negative and significant. In fact, the only significant effects for aid variables are *positive* for grants and *negative* for loans.

3.3 Option 3 MIMIC

The third option represents a middle way between treating the different possible dependent variables as interchangeable and treating them separately. A multiple indicators multiple cause (MIMIC) model can be applied in this case. The model was introduced by Jöreskog and Goldberger (1975), and an example of its use can be found in Giles (1999). To describe briefly, the model states that there is an unobserved latent variable y^* (tax revenue), but multiple indicators $y_i,\dots,y_m$ are observed (the various candidates for the dependent variable). The MIMIC model then considers two sorts of relationships. For the relationship between the causes and the latent variable, an equation of the following form is estimated:

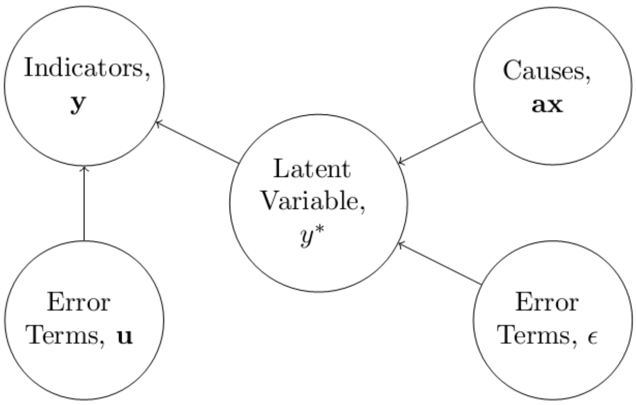
$$y^* = a_1x_1,\dots,a_kx_k + \varepsilon \qquad (1)$$

The relationship between the latent variable and the indicators is also modelled, using equations of the form:

$$y_1 = \alpha_1 + \beta_1 y^* + u_1,$$
$$\ldots, \qquad (2)$$
$$y_m = \alpha_m + \beta_m y^* + u_m$$

These are then jointly estimated using maximum likelihood.[5] Figure 1 is perhaps useful to visualise the model. The benefit of this approach should be seen predominately as a robustness check on the results presented in Table 5. The MIMIC model allows multiple candidate dependent variables to be used simultaneously, and can provide a larger sample than those using one candidate at a time.

Figure 1 The MIMIC model



One complication of using MIMIC models when there is not absolute overlap in sources is that it may be that models do not converge. Only one combination of sources works, and so Table 6 reports the coefficients from that regression. A further complication is that it is not feasible to use fixed effects in this case.

Table 6 MIMIC Model, 1980-2011

Equation (2)	
Aid loans	0.002
Lagged	(0.40)
Aid loans	0.029***
Lagged	(8.68)
Agriculture	-0.011***
	(-5.25)
Industry	0.007***
	(4.19)
Ln(GDP pc)	0.104***
	(5.56)
Imports	0.005***
	(6.19)
Exports	0.004***
	(3.39)
R squared	0.47

Equations of the Form (1)	
Dependent variable:	BA tax revenue (GFS)
Latent variable	1.000
	(constrained)
Constant	-3.028***
	(-15.71)
R squared	0.22
Dependent variable:	Revenue (WEO)
Latent variable	0.806***
	(9.87)
Constant	2.205***
	(16.19)
R squared	0.81
Log likelihood	69426.71
Observations	3281
Countries	114

Note: In Equation (2) on the left panel, the dependent variable is the estimated latent variable. This is constrained in the first equation in the form (1), such that the coefficient on the slope of GFS tax revenue is equal to one. Thus the latent variable can be thought of as the underlying concept of tax revenue, with two indicators. Z statistics are shown below parameter estimated in parentheses. R squared is calculated on an equation basis, and refers to the equation above it. See Table 1 for other details.

5 In Stata, the relevant command is sem. All Stata code and data is freely available at <https://sites.google.com/site/paulclist/data>.

Bearing in mind the various caveats regarding the difficulty in model convergence, I present only one converged specification which uses two candidate dependent variables: GFS's Budgetary Tax Revenue and WEO's Revenue variables. Table 6 contains the full results for all equations. One estimated coefficient must be constrained in the model, and so $\hat{\beta}$ for the latent variable's effect on BA Tax Revenue is constrained to equal 1. The benefit of this approach is evident in the sample size, as the sample is much larger than those presented in Table 5. Dealing first with the equations of the form (1) on the right panel, we can see that both indicators are positively correlated with the latent variable. The WEO variable measures all government revenue, and so it is no surprise that the constants are such that WEO is five units (where the variables are expressed as a percentage of GDP, logged) higher. The coefficient on the latent variable in the WEO equation is less than one, meaning that revenue is estimated to be less affected by tax revenue than the GFS measure. Moving to the left panel, we see familiar and reassuring parameter estimates in most cases. For the variables of interest, we see that both types of aid are consistent with higher tax receipts in the following year, but the result is only significant (at the 1 per cent level) for aid grants. The control variables are almost all as expected and all significant: higher tax receipts are associated with less agriculture, more industry, being richer and importing more. The only variable which is not in line with expectations is that related to exports, where a positive effect is found.

3.4 Option 4: New dataset

A fourth option in attempting to replicate the results of Benedek et al. (2012) has recently become available due to the newly created International Centre for Tax and Development *Government Revenue Dataset* (ICTD GRD). This new dataset is the culmination of a three-year effort to provide a consistent and accurate estimate of tax revenue in developing countries. Crucially, there was an emphasis on compatible sources in the construction of this dataset to allow comparisons across countries and time periods. Table 7 provides the results of applying my preferred specification (used in Table 5) to the new data. This exercise provides several benefits. First, it provides a further test of previous research in this area, by using what appears to be the most accurate dataset available. In maintaining a comparable specification with minimal adjustments, the new data can be used to evaluate previous findings in this area. Second, as the new dataset is used to tackle new issues, it is useful to benchmark it using a familiar econometric specification on older datasets.

Table 7 ICTD dataset, 1980-2009

Aid variables:	Concurrent (1)	Once lagged (2)	Twice lagged (3)
Loans/GDP, Lagged	0.002 (0.76)	0.004 (1.54)	0.002 (0.59)
Grants/GDP, Lagged	-0.004* (-1.78)	-0.001 (-0.76)	-0.001 (-0.49)
Agriculture	-0.007** (-2.39)	-0.006** (-2.24)	-0.005* (-1.94)
Industry	-0.005* (-1.84)	-0.005** (-1.99)	-0.006** (-2.25)
Ln(GDP pc)	0.043 (0.52)	0.081 (0.93)	0.064 (0.75)
Imports	0.008*** (5.40)	0.006*** (3.81)	0.006*** (3.82)
Exports	-0.003* (-1.91)	-0.003* (-1.69)	-0.003* (-1.77)
Overall R squared	0.39	0.38	0.38
Between R squared	0.41	0.42	0.42
Observations	2173	2138	2102
Countries	113	112	112

Note: As for Table 1, but with the dependent variable from ICTD GRD. The aid variables are either concurrent, or lagged one or two years.

Table 7 has the now familiar pattern for the majority of coefficients: tax revenue is positively associated with having higher imports, a smaller agriculture sector, lower exports and a smaller industrial sector. Of the six coefficients on the aid variables, only the coefficient for aid grants in column (1) is significant, with a negative sign. When a lag of one or two years is used, however, the coefficient for aid grants is both smaller and insignificant. This specification and data provide the only negative and significant effect of aid grants on domestic tax revenue I have found. However, a negative contemporary association is not necessarily evidence that aid grants cause lower domestic tax revenue. In any given year, a contemporary association is more likely to be the result of aid donors compensating recipients in adverse conditions (e.g. an economic downturn). If aid does corrode tax revenue, it is more likely to do so over a period of time, and yet the effect becomes insignificant over this time period.

4 Discussion and conclusion

Replication has long been recognised as performing a vital role in assessing the credibility of empirical results (see Camfield and Palmer-Jones (2013) for a recent review of the issues). For this reason I have attempted to replicate the results presented by Benedek et al. (2012), as it is the most high profile recent empirical research paper that reports differential effects from aid grants and loans on tax revenue. Like Carter (2013), I have failed to replicate the specific results. In only one case have I found a reported negative effect for aid grants: when contemporary aid grants are used. The specific replication failed even when using the exact dataset for the dependent variable reportedly used by Benedek et al. (2012). While Carter (2013) also provides an extensive critique of the methods used by Benedek et al. (2012) to combat endogeneity (mainly the use of GMM), I argue that there appear to be much more fundamental issues with the original data used. Specifically, several datasets are apparently used to construct a single dependent variable despite different data definitions. Tests show that previous specifications are not robust to the use of source dummies or running separate regressions by the dependent variable's source. Furthermore, the different candidate dependent variables are statistically and conceptually measuring different things. The use of system GMM in this setting is particularly worrying, as large yearly differences may simply be due to the use of different sources in consecutive years. The issue is wider than simply poor data quality: the empirical results are potentially misleading. For example, a substantial portion of the variable termed ‘tax revenue’ is provided by a variable which explicitly measures total government revenue. It is worrying that this misleading empirical result may influence policy, and underscores the value of replication.

Given the data quality, it is heartening that the various empirical approaches employed here to deal with the issue of multiple candidate dependent variables point in a single direction: there is very little evidence that aid grants undermine domestic tax revenue. The only significant and negative coefficient relating to aid in Table 5 and Table 6 actually relates to aid loans. The only significant and negative coefficient for aid grants relates to contemporaneous aid, and is thus dogged by endogeneity concerns. The weight of the empirical evidence suggests that aid has a relatively small, possibly positive, influence on domestic tax revenue. There is not sufficient evidence to suggest that the composition of aid should be influenced by concerns relating to differential effects on tax revenue. Wider concerns relating to a longer-term corrosive effect on government cannot be dismissed on the basis of these results (even if aid leads to higher tax revenue, it may be that tax is a lower proportion of total government revenue than it would be without aid), but neither can it be used to support them.

Appendices

Appendix 1 System break

Table A1 provides evidence that, as suspected, there is no system break in the data provided by Benedek et al. (2012).

Table A1 System break

Variable	Model 1	Model 2
Aid	0.000 (0.03)	
*dummy	0.005 (0.64)	
Aid squared	-0.000 (-0.28)	
*dummy	-0.000 (-0.30)	
Loans		-0.018* (1.73)
*dummy		-0.009 (-0.80)
Loans squared		-0.000** (-2.10)
*dummy		0.000 (0.99)
Grants		-0.005 (-0.43)
*dummy		0.010 (0.84)
Grants squared		0.000 (0.06)
*dummy		-0.000 (-0.74)
Agriculture	-0.006* (-1.80)	-0.007* (-1.95)
Industry	-0.007 (-1.19)	-0.003 (-0.67)
Ln(GDP pc)	0.251** (2.12)	0.212 (1.64)
Trade openness	0.003*** (3.03)	0.003** (2.38)
Overall R squared	0.28	0.32
Between R squared	0.28	0.39
Observations	2174	1968
Countries	99	97

Note: '*dummy' denotes the preceding variable multiplied by a dummy which takes the value 1 if the year is 1985 or later. For all other details, see Table 1.

Appendix 2 Data sources

All data and Stata code can be accessed from <https://sites.google.com/site/paulclist/data>.

Tax Revenue, Organisation of Economic Cooperation and Development (OECD). Total tax revenue as a percentage of GDP.

Tax Revenue, Government Finance Statistics (GFS). Two vintages are used by Benedek et al. (2012): 1986 and 2001. These refer to two different data definitions. Furthermore, three types of variable are used: budgetary, central government or general government.

Tax Revenue, World Economic Outlook (WEO) is in fact general government revenue (percentage of GDP). To quote from the data definition: ‘Revenue consists of taxes, social contributions, grants receivable, and other revenue. Revenue increases government's net worth, which is the difference between its assets and liabilities’ (IMF 2001, paragraph 4.20).

Aid: Total, Grants and Loans. Net ODA, ODA grants, and ODA loans, relative to GDP, are from the OECD database.

Agriculture and Industry. These are taken from the World Bank's *World Development Indicators* (WDI), and describe the added value as a percentage of GDP.

Trade Openness, Exports and Imports. The numerator is taken from the IMF's *International Finance Statistics* (IFS) database. The denominator (GDP) is taken from the *World Economic Outlook*.

GDP per capita is taken from the World Bank's *World Development Indicators* (WDI) database, and calculated in constant (2000) US dollars.

References

Benedek, D., Crivelli, E., Gupta, S. and Muthoora, P. (2012) *Foreign Aid and Revenue: Still a Crowding Out Effect?*, IMF Working Papers WP/12/186, International Monetary Fund

Camfield, L. and Palmer-Jones, R. (2013) ‘Three ‘Rs’ of Econometrics: Repetition, Reproduction and Replication’, *Journal of Development Studies* 49(12): 1607-1614

Carter, P. (2013) ‘Does Foreign Aid Displace Domestic Taxation?’, *Journal of Globalization and Development* 4(1): 1-47

—— (2010) *Foreign aid and taxation, revisited*, CSAE Conference Paper 180, <http://www.csae.ox.ac.uk/conferences/2010-EdiA/papers/180-Carter.pdf>

Clist, P. and Morrissey, O. (2011) ‘Aid and tax revenue: Signs of a positive effect since the 1980s’, *Journal of International Development* 23(2): 165-180

Deaton, A. (2013) *The Great Escape: Health, Wealth, and the Origins of Inequality*, Princeton University Press: Princeton

Giles, D. (1999) ‘Measuring the hidden economy: Implications for econometric modelling’, *The Economic Journal* 109(456): 370-380

Gupta, A. (2007) *Determinants of Tax Revenue Efforts in Developing Countries*, IMF Working Paper WP/07/184, International Monetary Fund

Gupta, S., Clements, B., Pivovarsky, A., and Tiongson, E. (2004) ‘Foreign Aid and Revenue Response: Does the Composition of Aid Matter?’, in S. Gupta, B. Clements and G. Inchauste (eds), *Helping Countries Develop: The Role of Fiscal Policy*, Washington: International Monetary Fund

House of Commons International Development Committee (2013) *Pakistan Tenth Report of Session 2012-13, Volume 1*, London: The Stationery Office Limited

IMF (2001) *Government Finance Statistics Manual 2001*, Washington DC: International Monetary Fund

Jöreskog, K. and Goldberger, A. (1975) ‘Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable’, *Journal of the American Statistical Association* 70(351): 631-639

Schmidt, W. (1964) ‘The Economics of Charity: Loans Versus Grants’, *The Journal of Political Economy* 72(4): 387-395

International Centre for Tax and Development
at the Institute of Development Studies
Brighton BN1 9RE, UK
T: +44 (0)1273 606261
F: +44 (0)1273 621202
E: info@ictd.ac
www.ictd.ac