www.elsevier.com/locate/worlddev

World Development Vol. 78, pp. 554–571, 2016

http://dx.doi.org/10.1016/j.worlddev.2015.10.006

What We Miss By Missing Data: Aid Effectiveness Revisited

ANJA BREITWIESER and KATHARINA WICK*
University of Vienna, Austria

Summary. — Missing data is an issue in many empirical applications, as it may entail efficiency losses as well as biased results. We address these problems within the literature that investigates the effect of foreign aid on welfare. Using multiple imputation techniques to account for the missing data, we find lower aid effectiveness. In addition, imputation allows for comparison of different welfare indicators within the same framework. We find that the respective indicator choice can matter for the ensuing results.

Key words — aid effectiveness, welfare indicators, sample selection, missing data, multiple imputation techniques, developing countries

1. INTRODUCTION

The empirical aid effectiveness literature mainly focuses on the effect of foreign aid on economic growth. [1] Only a limited number of studies discusses the effect of aid on other welfare indicators, such as school enrollment for the education sector and immunization rates for the health sector (Christensen, Homer, & Nielson, 2010; Dietrich, 2011; Dreher, Nunnenkamp, & Thiele, 2008; Gauri & Khaleghian, 2002; Mishra & Newhouse, 2009; Michaelowa, 2004; Michaelowa & Weber, 2007a, 2007b; Wilson, 2011). While these studies are a most welcome diversion from the strong focus on economic growth, they entail challenges of their own. In particular, the choice of a specific welfare indicator may influence estimated aid effectiveness in two conceivable ways: directly, via a distinct effect of aid on any particular indicator, or indirectly via the respective indicators data coverage (Clemens *et al.*, 2012; Dreher & Gehring, 2012).

For illustrative purposes, consider the health sector: here, DPT immunization (Dietrich, 2011), child mortality rates (Mishra & Newhouse, 2009), and life expectancy (Wilson, 2011) have been used as indicators of aid effectiveness. The estimated results however, may be directly related to the indicator choice. In this case, the effect of foreign aid on life expectancy may simply be innately different from the effect of aid on DPT immunization. In addition, the indirect effect—via the respective indicator's data coverage—may influence the results. [2] If the available data points are not a random sub-sample of the (unobservable) complete data, missing data cause not only efficiency losses, but may produce biased results.

Reflecting the above-mentioned problems, results within the literature of aid's effect on welfare are indeed inconclusive. The present article uses multiple imputation techniques to disentangle the two effects that may be responsible for this scattered evidence.

The most common approach to address missing data issues is the method of listwise deletion, i.e., in a panel setting dropping all country–year pairs for which any variable included in the estimated model is missing. Using listwise deletion, a relatively small fraction of missing data may result in considerable losses of observations. Ross (2006) for example reports a fraction of 18% of the data missing, which subsequently results in the loss of 75% of observations. With respect to our data, the method of listwise deletion results in a loss of up to 58% of the observations. An immediate consequence of this reduction in sample size is an efficiency loss in the analysis.

If, however, the observed data are not a random sub-sample of the complete data, [3] the consequences are even more profound. Ross (2006), for instance, finds that the positive effect of democracy on the welfare of the poor has been overestimated in the past, since mainly well-performing nondemocratic states are missing from existing empirical analyses. Similarly, if low performing countries are prone to missing observations—for example due to monetary and human resource constraints on data collection—while at the same time receiving large aid flows, excluding these observations from the analysis, may overestimate the positive effect of foreign aid. In short: Ignoring the missing data pattern entails ignoring sample selection bias.

In this paper, we apply multiple imputation techniques [4] to deal with the issue of missing data. We use the available information to reintroduce the missing observations into the analysis, while taking the inherent uncertainty of the imputed values into account. From careful inspection of the missing data mechanism, we argue that the multiple imputation method is the preferred approach to address potential sample selection when investigating the issue of aid effectiveness.

Using the imputed data, our empirical results indicate that (1) analyzing the effect of aid on welfare indicators [5] may be biased due to sample selection and (2) foreign aid has distinctively different effects on different indicators arguably capturing progress in the same sector (such as enrollment rates and the pupil–teacher ratio for the education sector).

The paper is structured as follows. First, we start with a short review of the literature on the effect of foreign aid on welfare. In the next section, we analyze the missingness pattern of our data and argue for the expedience of imputation techniques in the context of our research question. Following, we describe the multiple imputation approach and its results when applied to our data. In addition, we present our estimation method. After a short description of the data, our estimation results are presented. The last section discusses the main findings.

* We thank Alejandro Cuñat, Neil Foster, Jean-Philippe Platteau, Robert Stehrer, an anonymous referee, and participants of the 2013 Conference on Development Economics and Policy of the German Economic Association for their helpful comments and suggestions. The usual disclaimer applies. Final revision accepted: October 2, 2015.

2. EMPIRICAL LITERATURE

Studies investigating the effect of foreign aid on welfare indicators mainly use specific sector aid to measure foreign aid flows. Dietrich (2011), Mishra and Newhouse (2009), and Gauri and Khaleghian (2002) focus on the health sector and analyze the effect of aid on infant mortality and immunization rates (DPT, measles and hepatitis B) respectively. While Gauri and Khaleghian (2002) find at best weak evidence for a negative impact of health aid on infant mortality, Mishra and Newhouse (2009) and Dietrich (2011) conclude that there is a robust and significant positive effect of health aid on immunization rates.

With respect to the education sector, Michaelowa (2004), Dreher *et al.* (2008), Michaelowa and Weber (2007a, 2007b), and Christensen *et al.* (2010) investigate the effect of foreign aid on net and gross enrollment rates (primary, secondary, and tertiary), and primary completion rates. Michaelowa (2004), Michaelowa and Weber (2007a, 2007b) and Dreher *et al.* (2008) find some evidence for a significantly positive but quantitatively small effect of education aid on primary enrollment and completion rates. Christensen *et al.* (2010) on the other hand find no significant impact of education aid on enrollment rates using a latent growth model.

Two studies take a broader approach and analyze effectiveness on a variety of indicators. Wolf (2007) uses cross-section data for estimating a simultaneous equation model to examine the effect of education aid, health aid, and water and sanitation aid on literacy and primary school completion rates, infant and child mortality rates, and access to improved sanitation and water sources respectively. She finds a significant positive effect of education and health aid. Potential endogeneity of aid however, is not accounted for in this analysis. Findley, Hawkins, Nielsen, Nielson, and Wilson (2010) use data of aid commitments by multilateral organizations in their analysis. They define five categories of aid and examine their effect using matching techniques to form subclasses and then estimate latent growth models. While their results differ by categories they find some evidence of effectiveness (i.e., aid fostering the intended purpose) for democracy, human rights, and environmental aid.

While these studies' focus on sectoral aid has its advantages (e.g., fewer nuisances) sector aid is not in all circumstances preferable to an aggregate aid measure. For instance, aggregate aid measures allow to (at least partially) account for aid fungibility and external effects. [6] In addition sector aid data availability is either limited to data on aid commitments or to a sample starting in the 1990s. Considering this, we thus resort to an aggregate aid measure.

3. MISSING DATA

The maximum number of observations (N), that is, the number of country–year pairs for which aggregate foreign aid inflows have been reported to the OECD during 1970–2009, is 5,274. [7] An analysis of the effect of these aid flows on welfare indicators, however, would be based on a considerably smaller sample due to missing observations.

Rubin (1976) develops a framework for the different types of missing data patterns. In this framework, missingness is classified into missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). For any data set a matrix R can be defined such that it identifies the observed and the missing observations within the data. That is, R contains the value zero for missing and one for observed values. R can be considered as a combination of random variables with a joint probability distribution. MCAR implies that the probability of missingness neither depends on the observed data nor on the unobserved parts of the data. That is, $P(R|Y^c) = P(R)$, where Y^c denotes the complete data. MAR is less restrictive and allows the distribution to depend on the observed data, such that: $P(R|Y^c) = P(R|Y^o)$, where Y^o denotes observed parts of the data. [8] In other words, conditional on the observed data the probability of missingness does not depend on the missing data. The missing data are considered missing not at random (MNAR) if the condition $P(R|Y^c) = P(R|Y^o)$ does not hold. That is, if conditional on Y^o, the probability of missingness does depend on the missing data (Y^m) (Schafer & Graham, 2002).

If the data are MCAR, that is the observed sample is a random sub-sample of the unobserved complete sample, observed case analysis is subject to efficiency losses but unbiased. In the context of our analysis this assumption, however, is rather strong. Countries with higher income have more monetary and human resources available for data collection and are thus less likely to have missing data (Bueno de Mesquita, Smith, Siverson, & Morrow, 2003). Dictatorships on the other hand may be less willing to collect and report data on certain variables (Hollyer, Peter Rosendorff, & Vreeland, 2011; Bueno de Mesquita *et al.*, 2003). In the context of foreign assistance, outside agencies (bilateral or multilateral donors) may ask governments to regularly report certain data. Since the deterrent effect of non-compliance increases with the flow of foreign aid potentially withdrawn, missingness would in that case be a function of the amount of aid received. These examples support the notion that the observed data (e.g., on income, political regime, etc.) influence missingness and that the MCAR assumption is too strong. Whether the observed data can predict missingness can be easily tested. Running a fixed effects binary logit regression on generated indicator variables, which are 1 if the observation for the respective variable is missing and zero otherwise, we indeed find that income, form of government, civil liberties, and other variables significantly predict missing observations. [9] In this case, however, results from complete case analysis cannot generally be assumed to be unbiased. [10]

Eliminating MNAR from the set of possible missing data mechanisms is less straight forward since testing is not possible in this case. MNAR implies that the pattern of missingness depends on the values of the variable under consideration itself. One could argue that for example school enrollment rates are more likely to be missing for lower enrollment rates as governments might be more reluctant to collect or report these numbers. Yet, governments are usually not exempt from internal or external public pressure, which limits their autonomy of decision. In a democratic environment for instance, citizens could pressure their governments into collecting the respective data and hence make missingness a function of a country's regime type. Similarly, in particular when considering aid recipient countries, external pressure for data collection might be exerted by bilateral and multilateral donors. Inclusion of auxiliary variables (e.g., civil liberties, freedom of press, etc.) in the imputation model can thus support the MAR assumption (White, Royston, & Wood, 2011).

Table 1 reports the median, the number of observations, and the number of countries for each dependent variable for two different samples. [11] For each welfare indicator used in the subsequent analysis, countries are split into two samples with respect to their degree of missing data. Table 1 reports on the