Bangladesh Development Studies
*Vol. XXXIX, March-June 2016, Nos. 1& 2*

# Challenges of Missing Data in Analyses of Aid Activity: The Case of US Aid Activity

BEÁTA UDVARI*
GÁBOR DÁVID KISS**
JULIANNA PONTET***

Analysing aid activities has been in the centre of academic research; nevertheless, it is demanding to conduct long-term time series analyses due to missing data. Although there are several methods available to overcome this challenge, their distortion effect may result in unpredicted impacts on aid allocation. Thus, this paper aims to analyse the long-term motivations of US aid allocation with panel regression models. Two methods of handling missing data were tested in order to answer the question whether there is a significant difference in the results or not. Results suggest that there are several tools in the hands of a researcher to overcome missing data problems without any distorting effects. Furthermore, results reinforce the idea that US aid allocation has mainly been motivated by its economic drivers (export possibilities) rather than by war or conflict fears in the long run.

**Keywords**: Aid Allocation, Missing Data, Panel Regression

**JEL Classification**: F35, C15, C33

## I. INTRODUCTION

Nowadays, the migration crisis in the European Union has raised the attention to developing countries and their economic and social challenges: the consequences of their unhandled economic and social problems have resulted in significant and costly global issues. As it is well-known, there are several attempts and international initiatives to promote development in these countries. They appear both on international and national level, and there are local initiatives as well. There is still debate on the impacts of foreign direct investment (e.g. Gohou and Soumaré 2012, Hossain 2015) or, for example,

*Assistant Professor, Faculty of Economics and Business Administration, University of Szeged, Hungary.
**Assistant Professor, Faculty of Economics and Business Administration, University of Szeged, Hungary.
***Ph.D. student, Faculty of Economics and Business Administration, University of Szeged, Hungary.

microfinance (Bangoura 2012) on poverty reduction and development. However, one of the most important global attempts is the international development cooperation and allocation of foreign aid. The effectiveness of foreign aid provided by either bilateral donors (developed countries, emerging donors) or multilateral agencies (international organisations) is widely analysed with different statistical and econometric methods. Since these results are the ground for further attempts of aid activity, at the same time, influencing policy decisions, the reliability of data, the methods used and the results of empirical aid analyses are essential.

Aid data mainly meet the accredited reliability criteria. For instance, the Organisation for Economic Cooperation and Development's (OECD) Creditor Reporting System is widely used in aid analyses. The OECD and its Development Assistance Committee (DAC) are responsible for reliability; however, donors themselves submit reports and data on their own aid activity. Despite this fact, in several cases, it is very difficult to conduct (long-term) time series analyses because of missing data (due to the lack of report by donors). As a result, analysing long-term changes in motivations of donor countries or examining long-term economic effects of aid is still a great challenge. In order to handle the missing data problem, there are numerous statistical and econometric methods. This paper aims to analyse whether there is any difference between the different methods regarding their results on the field of international development cooperation and aid allocation**.** In order to answer the question, we used the United States as an example and analysed what kind of factors determined its aid allocation between 1967 and 2014. This is the longest time series we could gain from the OECD database covering more than 60 recipient countries.

The structure of the paper is as follows. Section I gives an overview on the aid allocation of the United States detailing the motivation background. This theoretical background gives the basis for selecting indicators. Section II outlines the methodological background describing the challenges of missing data by introducing their types and possible solutions. The last section contains the empirical analysis of the aid allocation of the United States. This empirical analysis introduces different solutions to the problem of handling missing data; therefore, we can learn which method could be the most effective.

## II. MOTIVATIONS BEHIND AID ALLOCATIONS

The United States of America is one of the largest donors in the world. According to the OECD aid database (OECD CRS 2016), the United States provides slightly more than 18 per cent of the global aid, while it is a bit more (above 20 per cent) if we take only the performance of the OECD Development Assistance Committee into consideration. These data also suggest that the United States finds international development cooperation as an important part of its own foreign policy (Tarnoff and Nowels 2004); nevertheless, the US motivations behind aid allocation have changed over the years.

### 2.1 General Motivations Behind Aid Allocation

By now, aid effectiveness and aid motivations have been thoroughly analysed (see, for instance, Doucouliagos and Paldam 2009, Hansen and Tarp 2001, McGillivray 2003); nonetheless, the results are not unambiguous. Regarding the motivations of aid allocations, it is observed that donors provide aid for several reasons, although their motivations and priorities have changed a lot during the last several decades (Szent-Iványi and Lightfoot 2015). To understand motivations, it is essential to consider both the point of view of donors and the point of view of the recipient countries, since unpredictable aid may increase corruption level in developing countries (Kangoye 2013). Besides some moral reasons, the motivations behind aid allocations of donor countries can be grouped as follows:

*Political interests of the donors*: Alesina and Dollar (1998), Younas (2008) and Einarsdottir and Gunnlaugsson (2016) point out that strategic and political interests of donors are more important than the economic needs of recipient countries, and ethical background does not influence aid allocation significantly either. Furthermore, colonial past and political alliances also determine the amount of the aid.

*Governance of the recipient countries*: In their research, Alesina and Dollar (1998) and Berthelemy and Tichit (2004) find that countries in the process of democratisation are supported more. Collier and Dollar (2001) emphasize that foreign aid can contribute to the improvement of the public services provided by governments; and these results are more spectacular if aid flows to poorer countries. Hossain (2015) also claims that stable macroeconomic and fiscal policies and political stability are needed for least developed countries to attract more capital.

*Policies and internal changes in donor countries* also influence aid allocation, as the studies of Round and Odedokun (2004) and Chong and Gradstein (2008) proved. Furthermore, if there is a conservative government in a donor country, aid allocations to low income countries will likely decline (Tingley 2010). The internal situation (inequality, corruption, political leaning and taxes) in a donor country also influences aid allocation, and these circumstances seem to have larger effects than the economic conditions in the recipient country (Chong and Gradstein 2008).

*The 4P* also affects aid allocation as Clist (2011) details. In this sense, 4P refers to poverty, population, policy (including good governance and freedom) and proximity (including religion, language, colonial past and trade relations).

*War on terrorism* has recently become a crucial motivation in aid allocation, too. It is seen that more aid is provided to countries where positive regime changes and armed conflicts are expected. A similar tendency appears in neighbouring countries of a war-torn country (Brück and Xu 2012), the aim being to avoid further conflict and preserve peace.

*Anti-corruption movement also deserves* special attention, since 1997 multilateral donor agencies have provided less aid to more corrupt countries; while bilateral donors have supported countries regardless their corruption level. However, the difference between the two main donor groups is decreasing: before 1997, donor countries either did not pay attention to corruption level or provided aid to more corrupt countries, while nowadays donor countries have become more sensitive to corruption level, so the anti-corruption movement of international organisations seems to be successful (Charron 2010).

### 2.2 Effects of Democratisation Process on Aid Allocation

The Monterrey Consensus of 2002 emphasised that sound policies, good governance and rule of law should be a priority for donors while allocating aid (UN 2002, Dollar and Levin 2006). However, empirical results show that multilateral agencies started to pay attention to this requirement only after the Millennium, while bilateral aid allocation does not seem to consider these very priorities. Empirical results, on the other hand, indicate that aid effectiveness is better in countries with better and more stable governance. For instance, Hoeffler and Outram (2011) emphasize that development aid is more effective in countries where democracy and good governance exist, and more open and democratic countries receive 36 per cent more aid than less democratic countries; however, there are some exceptions: for instance, France pays less attention to these criteria (Alesina and Dollar 1998). Looking at human rights, donor countries are more likely to give aid to countries with better human rights records, but this result is reversed when other factors, for instance democracy, are controlled in the model (Clist 2011). On the other hand, the author emphasizes that donors tend to prefer democracy to human rights. Gates and Hoeffler (2004) found that globally a higher number of democratic countries had received more aid on average, and their findings suggest that aid allocations of the Nordic donors follow this trend as well. Furthermore, the donor community cuts aid disbursements in countries where there are coups d'e´tat (Masaki 2016).

There are also some counter-opinions. For instance, Reinsberg (2014) suggests that donor selectivity in favour of democratic policies and good governance remained very low, without any significant improvement over the last two decades. The author adds that more corrupt countries seem to receive more aid (similar to the results of Alesina and Dollar 1998). This result contradicts the findings of Charron (2010). To sum up, it seems that there is no consensus on the question whether a democratic country attracts more or less aid. It suggests that donors should be analysed separately.

### 2.3 Motivations behind the US Aid Allocation

Many studies suggest that the United States often behaves in a different way from other donor countries (Gates and Hoeffler 2004, Harrigan and Wang 2011, Masaki 2016). For example, there is no difference between the pre- and post-Cold War aid allocation behaviour of the United States, as Balla and Reinhardt (2004) claim: the United States allocated a huge amount of aid to the neighbouring countries of conflict-affected areas and still continues to do so.

The United States regularly follows its own geopolitical, commercial and other interests while allocating aid to developing countries (Harrigan and Wang 2011). It does not accept the international norm of political conditionality, which results in higher amounts of aid granted to countries with coups, as Masaki (2016) adds. This norm was a feature of US aid allocation both during and after the Cold War.

This may be explained by the fact that war on terror has become a crucial aspect of US aid allocation. Fleck and Kilby (2010) found that countries which had received military aid from the United States received economic aid as well almost in every case; however, the need of the countries (e.g. poverty) is not an important principle in United States aid allocation. As the findings of Hoeffler and Outram (2011) show, the United States provides more aid to countries with poorer human rights records, and the United States inherently values democracy for ideological reasons rather than any positive effect on poverty reduction (Clist 2011).

The United States Agency for International Development (USAID) has supported six important development objectives based on five main strategic goals in recent years. Table I shows net costs of these strategic goals and objectives for the years of 2014 and 2015. The six main objectives of the USAID are *Peace and Security*, *Governing Justly and Democratically*, *Investing in People*, *Economic Growth*, *Humanitarian Assistance*, and *Operating Unit Management.* These objectives include further programme areas.[1]

The data suggest that supported areas of US aid are largely in line with the classical objectives of foreign assistance (economic growth, humanitarian assistance, investing in people) and naturally also with the US foreign policy (peace, security, support democratisation process). There was a small increase in the total net cost of USAID operations: in 2014 it was 11,671,109 and in 2015 it was 12,528,594 thousand dollars. The two mostly supported areas in both years were *Investing in People* and *Economic Growth*, but the cost of the latter decreased from 2014 to 2015. The first two supported areas are followed by *Humanitarian Assistance* (2,121,191 thousand USD in 2014 and 2,783,754 thousand USD in 2015), *Governing Justly and Democratically* (1,420,292 thousand USD in 2014 and 1,400,277 thousand USD in 2015) and *Operating Unit Management* (718,970 thousand USD in 2014 and 788,835 thousand USD in 2015). Finally, the least supported area was surprisingly *Peace and Security* with a small increase from 2014 to 2015 (671,264 thousand USD and 718,411 thousand USD).

---

[1]Peace and Security: counterterrorism, combating weapons of mass destruction, stabilisation operations and security sector reform, counternarcotics, transitional crime, conflict mitigation and reconciliation.

Governing Justly and Democratically: rule of law and human rights, good governance, political competition and consensus-building, civil society.

Investing in People: health, education, social and economic services, and protection for vulnerable populations.

Economic Growth: macroeconomic foundation for growth, trade and investment, financial sector, infrastructure, agriculture, private sector competitiveness, economic opportunity, environment.

Humanitarian Assistance: protection, assistance and solutions; disaster readiness, migration management.

Operating Unit Management: crosscutting management and staffing, programme design and learning, administration and oversight.

TABLE I

**2014-2015 NET COST PROGRAMME AREAS (IN THOUSAND USD)**

| Strategic goal | Objective | 2014 | 2015 |
|---|---|---|---|
| 1. Counter threats to the US and the international order, and advance civilian security around the world<br>2. Effectively manage transitions in the frontline states | Peace and Security | 671,264 | 718,411 |
| 3. Expand and sustain the ranks of prosperous, stable and democratic states by promoting effective, accountable, democratic governance; respect for human rights; sustainable, broad-based economic growth and well-being | Governing Justly and Democratically | 1,420,292 | 1,400,277 |
| | Investing in People | 2,640,080 | 2,861,007 |
| | Economic Growth | 4,099,312 | 3,976,310 |
| 4. Provide humanitarian assistance and support disaster mitigation | Humanitarian Assistance | 2,121,191 | 2,783,754 |
| 5. Build a 21$^{st}$ century workforce; transparency and accountability | Operating Unit Management | 718,970 | 788,835 |
| **Total Net Cost of Operations** | | **11,671,109** | **12,528,594** |

**Source:** USAID 2014, 2015.

It is not easy to decide what motivates US aid allocations in the long run: democracy-building or poverty decrease? Table I shows that the need of recipient countries is essential, while empirical studies suggest the opposite. Which countries can receive more aid from the United States? Can we notice any long-run tendency? In order to answer these questions, we conducted a time-series analysis in which we handled the missing data problem in two ways.

## III. METHODOLOGICAL BACKGROUND

Missing values may distort the quality of panel data, as list wise deletion weakens their statistical power (Park 2011). Missing data in empirical investigations cause several challenges for researchers working in the field of market research, financial analysis or medical issues, especially in the case of longer time-series. The challenge is how to handle the missing data problem without significantly influencing the final results of an empirical investigation. Since empirical analyses have influence on policies, it is worth analysing whether there is a significant difference in terms of research findings when we use different methods, and, if there is, which method suggests the most reliable result. The aim of this paper is to study whether there is a distortion effect of using different models for handling missing data. In order to answer this question, we analyse the aid allocation of the United States empirically by comparing two of the three mainstream missing value handling methods.

The relevant aid literature provides some methods available to answer this challenge:

- Researchers often give a minimal value to the missing data which was marked zero before, and they calculate its natural log. For example, instead of 0 as aid amount, researchers calculate with 0.01. This method is used, for instance, by Hansen and Tarp (2001), Dollar and Levin (2004) and Younas (2008).
- Wagner (2003) and Cali and te Velde (2011) shared the opinion that the formula of (1+aid) had a distortion effect; therefore, they used the following formula: ln(max(1, aid)) and they also added a dummy variable, where the dummy was 0 when the country received aid and 1 when the country did not receive any support.
- Udvari (2014) used a similar formula to Wagner (2003) and Cali and te Velde (2011), but she left the dummy variable out of the model arguing that this variable had no economic content.

Since the way of handling the problem of missing data may have distorting effects (as Wagner 2003 showed in his research), and an out-of-date method may lead to unrealistic results and cause bad policy implications (McGillivray 2003), we find it crucial to understand missing data as a methodological issue, too.

### 3.1 Types of Missing Data

Three forms of missingness are defined by literature (Graham 2012, Junger and Leon 2015): the first one assumes that data is missing completely at random (MCAR), suggesting that missingness does not depend on the values of the data or on other observed particular variable, and their exclusion does not bias estimations due to their homogeneity (Junger and Leon 2015, Kang 2013). The second type, missing at random (MAR), occurs when dropout is conditionally independent of the variable (Kang 2013), still one can assume some sort of mechanism behind missingness (Graham 2012). Their exclusion may corrupt temporal structures such as autocorrelation, trends, and seasonality (Junger and Leon 2015). The third type, missing not at random (MNAR), happens when it is possible to make an unbiased estimation to model the missing data. When missingness is beyond the researcher's control (their distribution is unknown), MAR is the primary assumption (Graham 2012).

### 3.2 Methods for Handling Missing Aid Data

There are three different approaches to assess the missing data problem (Baraldi *et al*. 2015). First, we can remove the time intervals where there is at least one missing data for a specific date. Listwise deletion or the last observation carried forward scheme can make time series more fragmented or may introduce bias in the estimation of the parameters unless there is a chance that our missingness is MCAR (Kang 2013).

The second approach substitutes the missing data with an unconditional mean value or median (for skewed data, suggested by Junger and Leon 2015) of the available historical data. It has an impact similar to the last observation carried forward scheme until data is converted to differentials for future calculation and the returns reach a standard level (like zero mean and mode in financial time series). This solution is not recommended by Graham (2012) due to its distortions, since it results in a higher concentration around the mean and underestimates errors and variance at MCAR states (Junger and Leon 2015).

Third, some computation based approaches reconstruct missing data through the minimisation of an error function derived from mean, variance or a likelihood ratio (Baraldi et al. 2015, Ceylan *et al*. 2013, Juan Carlos 2010). Expectation maximisation (EM) models apply maximum likelihoods to estimate the variance and covariance matrices of the data, while neural networks-based and genetic structure-based approaches (Ceylan *et al*. 2013, Juan Carlos 2010) are also available. The expectation maximisation process required more computation time in the past (Ruud 1991), and it needs a well specified data generation model (Houari *et al*. 2013), not relying on the MCAR assumption. "Unbiasedness under MAR and higher efficiency under MCAR make maximum likelihood the method of choice in a situation with incomplete multinormal data" (Wothke 1998:19). This approach provides less biased data for listwise or pairwise deletion and mean-imputation methods; however, we should note that this advantage depends on the missing data rate, the covariance structure of the data and the size of the sample (Wothke 1998).

The present paper applies and compares two of the three mainstream missing value handling methods to capture their ability to influence panel regression coefficients. Let us assume $n$ aid data (1), where country $i$ ($1 \leq i \leq n$) has the following *d* value for every *y* year with *v* sample size:

$$D_i = \begin{bmatrix} y_1 & d_{i,1} \\ \dots & \dots \\ y_v & d_{i,v} \end{bmatrix}. \tag{1}$$

There is also a *k*th ($1 \leq k \leq n, and\ k \neq i$) country (2) with $w$ data, and $z$ ($z \neq y$) time indices:

$$D_k = \begin{bmatrix} z_1 & d_{k,1} \\ \dots & \dots \\ z_w & d_{k,w} \end{bmatrix}. \tag{2}$$

Upper $D_{1,..,i,k,...n}$matrices should be united for purposes of multivariate analysis which requires the synchronisation of time indices.

Listwise deletion (3) means a $T$ cap of specific time indices to exclude all cases where at least one value is missing:

$$T = Y \cap Z. \tag{3}$$

The effectiveness of this approach can be limited in case of a great number of missing data, with an empty *T* matrix as a result.

The Last Observation Carried Forward (LSCF) (4) scheme replaces missing data with the last available data:

$$T = (Y \cup Z)\ with\ d_{i,o} = d_{i,o-1}. \tag{4}$$

The LSCF procedure requires the addition of a very small positive $\varepsilon = 10^{d}$, $d \rightarrow +\infty$ number to satisfy the $e^{p_{i,o}-p_{i,o-1}} \neq 0$ requirement for a $p_{i,o} = p_{i,o-1}$case if we would like to use logarithmic returns. The inclusion of ε will provide an asymptotical result $e^{\varepsilon+p_{i,o}-p_{i,o-1}} \neq 0$ for $p_{i,o} = p_{i,o-1}$cases as well: $e^{\varepsilon} \approx 0$.

Regularised expectation-maximisation (EM) algorithm is based on iterated linear regression analyses, but it replaces the conditional maximum likelihood estimation of regression parameters for Gaussian data (5), following Schneider (2001). For each $d_{t,i} \in D$ with missing values, the relationship between the available and missing values of data matrix is modelled by a linear regression model:

$$d_{NaN} = \mu_{NaN} + (d_a - \mu_a)B + \varepsilon \tag{5}$$

where $a$ represents available data, and $B \in \Re^{n_a \times n_{NaN}}$ is a matrix of regression coefficients with a covariance matrix with missing and available data from $n$ all sample markets. The $\varepsilon \in \Re^{1 \times n_{NaN}}$ residual is assumed to be a zero-mean and $C \in \Re^{n_{NaN} \times n_{NaN}}$ is an unknown covariance matrix vector. In each iteration of the EM algorithm, estimates of the mean $\mu \in \Re^{1 \times n}$ and of the $\Sigma \in \Re^{n \times n}$ covariance matrix are taken as given, and from these estimates, the conditional maximum likelihood estimates of the matrix of regression coefficients B and of the covariance matrix C of the residual are computed for each record with missing values, in order to fill each missing value with imputed values, before recomputation of the entire $\mu$ vector and $\Sigma$ matrix. Then, the estimated regression coefficients will be the product of the two (missing-missing and available-missing) estimated covariance matrices: $\hat{B} = \widehat{\Sigma_{aa}^{-1}} \widehat{\Sigma_{aNaN}}$ to estimate the residual covariance matrix later. However, the regularised EM algorithm for each record with missing values uses $\hat{B} = \left(\widehat{\Sigma_{aa}} + h^2 Diag(\widehat{\Sigma_{aa}})\right)^{-1} \widehat{\Sigma_{aNaN}}$ with a $h$ regularisation parameter to inflate diagonal elements with a $1 + h^2$ factor.

### 3.3 Panel Regression

Panel regression requires consistent, balanced and fixed database to group country- and time-specific effects, and to manage heterogeneity that can or cannot be observed (Park 2011). The first generation of panel unit root tests like Im, Pesaran and Shin (2003) test requires the cross-sectional independence, with individual effects and no time trend:

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta y_{i,t-z} + \varepsilon_{i,t}$$

Null hypothesis: $\rho_i$=0 for all i=1,…,N and alternative hypothesis is $\rho_i$<0 for i=1,…,$N_1$ and $\rho_i$ =0 for i=$N_1$+1,…,N, with 0<$N_1$ ≤N alternative hypothesis allows for some (but not all) of the individual series to have unit roots. This test uses separate unit root tests for each cross-section units based on the (augmented) Dickey-Fuller statistics averaged across groups (Hurlin and Valérie 2007).

The present paper applies the Panel Data Toolbox,[2] following Alvarez, Barbero and Zofio (2015). Panel data (6) contains data matrices (with $i$ columns and $t$ rows) that were observed over a long period of time with $y$ dependent and $X$ independent variables with the following representation:

$$y_{it} = \alpha + \beta X_{it} + \mu_i + v_{it}, \text{ i=1,…, n, t=1,…, } T_i. \quad (6)$$

where $\mu_i$ represents the i-th invariant time individual effect (or unobserved component, latent variable, and unobserved heterogeneity) and $v_{it} \sim i.i.d(0, \theta_v^2)$

[2] http://www.paneldatatoolbox.com

refers to the disturbance (or idiosyncratic errors or idiosyncratic disturbances, because these change across t as well as across i) with the following properties: $E(v_i) = 0, E(v_i v_j) = 0, E(v_i v_i^T) = \theta_v^2 I_T$ for $i \neq j$, $I_T$ being the $T \times T$ identity matrix. In panel data models $\mu_i$ is called as a ''random effect'' when it is assumed as a random variable and a ''fixed effect'' when it is treated as a parameter to be estimated for each cross section observation $i$. It means that fixed effect approach allows arbitrary correlation between the unobserved effect $\mu_i$ and the observed explanatory variables $X_{it}$. Fixed effects analysis is more robust than random effects analysis, but time-constant factors cannot be included as $X_{it}$ – this approach is for time-varying explanatory variables (Wooldridge 2010).

The classical least squares model contains random error as a sole random component; all other effects are assumed to be fixed constants (Rawlings *et al*. 1998). Fixed and random effects models were used in this paper to compare the impacts of different missing data handling methods on panel regression coefficients. Under standard fixed effect specifications, individual effects are correlated with the explanatory variables ($COV(X_{it}, \mu_i) \neq 0$), their inclusion results in a biased OLS (ordinary least squares) estimation. To avoid such a bias, the within estimator of the parameters (7) – taking into account the variations in each group – is computed using OLS:

$$\hat{\beta}_{fe} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y} \tag{7}$$

where "within" estimator $\tilde{y} = y - \bar{y}$ and $\tilde{X} = X - \bar{X}$ are transformed variables to represent deviations from the group means $\bar{y}$ and $\bar{X}$ (unbiased and consistent for $n \to \infty$). Statistical inference (checked by the standard $t$ and $F$ tests) is generally based on the asymptotic variance-covariance matrix (8):

$$VAR(\widehat{\beta_{fe}}) = \frac{(\tilde{y} - (\tilde{X}\widehat{\beta_{fe}}))^T (\tilde{y} - (\tilde{X}\widehat{\beta_{fe}}))}{(nk) - n - k} \tilde{X}^T \tilde{X}^{-1}, \tag{8}$$

where *n* denotes the elements of the panel (*countries*), *k* represents time (*years*).

The individual effects, with their standard errors and significance test, can be computed as follows:

$$\hat{\mu} = \bar{y} - \bar{X}\beta, \tag{9}$$

$$VAR(\mu_i) = \frac{\tilde{\sigma}_v^2}{T_i} + \bar{X} VAR(\hat{\beta}) + \bar{X}'. \tag{10}$$

In the general panel data model (6), the loss of degrees of freedom can be avoided if the individual effects can be assumed random, where the error component $u_{it} = \mu_i + v_{it}$ includes the *i*-th invariant time individual effects $\mu_i$ and the disturbance $v_{it}$ ($\mu_i$ is assumed independent of the $v_{it}$ as well as they are independent of the explanatory variables: COV($X_{it}, \mu_i$)=0 and COV($X_{it}, v_{it}$)=0 for all *i* and *t*).

$$y_{it} = \alpha + X_{it}\beta + u_{it}, i = 1, \ldots, n \text{ and } t = 1, \ldots, T_i \quad (11)$$

The random effects model (11) is an appropriate specification in the analysis of large *n* number of individuals, randomly drawn from a large population. The composed error component has the following properties:

$$E(\mu_i) = E(v_{it}) = E(\mu_i v_{it}) = 0, \quad (12)$$

$$E(\mu_i \mu_j) = \begin{cases} \sigma_\mu^2 \; i \neq j \\ 0 \; i = j \end{cases} \; E(v_i v_j) = \begin{cases} \sigma_v^2 \; i \neq j \\ 0 \; i = j \end{cases}. \quad (13)$$

The block-diagonal covariance matrix can have serial correlation over time only between disturbances of the same individual, otherwise it is zero:

$$COV(u_{it} u_{js}) = \begin{cases} \sigma_\mu^2 + \sigma_v^2 \; i = j \; t = s \\ \sigma_\mu^2 \; i = j \; t \neq s \end{cases} \quad (14)$$

The GLS (generalized least squares) method yields an efficient estimator of the parameters:

$$\hat{\beta}_{re} = (X^T(\frac{1}{(T\sigma_\mu^2 + \sigma_v^2)P} + \frac{1}{\sigma_v^2 Q})^{-1}X)^{-1}X^T(\frac{1}{(T\sigma_\mu^2 + \sigma_v^2)P} + \frac{1}{\sigma_v^2 Q})^{-1}y = (\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T y \quad (15)$$

The P and Q are the matrices that compute the group means and the differences with respect to the group means. The asymptotic variance-covariance matrix will be similar to (8):

$$VAR(\widehat{\beta_{re}}) = \frac{(\tilde{y} - (\tilde{X}\widehat{\beta_{re}}))^T(\tilde{y} - (\tilde{X}\widehat{\beta_{re}}))}{(nk) - k}\tilde{X}^T\tilde{X}^{-1}, \quad (16)$$

Either fixed or random effect is an issue of unmeasured variables or omitted relevance variables-the main difference between these models is the relation of the individual specific error component to regressors (Kennedy 2008).

Several canonical tests should be done on the panel data regression models to identify serial correlation in the error term or to select the efficient estimator between fixed and random effects models – like the Hausman's test. The Hausman's test compares the GLS estimator of the random effects model $\widehat{\beta_{re}}$, and the within estimator in the fixed effects model $\widehat{\beta_{fe}}$, both of which are consistent under the null hypothesis ($H_0$: $\beta_{fe} - \beta_{re} = 0$). Under the alternative, only the GLS estimator of random effects is consistent. The computation is based on the difference between both estimators:

$$H = (\widehat{\beta_{fe}} - \widehat{\beta_{re}})'VAR(\widehat{\beta_{fe}} - \widehat{\beta_{re}})^{-1}(\widehat{\beta_{fe}} - \widehat{\beta_{re}}), \quad (17)$$

under the assumption of homoscedasticity:

$$VAR\left(\widehat{\beta_{fe}} - \widehat{\beta_{re}}\right) = VAR\left(\widehat{\beta_{fe}}\right) - VAR(\widehat{\beta_{re}}). \tag{18}$$

In applications where *n* is relatively large with respect to *T*, it can be used to choose between estimators. Fixed models are better under $p<0.05$ cases.

### 3.4 Final Model

Keeping in mind the purpose of the present study, we selected indicators which closely describe the US motivations behind its own aid allocation. For this purpose, we considered indices which appeared in relevant literature. As a result, we collected data on the following variables:

- Official Development Assistance provided by the United States as a dependent variable was collected from the OECD aid database[3] (OECD 2016). In our calculations, we used net disbursements which show the amount of aid paid in the reality. These amounts are paid on an agreement signed by the recipient country and the donor country. The agreement contains commitments on aid, but the real disbursements are mainly lower than the commitments. That is, disbursements show more precise picture on the processes.
- US exports to the selected recipient countries describe the economic and commercial interests of the United States. These data were mainly collected from the detailed database of the Observatory of Economic Complexity (OEC 2016); although, because of the growing number of missing data after 2010, the UNCTADStat database (UNCTADStat 2016) was also used for these last few years. The UNCTADStat contains bilateral trade data only from 1995 onwards. In order to control any distortion effect of using two databases, we checked whether there was any significant difference in data of the overlapping years (1995-2010). We experienced that the data from the two databases were very close to each other.
- GDP and GDP per capita data were collected from the World Bank database. GDP per capita could be a proxy for poverty (as Fleck and Kelby 2010 described).

[3]The most studies on aid effectiveness are based on data collected from the OECD's Creditor Reporting System (e.g. Ahmad *et al*. 2014, Cali and te Velde 2011, or Vijil and Wagner 2012), but there is another database for aid: AidData which is mainly used for the analysis of emerging donors.

- Democracy score describes the democracy status of the recipient country. We used Polity IV as an appropriate indicator: it is a combined polity score computed by the autocratic and democratic scores of a country; the scale ranges from +10 (strongly democratic) to -10 (strongly autocratic). There are several studies in which the same indicator was used (see, for instance, Fleck and Kelby 2010 or Vijil and Wagner 2012).
- The number of armed conflicts in recipient countries refers to the war on terrorism aspect. To collect the data, we used the UCDP/PRIO database (UCDP/PRIO 2016): we added the number of conflicts in the years concerned in the analysis.

The present study aims to analyse the motivations behind the aid allocation preferences of the United States. From this point of view, the readiness and real need of accepting aid in a recipient country would be also worth analysing: share of people living in absolute poverty, income inequality, health issues (epidemics, hospitals, or doctors), needs in education (illiteracy rate, enrolment in primary education), and so on. However, in the long-run, there are several missing data in the case of these social indicators, and in order to cover a relatively large sample, we decided to leave these indicators out of the analysis. If one decides to analyse a smaller number of countries, these indicators can be involved, too.

In our analysis, we aimed to involve as many recipient countries as possible and take the longest time period into account. The OECD database enabled us to collect US aid data from 1960 to 2014 covering 97 countries; however, unfortunately, we had to exclude some countries from the analysis and narrow the time period due to the lack of data regarding other indices. Although there are many methods to handle the missing data problem (as described in the previous part), we concentrated on the missing data of the dependent variable (aid) and avoided making artificial data of independent variables because of the relatively high risk of its distorting effects on final results. Finally, 60 developing countries were involved in the empirical analysis and the time series covered the years between 1966 and 2014. This time series of nearly 50 years is appropriate to analyse missing data problems.

Furthermore, we followed the standard analysis method in aid investigations (see, for instance, Fleck and Kelby 2010, Wagner 2003, Cali and te Velde 2011, Vijil and Wagner 2012): we calculated the natural logarithm of all variables in order to reduce the influence of outliers. The final model is as follows:

$$logUS\,aid_{it} = \alpha + \beta \left[\text{logUS export}, \text{logGDP}, \log\frac{\text{GDP}}{\text{capita}}, \text{Polity}, \text{armed conflict}\right]_{it-1} + \mu_i + v_{it}, \quad (19)$$

where the dependent variable, $USaid_{it}$ refers to the US aid (official development assistance) allocated to country *i* in year *t*. In the model, the *USexport* shows the exports of the United States into recipient country *i* in year *t-1*; the GDP and GDP per capita indicators are features of the recipient country *i* in year *t-1*; *Polity* refers to the level of democracy in country *i* in year *t-1*, while a*rmed conflict* variable counts the number of civil and armed conflicts in country *i* in year *t-1*. In the model, $\mu_i$ represents the i-th invariant time individual effect and $v_{it} \sim i.i.d(0, \theta_v^2)$ refers to the disturbance as equation (6) presented.

Endogeneity may distort the results of regression models. In empirical analyses on aid allocation, this challenge is handled in three different ways:

1. involvement of instrumental variable into the analysis (Acemoglu *et. al*. 2001, Angeles and Neanidis 2009, Roodman 2007);
2. calculation of averages following Vijil and Wagner (2012);
3. most frequent method: use of lagged data (Cali and te Velde 2011, Kimura and Todo 2010, Wagner 2003, Younas 2008), but there is no consensus on the extent of the lag (Doucouliagos and Paldam 2009).

We followed the most frequent method and used 1-year lagged data in our analysis. Its economic sense is that the US aid allocation was determined by the economic performance of previous years in the recipient countries.

## IV. RESULTS

Seven per cent of the US Official Development Assistance data was missing between 1967 and 2014 (N= 2,784), and 47 per cent of the sample countries were affected by this issue. This temporary suspension of data can be originated from non-transparent aid decisions, or from lack of reporting to the OECD. Data from the 1960s and 1970s suffer more from missingness, so the phenomenon has a temporal property. We applied two methods: Last observation carried forward (LOCF) and EM methods.

Comparing the two methods with t-test (Table II), we expected that LOCF and EM methods would provide similar outputs: they were not significantly different in their first four moments according to the t-test.

TABLE II

**MISSING VALUE DIFFERENCES IN THE FIRST FOUR MOMENTS**

| | mean | standard deviation | skewness | kurtosis |
|---|---|---|---|---|
| t-test (p) | 0.34 | 0.32 | 0.35 | 0.37 |

**Source:** Authors' calculation.

Unit root was rejected by the Im, Pesaran and Shin Panel Unit Root Test (see Appendix 1). Decision on fixed or random effects was based on the Hausman's test (Table III). In the context of the LOCF model, fixed effect models are proved to be more valid in the test, despite the fact that armed conflicts were significant in the random effect model.

TABLE III

**HAUSMAN'S TEST OF SPECIFICATION ON LOCF DATA**

| Varname | A:FE | B:RE | Coef. Diff | S.E. Diff |
|---|---|---|---|---|
| logUS_export | 0.067860 | 0.156730 | -0.088869 | 0.004387 |
| logGDP | 2.405092 | 1.403750 | 1.001342 | 0.063011 |
| logGDP/capita | -2.886051 | -1.678252 | -1.207799 | 0.076868 |
| Polity | -0.000973 | -0.000933 | -0.000040 | 0.000000 |
| armed_conflict | 0.068831 | 0.129183 | -0.060352 | 0.000000 |

**Note:** A is consistent under H0 and H1 (A = FE);B is consistent under H0 (B = RE);H0: coef(A) - coef(B) = 0;H1: coef(A) - coef(B) != 0;H = 255.134992 ~ Chi2(5);p-value = 0.0000.

**Source:** Authors' calculation.

According to the results presented in Table IV, the long term aid policy of the United States between 1967 and 2014 was motivated largely by the GDP of the recipient country (as a proxy of the market), while the level of the GDP per capita had a decreasing effect on it, that is, the richer a recipient country is, the less aid is provided to it. Surprisingly, political regimes or armed conflicts had no significant impact on these decisions, although the sample was largely dominated by cold war or war on terror periods.

TABLE IV

**PANEL: FIXED EFFECTS (WITHIN) (FE) ON LOCF DATA**

| logUS_aid | Coefficient | Std. Error | t-stat | p-value |
|---|---|---|---|---|
| constant | −9.2521 | 0.9872 | −9.372 | 0.000 *** |
| logUS_export | −0.0241 | 0.0271 | −0.8903 | 0.3734 |
| logGDP | 0.6748 | 0.0756 | 8.924 | 0.000 *** |
| logGDP/capita | −0.7402 | 0.0915 | −8.085 | 0.000 *** |
| Polity | −0.0010 | 0.0009 | −1.097 | 0.2726 |
| armed_conflict | 0.0593 | 0.0465 | 1.277 | 0.2016 |
| logUS_aid t-1 | 0.7472 | 0.0127 | 58.64 | 0.0000 *** |

**Note:** N = 2,784 n = 58 T = 48 (Balanced panel), R-squared = 0.8471, Adj R-squared = 0.6838, Wald F(5, 2721) = 234.0345, p-value = 0.0000, Durbin-Watson 2.188032.

**Source:** Authors' calculation.

Fixed effects models were not affected by serial correlation as the nearly 2 value of Durbin-Watson test suggested.

In the context of the second model (EM), fixed effect models proved to be more valid by the Hausman's test in each case for EM data as well, despite the fact that armed conflicts were significant in the random effect model (Table V). This result is similar to those of the LOCF model.

TABLE V

**HAUSMAN'S TEST OF SPECIFICATION ON EM DATA**

| Varname | A:FE | B:RE | Coef. Diff | S.E. Diff |
|---|---|---|---|---|
| logUS_export | 0.0678 | 0.1567 | -0.0889 | 0.0044 |
| logGDP | 2.4050 | 1.4038 | 1.001 | 0.0630 |
| logGDP/capita | -2.886 | -1.678 | -1.2078 | 0.0769 |
| Polity | -0.001 | -0.001 | -0.00004 | 0.0000 |
| armed_conflict | 0.0688 | 0.1292 | -0.0603 | 0.0000 |

**Note:** A is consistent under H0 and H1 (A = FE);B is consistent under H0 (B = RE);H0: coef(A) - coef(B) = 0;H1: coef(A) - coef(B) != 0;H = 255.134992 ~ Chi2(5);p-value = 0.0000.

**Source:** Authors' calculation.

As for the results presented in Table VI, we see that the long term aid policy of the United States between 1967 and 2014 was largely motivated by the GDP, at the same time the GDP/capita had a reduction effect. That is, the results are the same as the results of the LOCF model. However, the EM data decreased the importance of coefficients: for logGDP with 0.5, for the logUS export with 0.01 and for logGDP/capita with 0.12. It also provided high R-square by 0.8209, indicating a relatively strong dependency between the variables and the previous value of the aid.

TABLE VI

**PANEL: FIXED EFFECTS (WITHIN) (FE) ON EM DATA**

| US_aid | Coefficient | Std. Error | t-stat | p-value |
|---|---|---|---|---|
| constant | −10.1577 | 1.04272 | −9.742 | 0.000 *** |
| logUS_export | −0.0069 | 0.0288 | −0.2410 | 0.8095 |
| logGDP | 0.7329 | 0.0801 | 9.144 | 0.000 *** |
| logGDP/capita | −0.8309 | 0.0969 | −8.572 | 0.000 *** |
| Polity | −0.0010 | 0.0010 | -0.9910 | 0.3218 |
| armed_conflict | 0.0705 | 0.0494 | 1.428 | 0.1533 |
| logUS_aid t-1 | 0.6969 | 0.0137 | 50.72 | 0.0000 *** |

**Note:** Durbin-Watson 2.216347, N = 2,784 n = 58, T = 48 (Balanced panel), R-squared = 0.8209, Adj R-squared = 0.6229, Wald F(5. 2721) = 193.6990, p-value = 0.0000.

**Source:** Authors' calculation.

Similar to the LOCF model, the fixed effects model of EM was not affected by serial correlation as the nearly 2 level of Durbin-Watson test suggested.

In conclusion, in the case of US aid allocation, the results strengthen the view that there is no significant difference between the two methods of handling missing data (LOCF and EM), thus, in long term analyses both methods offer reliable results without a significant distortion effect.

These results can be used by policy makers, too. It is always an important issue to analyse the motivations and effectiveness of foreign aid not only in the short term but also in the long run. Aid can have impacts on the long-run and the impacts may appear years later than aid was really provided. As a result, long-run analyses may provide more precise information on the effectiveness of aid. Furthermore, analysing long-term effectiveness is important from the point of view of accountability and transparency: a government should provide precise information for the public (and tax-payers) how aid was spent and what its main effects are. Furthermore, the long-run analyses of motivations behind aid allocation may improve aid effectiveness with considering the changes in motivations over the decades.

## V. CONCLUSIONS

The problems of developing countries like poverty, migration, terrorism, and others mean a huge challenge not only for underdeveloped states but also for developed ones. To handle this global issue, promoting development in the third world would be an essential objective for the international community. One of the several international initiatives to support developing countries to catch up is international development cooperation and allocation of foreign aid to them. The system of international development cooperation has been operating for almost sixty years, but the effectiveness of foreign aid is still a central question and widely analysed with different methods. In most cases, it is the missing data that make long-term analyses difficult; hence, how to investigate the changes in the motivations of donor countries, the long-term economic effects of aid or other aid-related issues is not always evident for researchers.

The purpose of this study was to compare two different methods of handling the missing data problem and to answer the question whether there is any difference between the results achieved with the two methods. To answer the question, the United States was an example: the motivations behind US aid allocation were analysed between 1966 and 2014.

This paper applied the panel regression model and compared two of the three mainstream missing value handling methods. As the results show, the long term aid policy of the United States between 1966 and 2014 was motivated largely by the GDP of the recipient country, while the level of the GDP per capita had a decreasing effect on it. That is, economic issues were more important for the United States than political ones, since political regimes or armed conflicts had no significant impact on aid decisions of the United States. Furthermore, concerning the possible divergence between the two methods of handling missing data, the results also show there is no significant difference between them, thus, in long term analyses both methods may offer reliable results without any significant distortion effect.

However, the results raised more questions which still need to be investigated. The present paper only analysed the motivation side, but the long-term aid effectiveness should be analysed, too: how the aid affects long-term poverty, GDP, education, and the overall development of a recipient country. Furthermore, aid volatility should be analysed. Since we could show that the problem of missing aid data can be overcome with different methods without any distorting effects, the opportunities for long-term analysis are now available.

## REFERENCES

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 5: 1369-1401.

Ahmad, Khalil, Amjad Ali and Muhammd Irfan Chani. 2014. "Does Foreign Aid to Social Sector Matter for Fertility Reduction? An Empirical Analysis for Pakistan." *Bangladesh Development Studies*, 37(4): 65-76.

Alesina, Alberto and David Dollar. 1998. "Who Gives Foreign Aid to Whom and Why?" *Working Paper*, 6612. National Bureau of Economic Research, Cambridge.

Álvarez, Inmacualda C., Javier Barbero and José L. Zofio. 2015. "A Panel Data Toolbox for MATLAB." *Economic Analysis Working Paper Series.* Universidad Autonoma Madrid Working Paper 05/2013, ISSN: 1885-6888

Angeles, Luis and Kyriakos C. Neanidis. 2009. "Aid Effectiveness: The Role of the Local Elite." *Journal of Development Economics*, 1:120-134.

Balla, Eliana and Gina Yannitell Reinhardt. 2004. "Giving and Receiving Foreign Aid: Does Conflict Count?" *World Development*, 36(12): 2566-2585.

Bangoura, Lansana. 2012. "Microfinance as an Approach to Development in Low Income Countries." *Bangladesh Development Studies*, 35(4): 87-111.

Baraldi, Piero, F. Di Maio, D. Genini and Enrico Zio. 2015. "Reconstruction of Missing Data in Multidimensional Time Series by Fuzzy Similarity." *Applied Soft Computing Journal*, 26: 1–9. http://dx.doi.org/10.1016/j.asoc.2014.09.038

Berthelemy, Jean-Claude and Ariane Tichit. 2004. "Bilateral Donors' Aid Allocation Decisions—a Three-Dimensional Panel Analysis." *International Review of Economics and Finance*, 13(3): 253-274.

Brück, Tilman and Guo Xu. 2012. "Who Gives Aid to Whom and When? Aid Accelerations, Shocks and Policies." *European Journal of Political Economy*, 28(4): 593-606.

Cali, Massimiliano and Dirk Willem te Velde. 2011. "Does Aid for Trade Really Improve Trade Performance?" *World Development*, 5: 725-740.

Ceylan, Yozgatligil, Aslan Sipan, Iyigun Cem and Batmaz Inci. 2013. "Comparison of Missing Value Imputation Methods in Time Series: The Case of Turkish Meteorological Data." *Theoretical & Applied Climatology*, 112(1-2): 143–167.

Charron, Nicholas. 2011. "Exploring the Impact of Foreign Aid on Corruption: Has the "anti-Corruption Movement" been Effective?" *The Developing Economies*, 49(1): 66-88.

Chong, Alberto and Mark Gradstein. 2008. "What Determines Foreign Aid? The Donors' Perspective." *Journal of Development Economics*, 87(1): 1-13.

Clist, Paul. 2011. "25 Years of Aid Allocation Practice: Whither Selectivity?" *World Development*, 39(10): 1724-1734.

Collier, Paul and David Dollar. 2001. "Can the World Cut Poverty in Half? How Policy Reform and Effective Aid can Meet International Development Goals." *World Development*, 29(11): 1787-1802.

Dhasmana, Anubha. 2010. "Welfare Gains of Aid Indexation in Small Open Economies." *The Developing Economies*, 48(2): 247-276.

Dollar, David and Victoria Levin. 2006. "The Increasing Selectivity of Foreign Aid, 1984–2003." *World Development*, 34(12): 2034-2046.

Doucouliagos, Hristos and Martin Paldam. 2009. "The Aid effectiveness Literature: the Sad Results of 40 Years of Research." *Journal of Economic Surveys*, 23(3): 433-461.

Einarsdottir, Jónína and Geir Gunnlaugsson. 2016. "Applied Ethics and Allocation of Foreign Aid: Disparity in Pretensions and Practice." *Development Policy Review*, 34(3): 345-363.

Fleck, Robert K. and Christopher Kilby. 2010. "Changing Aid Regimes? US Foreign Aid from the Cold War to the War on Terror." *Journal of Development Economics*, 91(2): 185-197.

Gates, Scott and Anke Hoeffler. 2004. "Global Aid Allocation: Are Nordic Donors Different?" *CSAE WPS/2004-34.* Centre of the Study of African Economies, University of Oxford and International Peace Research Institute, Oslo (PRIO).

Gohou, Gaston and Soumaré Issouf. 2012. "Does Foreign Direct Investment Reduce Poverty in Africa and are There Regional Differences?" *World Development*, 40( 1):637-651.

Graham, John W. 2012. *Missing Data Analysis and Design.* New York: Springer.

Hansen, Henrik and Finn Tarp. 2001. "Aid and Growth Regressions." *Journal of Development Economics*, 64(2):547-570.

Harrigan, Jane and Chengang Wang. 2011. "A New Approach to the Allocation of Aid Among Developing Countries: is the USA more Selfish than the Rest?" *World Development*, 39(8): 1281-1293.

Hoeffler, Anke and Verity Outram. 2011. "Need, Merit or Self-Interest – What Determines the Allocation of Aid?" *Review of Development Economics*, 15(2): 237-250.

Hossain, Monzur. 2015. "Capital Flows to Least Developed Countries: What Matters?" *Bangladesh Development Studies*, 38(2).

Houari, Rima, Ahcéne Bounceur, Tahar Kechadi, Tari Abdelkamel and Reinhardt Euler. 2013. "A New Method for Estimation of Missing Data Based on Sampling Methods for Data Mining." CCSEIT http://dx.doi.org/10.1007/978-3-319-00951-3_9

Hurlin, Christophe and Mignony Valérie 2007. *"Second Generation Panel Unit Root Tests."* Mimeo, University of Paris X.

Im, K.S. and M. H. Pesaran. 2003. *"On the Panel Unit Root Tests Using Nonlinear Instrumental Variables."* Mimeo, University of Southern California.

Juan Carlos, Figueroa García, Kalenatic Dusko and Cesar Amilcar López Bello. 2010. "An Evolutionary Approach for Imputing Missing Data in Time Series." *Journal of Circuits, Systems & Computers*, 19(1):107–121.

Junger, Washington and Antonio Ponce de Leon. 2015. "Imputation of Missing Data in Time Series for Air Pollutants." *Atmospheric Environment*, 102: 96–104. http://dx.doi.org/10.1016/j.atmosenv.2014.11.049

Kang, Hyun. 2013. "The Prevention and Handling of the Missing Data." *Korean Journal of Anesthesiology*, 64(5): 402–406. http://dx.doi.org/10.4097/kjae.2013.64.5.402

Kangoye, Thierry. 2013. "Does Aid Unpredictability Weaken Governance? Evidence from Developing Countries." *The Developing Economies*, 51( 2):121-144.

Kennedy, Peter. 2008. "*A Guide to Econometrics."* 6th ed. Malden, MA: Blackwell Publishing.

Kimura, Hidemi and Yasuyuki Todo. 2010. "Is foreign aid a vanguard of foreign direct investment? A Gravity-equation Approach." *World Development*, 4: 482-497.

Masaki, Takaaki. 2016. "Coups d'État and Foreign Aid." *World Development*, 79 : 51-68.

McGillivray, Mark. 2003. "Modelling Aid Allocations: Issues, Approaches and Results." *Journal of Economic Development*, 28(1): 171-188.

OECD CRS. 2016. *Query Wizard for International Development Statistics*. http://www.oecd.org/dac/stats/crsguide.htm (accessed: May 20, 2016).

Park, Hun Myoung. 2011. "Practical Guides To Panel Data Modeling: A Step-by-Step Analysis Using Stata." Tutorial Working Paper. Graduate School of International Relations, International University of Japan.

Rawlings, John O., Sastry G. Pantula and David A. Dickey. 1998. *Applied Regression Analysis*. Springer.

Reinsberg, Bernhard. 2014. "Foreign Aid Responses to Political Liberalization." *World Development*, 75: 46-61.

Roodman, David. 2007. "Macro Aid Effectiveness: A Guide for the Perplexed." *Working Paper*, 134. Centre for Global Development.

Round, Jeffery I. and Matthew Odedokun. 2004. "Aid Effort and Its Determinants." *International Review of Economics & Finance*, 13(3): 293-309.

Ruud, Paul A. 1991. "Extensions of Estimation Methods Using the EM Algorithm." *Journal of Econometrics*, 49(3):305–341. http://dx.doi.org/10.1016/0304-4076(91)90001-T

Schneider, Tapio. 2001. "Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values." *Journal of Climate*, 14: 853–871. http://dx.doi.org/10.1175/1520-0442(2001)014%3C0853:AOICDE%3E2.0.CO;2

Szent-Iványi, Balázs and Simon Lightfoot. 2015. *New Europe's New Development Aid.* Routledge, Abingdon – New York.

Tanaka, Kiyoyasu and Kenmei Tsubota. 2013. "Does Aid for Roads Attract Foreign or Domestic Firms? Evidence from Cambodia." *The Developing Economies*, 51(4): 388-401.

Tarnoff, Curt – Nowels, Larry (2004): Foreign Aid: An Introdutory Overview of U.S. Programs and Policy. http://www.au.af.mil/au/awc/awcgate/crs/98-916.pdf (Accessed: June 5, 2016)

Tingley, Dustin. 2010. "Donors and Domestic Politics: Political Influences on Foreign Aid Effort." *The Quarterly Review of Economics and Finance*, 50(1): 40-49.

Udvari, Beáta. 2014. "Impacts of Aid for Trade on Trade with the EU: The Role of Old and New Member States." *Journal of Global Policy and Governance*, 3(1): 77-93.

UN. 2002. *Monterrey Consensus on Financing for Development. International Conference on Financing for Development*. United Nations, 18-22 March.

USAID. 2014. *Ending Extreme Poverty. Agency Financial Report Fiscal Year 2014*. http://www.usaid.gov/results-and-data/progress-data/agency-financial-report (accessed June 6, 2016)

USAID. 2015. *Agency Financial Report Fiscal Year 2015. Ending Extreme Poverty in This Generation.* https://www.usaid.gov/results-and-data/progress-data/agency-financial-report (accessed June 6, 2016).

Vijil, Mariana and Wagner Laurent. 2012. "Does Aid for Trade Enhance Export Performance? Investigating the Infrastructure Channel." *The World Economy*, 35(7): 838-868.

Wagner, Don. 2003. "Aid and Trade – An Empirical Study." *Journal of Japanese and International Economies*, 2:153-173.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.

Wothke, Werner. 1998. *Longitudinal and Multi-group Modelling with Missing Data*. Mahwah. NJ: Lawrence Erlbaum Associates.

Younas, Javed. 2008. "Motivation for Bilateral Aid Allocation: Altruism or Trade Benefits." *European Journal of Political Economy*, 24(3):661-674.

ANNEX 1

**IM, PESARAN AND SHIN (2003) PANEL UNIT ROOT TEST RESULTS ON BALANCED PANEL**

Aid (LOCF) data:

P-value of the W_bar statistic = 0.0000

P-value of the Z_bar statistic = 0.0000

P-value of the Z_bar_DF statistic = 0.0000

Lag = 1 Adj. sample size = 2782 ADF statistic = -10.7310 ADF p-value = 0.0100

Aid (EM) data:

P-value of the W_bar statistic = 0.0000

P-value of the Z_bar statistic = 0.0000

P-value of the Z_bar_DF statistic = 0.0000

Lag = 2 Adj. sample size = 2781 ADF statistic = -10.1174 ADF p-value = 0.0100

lUS_export, lGDP, lGDP/capita, Polity, armed conflict data:

P-value of the W_bar statistic = 0.0000

P-value of the Z_bar statistic = 0.0000

P-value of the Z_bar_DF statistic = 0.0000

Cross-unit lUS_export:

Lag = 0 Adj. sample size = 2783 ADF statistic = -8.8630 ADF p-value = 0.0100

Cross-unit lGDP:

Lag = 0 Adj. sample size = 2783 ADF statistic = -8.7252 ADF p-value = 0.0100

Cross-unit lGDP/capita:

Lag = 0 Adj. sample size = 2783 ADF statistic = -9.5289 ADF p-value = 0.0100

Cross-unit Polity:

Lag = 2 Adj. sample size = 2781 ADF statistic = -17.1571 ADF p-value = 0.0100

Cross-unit armed conflict:

Lag = 4 Adj. sample size = 2779 ADF statistic = -10.0910 ADF p-value = 0.0100