

# Methodology for measuring and estimating funding to data and statistics

## Technical note for the Partner Report on Support to Statistics

### Contents

1. Background .....	32
2. Monitoring funding to statistics with accuracy .....	32
2.1. OECD's Creditor Reporting System .....	32
2.2. PARIS21's annual online survey.....	64
3. Monitoring funding to statistics with reduced reporting lag .....	65
3.1. What is the reporting lag?.....	65
3.2. Estimating up-to-date support to statistics using CRS.....	86
3.2.1. Nowcasting: using commitments to predict current disbursements .....	97
3.2.2. Forecasting: anticipating future funding .....	129
3.3. Expanding the PRESS database .....	1311
3.3.1. Exploring alternative data sources for aid flows on statistics .....	1311
<i>The International Aid Transparency Initiative (IATI)</i> .....	1411
<i>Donors' transparency portals</i> .....	1411
<i>Multilateral donors' prospective reporting to the PRESS survey</i> .....	1411
3.3.2. Addressing gaps in the alternative data sources.....	1613
3.3.3. Linking the alternative sources: the new harmonised database .....	1714
3.4. Bringing them together – nowcasting and forecasting with the new harmonised database.....	1915
4. Conclusion .....	1915
References.....	2017

### Figures, Tables and Boxes

Figure 1. Flow of official aid in CRS .....	54
Figure 2. How the lag in the CRS data led to a lag in previous rounds of PRESS .....	75
Figure 3: Disbursements vs commitments reported in CRS.....	108
Figure 4: Disbursements vs Commitments in CRS reported by the UK .....	119

Figure 5: Comparison of PRESS and the new harmonised database for the share of projects in the final datasets, by sources of data, 2016-2018..... ~~18~~<sup>15</sup>

Table 1. Regression table from the analysis ..... ~~10~~<sup>9</sup>

Table 2: Comparison of data sources ..... ~~15~~<sup>13</sup>

Box 1. Identifying data and statistical projects in CRS using text analysis ..... ~~5~~<sup>5</sup>

Box 2: The Eurostat’s Donor Survey – and why it is not included in the PRESS database..... ~~16~~<sup>16</sup>

Draft methodological note prepared by Yu Tian and Archita Misra (PARIS21) under the supervision of Rajiv Ranjan (PARIS21). The authors are grateful to Eric Swanson and Lorenz Noe (Open Data Watch) and Simon Lange (OECD) for their valuable review and feedback. Further comments to the authors ([Yu.Tian@oecd.org](mailto:Yu.Tian@oecd.org); [Archita.misra@oecd.org](mailto:Archita.misra@oecd.org)) are welcome.

## 1. Background

PARIS21 produces the Partner Report on Support to Statistics (PRESS) annually to report on trends in support to statistics. The methodology is applied retrospectively for all previous years to ensure comparability over time. This document presents the methodology.

## 2. Monitoring funding to statistics with accuracy

This section provides information on how to monitor support to statistics and how the data of the PRESS, which is also used for reporting SDG indicator 17.19.1 (“Dollar Value of all resources made available to strengthen statistical capacity in developing countries”), is generated.

Prior to PRESS 2018, the PRESS only focused on borrowing countries of the International Development Association.<sup>1</sup> Since 2018, the PRESS covers the commitments received by all countries throughout the report to align the findings with the SDG indicator 17.19.1: “Dollar value of all resources made available to strengthen statistical capacity in developing countries”. Commitments were used as the main measurement instead of disbursements from the beginning of PRESS in 2006, when commitment data entries were made available more consistently. See section 3.2 for a discussion on using commitments vs disbursements in detail.

The PRESS aims to provide a full picture of international support to statistics. To achieve this goal, it mainly takes advantage of two data sources:

### 2.1. OECD’s Creditor Reporting System (CRS)

The Organisation for Economic Co-operation and Development (OECD)’s Creditor Reporting System (CRS) records data from OECD Development Assistance Committee (DAC) members (donors) and some non-DAC donors. This provides a comprehensive account of Official Development Assistance (ODA). Donors report to the CRS using specific codes for the sectors targeted by their aid activity. Statistical Capacity Building (SCB) is designated by the sector code 16062.<sup>2</sup> Each activity reported in CRS can only

---

<sup>1</sup> Eligibility for IDA support depends first and foremost on a country’s relative poverty, which is defined as GNI per capita below an established threshold and updated annually (\$1,185 in the fiscal year 2021). IDA also supports some countries, including several small island economies, that are above the operational cut-off but lack the required creditworthiness to borrow from the International Bank for Reconstruction and Development (IBRD). For more information, see: <http://ida.worldbank.org/about/borrowing-countries>

<sup>2</sup> Until 2019, this purpose was vaguely defined as “Both in national statistical offices and any other government ministries”. However, after a successful campaign to improve the description, this purpose is now defined as “All statistical activities,

be assigned with one of the over 100 purpose codes<sup>3</sup>. While CRS is one of the most reliable and comprehensive databases that accounts for aid flows, there are some concerns that need to be addressed while compiling support for data and statistics, such as: the cross-cutting nature of projects with statistical components, limited reporter knowledge about the code, the assignment of some ODA for data and statistics under other codes, lack of granularity in reporting, etc. These issues are covered in detail in PARIS (2019). PARIS21 is seeking to reduce this usually downward bias using a text analysis methodology (see [Box 1](#)~~Box 1~~).

---

such as data collection, processing, dissemination and analysis; support to development and management of official statistics including demographic, social, economic, environmental, and multi-sectoral statistics; statistical quality frameworks; development of human and technological resources for statistics, investments in data innovation”.

<sup>3</sup> In recent years, CRS reporters can also assign multiple voluntary purpose codes to the same project. Code 16062 is not a voluntary code. See the CRS code list for more information: <https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/dacandcrscodelists.htm>

### Box 1. Identifying data and statistical projects in CRS using text analysis

#### Step 1: Filtering project titles

Firstly, PARIS21 uses text analysis to search through titles of all CRS projects using a keyword list that contains statistical terms such as *census*, *survey*, *data*, *indicators*, *etc.* Because the project titles are usually concise, those which contain any of these keywords are identified as data and statistical projects.

#### Step 2: Filtering project descriptions (short)

Secondly, using a similar approach, PARIS21 searches through the short descriptions using a stricter keyword list, which contains terms such as *SDG data collection*, *census preparation*, *data analysis*, *etc.* This stricter standard is used to avoid “false-positive” descriptions that only mention data or statistical concepts for reference.

#### Step 3: Identification of false positives and blacklist

Similarly, throughout the first and the second steps, another list of keywords (a “blacklist”) is also used to avoid the identification of projects that are almost certainly not related to data and statistics. The most common example here is the landmine survey in conflict regions.

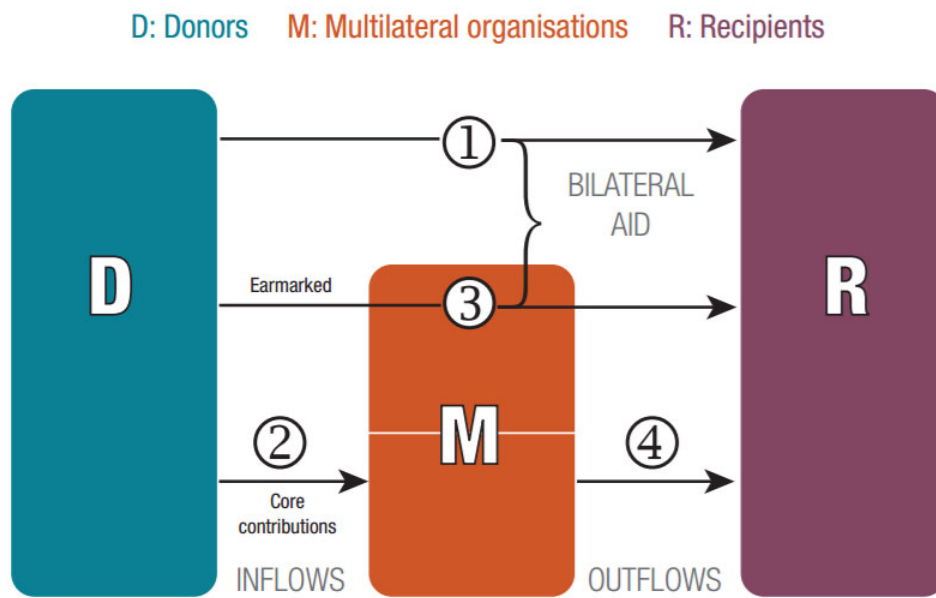
This process also helps to eliminate some monitoring programmes that do not contribute to the improvement of capacity in recipient countries, but which have been mistakenly marked as 16062 by reporters. However, for projects that have no or very simple descriptions, these mistakes may still lead to “false-positives” in the PRESS dataset.

#### Step 4: Filtering project descriptions (long) using machine learning

Finally, the projects identified in the first and second steps are combined with projects identified through the SCB purpose code. The machine learning approach is then used to summarise the key features of long descriptions in identified projects. The long descriptions of the rest of the projects are then compared against the identified features. Only descriptions that bear significant similarities with the identified projects are marked as data and statistical projects. Donors’ reporting patterns are taken into account in this process to mitigate the bias across different reporting agencies. The description pattern by the same donor will receive more weight during this searching process, while the global pattern will be examined against afterwards with less weights.

The CRS identifies a project donor by looking at the source of the funding. Countries are identified as donors if the flow is directly between them and the recipient country (type 1 in [Figure 1Figure 1](#)), or if the flow is earmarked for a certain project and channelled through multilateral organisations (type 3 in [Figure 1Figure 1](#)). If a project is funded by un-earmarked core contributions to multilateral organisations, the donors are marked as the multilateral organisations (types 2 and 4 in [Figure 1Figure 1](#)).

Figure 1. Flow of official aid in CRS



## 2.2. PARIS21's annual online survey

The PARIS21 Secretariat supplements the data from the CRS with an annual online survey that is completed by a global network of respondents, mostly non-DAC donors. The survey covers a subset of the variables collected in the CRS, as well as some additional variables specific to data and statistics. Responding to the online survey is voluntary and offers an opportunity for respondents to share information about their statistical activities. Respondents include non-DAC members, including non-DAC donor countries, multilateral organisations, regional statistical training institutes, and other philanthropic organisations. The percentage of these projects in the final PRESS database has decreased in recent years, as many multilateral organisations have improved the granularity of their reporting to the CRS, making these data equally useful as data collected from the PRESS survey. To reduce the burden on donors, these multilateral organisations are no longer required to fill in the PRESS survey.

## 3. Monitoring funding to statistics with reduced reporting lag

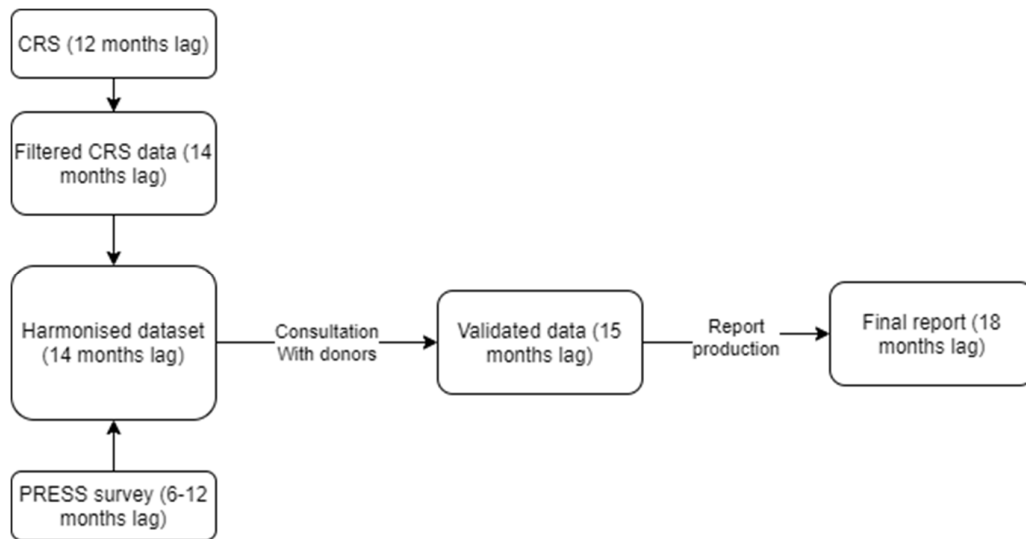
### 3.1. What is the reporting lag?

The workflows for combining the two main data sources of PRESS are described in Figure 2.

One key step when merging the PRESS data, reported by both donors and implementors, with the CRS data is avoiding duplication in a donor-implementor-recipient funding flow. To achieve this, the

projects are examined against their unique identifier in both sources. The projects reported by implementers (mostly from the PRESS survey) are not counted as contribution of the reporting agencies. These projects are counted as projects by the donor agencies, after duplication checks were applied when merging the projects reported by implementers and the projects reported by donors. As the data and final report of PRESS depend in large part on the CRS database, which has a 12-months lag in coverage, the previous editions of PRESS did not capture timely donor financial flows to statistics, leading to a structural lag in reporting.

**Figure 2. How the lag in the CRS data led to a lag in previous rounds of PRESS**



This lag meant that in its previous format, PRESS could not provide timely information for partners in data and statistics, including:

- Nowcasting the funding to statistics
- Forecasting funding to statistics

Hence, despite the many improvements in PRESS over the years, the lack of timely aid reporting is a persistent concern among its primary users, especially development aid providers. With a growing interest in supporting data and statistics, there is an increasing demand for timely data to plan activities and projects and coordinate development co-operation efforts. This issue has become particularly urgent in light of the coordinating efforts to fund the Cape Town Global Action Plan for

Sustainable Development Data (CT-GAP),<sup>4</sup> as well as in the context of a diverse data ecosystem comprising new actors.

PARIS21 addressed this request in its 2019 annual meeting by introducing the concept of a methodology extension. While the PRESS methodology will still be used to report information until 18 months before the publication, the methodology extension will provide stakeholders with PRESS-like information on more recent periods, therefore reducing the reporting lag significantly. This concept became more relevant in 2020, when the development co-operation community had to face the challenges that arose due to the COVID-19 pandemic in national statistical systems (PARIS21, 2020a) and funding to data by domestic and external stakeholders (PARIS21, 2020b).

### 3.2. Estimating up-to-date support to statistics using CRS

While the previous PRESS (from 2008 to 2019) captured the support to data and statistics by looking at global **commitments**<sup>5</sup> to statistics, the annual **disbursements**<sup>6</sup> received by a certain country are also informative for donors and countries when planning their activities, especially those short-term activities financed by a donor's annual or biannual budget. Leveraging this additional variable allows for the estimation of funding to data and statistics received by countries in the current and coming years while still using the same base data, i.e., the CRS and PRESS surveys (and many other data sources on development aid, see section 3), which include both variables for each project.

Looking at disbursements instead of commitments to estimate the support to data and statistics has two distinct advantages:

---

<sup>4</sup> See <https://unstats.un.org/sdgs/hlg/Cape-Town-Global-Action-Plan>.

<sup>5</sup> A firm obligation, expressed in writing and backed by the necessary funds, which is undertaken by an official donor. It provides specified assistance to a recipient country or a multilateral organisation. Bilateral commitments are recorded in the full amount of the expected transfer, irrespective of the time required for the completion of disbursements. Commitments to multilateral organisations are reported as the sum of (i) any disbursements in the year reported on, which have not previously been notified as commitments, and (ii) expected disbursements in the following year.

<sup>6</sup> The release of funds to or the purchase of goods or services for a recipient; by extension, the amount spent. Disbursements record the actual international transfer of financial resources, or of goods or services valued at the cost to the donor. In the case of activities conducted in donor countries, such as training, administration, or public awareness programmes, disbursement is assumed to have occurred when the funds have been transferred to the service provider or recipient. These may be recorded as gross (the total amount disbursed over a given accounting period) or net (the gross amount, less any repayments of loan principal or recoveries on grants received during the same period). It can take several years to disburse a commitment.



- 1) Disbursements capture the actual release of funds, so are more useful for donor planning purposes.
- 2) It can take several years to disburse a commitment and some commitments are never disbursed. Hence, by design, there are more data points available on disbursements than commitments over the same time period. The additional data on disbursements allows for better understanding of financing patterns and donor behaviour, leading to more robust data analysis.

This availability of more data points enables us to estimate support to statistics in the current year (nowcasting) through robust regression analysis. It also provides more substantial evidence of funding trends in the coming years (forecasting). The following sub-sections will focus on how to arrive at these estimates.

### ***3.2.1. Nowcasting: using commitments to predict current disbursements***

Given that CRS has a lag of 12 months for reporting both disbursements and commitments, one way we can estimate support to statistics disbursed in the current year is by looking at the relationship between the two variables. The literature on aid predictability indicates that these two variables may be closely related over time. A 2013 study examining aid predictability based on CRS data also shows that commitments have a significant impact on disbursements five years after they were made (Hudson, 2013).

For most development projects reported to the CRS and the PRESS survey, both commitment and disbursement data are reported. Even when these variables are not directly reported, however, the missing value can usually be imputed.<sup>7</sup>

Using these two variables, PARIS21 has developed a simple linear regression model to estimate the funding from donors based on historical data at activity level. Regression analysis was conducted to predict current disbursements based on reported commitments, captured by *Average\_Annual\_Spending*.

$$\text{Disbursement} = \text{Average\_annual\_spending} * k + d$$

---

<sup>7</sup> For example, the amount for technical support projects that do not have direct monetary transfers can be replaced by a cost estimate by the provider. In cases where the disbursement information is missing, the estimated disbursement amount can be calculated by dividing the unspent commitment amount using the number of years left before the expected end date.

Where **Average\_annual\_spending** = **Total Project Commitments/number\_of\_years**

$k$  is the regression coefficient and  $d$  is the error term. The number of years is the difference between the start date and end date of a reported activity. Reported dates are used for activities with missing value in those two variables. The analysis used the most recently accessible data from the CRS<sup>8</sup>.

This model shows a correlation between disbursements and average spending. Average annual spending is calculated based on the assumption that commitment without a detailed plan for disbursement will be distributed evenly by year, from the expected start year to the end year of the project.

**Table 1. Regression table from the analysis**

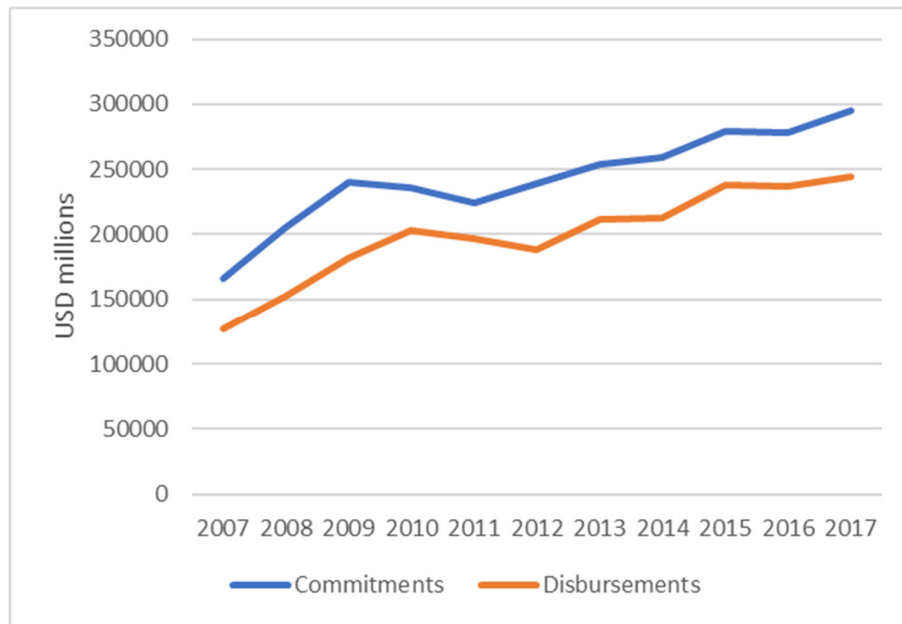
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	38.455	8.04901	-4.778	0.0139 *
<i>Average_annual_spending</i>	0.87461	0.05191	12.997	0.0456 *

The analysis of CRS data shows a significant correlation (90%) between disbursements and commitments each year. The value of  $k$  and the predictability of the model vary depending on the reporting pattern of each donor. For example, while the commitment numbers reported by most donors each year are usually higher than disbursements ([Figure 3](#)Figure 3), this is reversed in the case of a few donors ([Figure 4](#)Figure 4).<sup>9</sup>

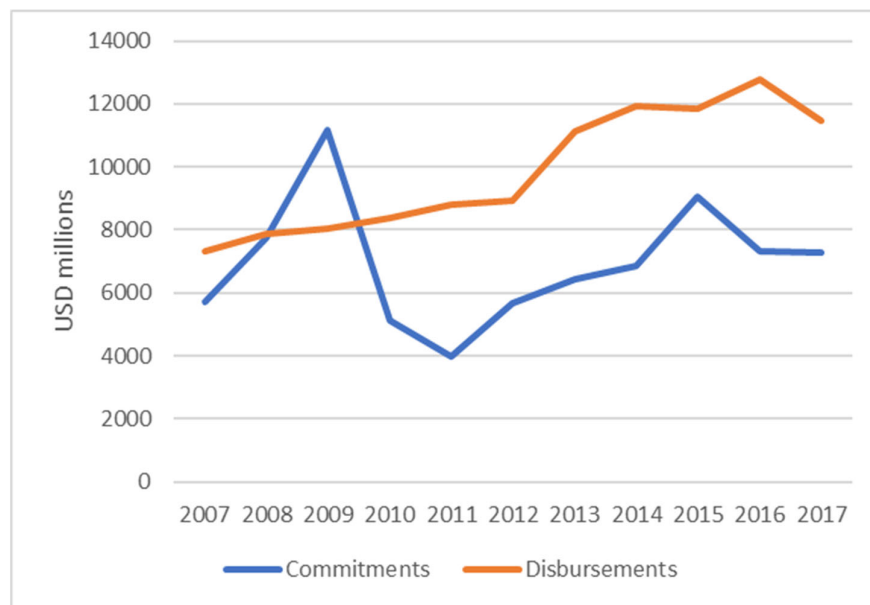
**Figure 3: Disbursements vs commitments reported in CRS (all donors)**

<sup>8</sup> See <https://stats.oecd.org/DownloadFiles.aspx?DatasetCode=CRS1>

<sup>9</sup> This particular reverse correlation can be explained by different factors. Firstly, the financial crisis impacted the continuity of some donors' ODA flow more than others. The significant variation of the exchange rate or inflation rate of a donor could also lead to a sudden change in the converted constant value of aid. In addition, some donors tend to make more long-term commitments than others, resulting in a distribution of disbursements over a long period of time, even after the donors had significantly reduced their overall international aid package.



**Figure 4: Disbursements vs Commitments in CRS reported in CRS by the UK**



The estimated value can deviate from the actual reported number by 10-20%, and the accuracy is higher for more recent years as donors have committed to better transparency and reporting granularity.

Using the above method, PARIS21 was able to nowcast the funding to statistics in years that the most recent CRS data yet to cover. For instance, although the CRS data available in early 2020 only includes full coverage of official aid until 2018, the nowcast is able to provide information on aid to statistics including 2019 and 2020. This is because the stable relationship between average annual spending and commitments is leveraged, which allows us to estimate the 2019 and 2020 disbursement values, from the 2018 reported commitment values.

As a result, for the first time, the 2020 edition of PRESS presented information on funding to data and statistics up to 2019, as opposed to two years prior as in editions including and before PRESS 2019.<sup>10</sup>

### ***3.2.2. Forecasting: anticipating future funding***

The predictability of disbursements and commitments used for nowcasting funding to statistics decreases greatly after the current year, since, for example, many projects which commenced in 2017 will end in 2020. However, this lack of predictability can be partially mitigated by the creation of a forecasting model based on a few well-informed assumptions, leveraging past PRESS data and PARIS21's institutional knowledge on support to statistics for over two decades. These assumptions, which can lead to better forecasting quality, are described below:

- 1) **Continuation of certain long-standing projects:** We can assume that large projects such as the support to the Demographic Health Survey driven by USAID, IMF's national and regional training on economic statistics, and the World Bank's programme on statistical development will remain stable in the near future. Significant changes on these programmes are also easier to target and detectable. These projects are generally stable and attract similar spends each year. Likewise, the upcoming censuses or major surveys in low-income countries are expected to be funded partially by donors.<sup>11</sup> This information accounts for *nearly half of the total amount for data and statistics*. Confirmation from donors of the continuation of these projects can further improve the accuracy of this analysis.
- 2) **Termination or reduction of funding for certain projects:** We can also anticipate the termination or reduction for funding tied to a project based on its specific nature. For instance,

---

<sup>10</sup> Due to the disruptive effect of the COVID-19 pandemic on the predictability of the model, the nowcasting results for 2020, although produced, were not presented in the PRESS 2020 report.

<sup>11</sup> The COVID-19 pandemic has brought great uncertainty to this assumption, especially for censuses planned to take place between 2020 and 2021. Although funding for most censuses has been secured, many countries have diverted their national budget to other, more urgent matters (CCSA, 2020) (UNFPA, 2020), resulting in the postponement of external funding.

the support for censuses is a one-off disbursement and will not reoccur until the next census round. Similarly, if a country become no longer eligible for ODA, graduate from IDA's borrower list or becomes an upper-middle-income country, it is then expected to receive a lower ODA grant and become ineligible for some loans. In those cases, support for statistics might be affected disproportionately, given its low priority.

It is also crucial to state that these predictions can only be accurate if the following additional assumptions are met:

- Development aid providers maintain their current levels of effort
- Existing programmes continue to run
- Commitments are fully disbursed
- There is a response to prioritised needs such as censuses

The estimation is also limited if donors agencies publishing the new initiatives with a lag. The predictability of both nowcast and forecast on funding to data and statistics also relies on aid providers committing to maintaining the transparency and timeliness of their aid data. The forecasting looks at contribution until  $n+2$ , given most international organisations' work plan don't go beyond that horizon.

Due to the uncertainty caused by the COVID-19 pandemic, the forecasting results from this methodology were not presented in PRESS 2020. These will appear in the future PARIS21 publications once more evidence becomes available.

In sum, the forecasting estimates should be interpreted with significant caution even if the above model indicates a relative increase in the coming years. According to many historical estimates, the funding gap for data and statistics (i.e. to find the entire CT-GAP) is far from being closed. This gap is likely to be exacerbated by the effects of the ongoing COVID-19 crisis in large parts of the globe.

### **3.3. Expanding the PRESS database**

#### ***3.3.1. Exploring alternative data sources for aid flows on statistics***

Apart from nowcasting and forecasting disbursements to statistics from PRESS data, another way to address the structural lag in aid-flow reporting can be by attempting to remedy the root cause of the problem – the dependency on the CRS database – and searching for more timely information in alternative data sources. PARIS21 has identified three main (types) of alternative data sources, outlined below:

### ***The International Aid Transparency Initiative (IATI)***

The IATI datastore is the largest alternative database outside of OECD-DAC data for official development assistance. With more than 100 donors reporting to this database, IATI has a much shorter lag than CRS. It also covers more projects by philanthropic foundations. The COVID-19 pandemic and the rising need for coordination has also incentivised aid providers to report to IATI with less delay. However, IATI data suffer from a lack of quality assurance and inconsistency within the dataset. Although it uses a similar data structure as the CRS, the reported projects in IATI may not include important granular information, such as the project description. Furthermore, as many donors only committed to reporting to IATI after 2014, the lack of historical data for drawing time series also affects its ability to forecast.

### ***Donors' transparency portals***

In recent years, global donors have strengthened their efforts in aid transparency. Many donors have developed online data portals or uploaded online datasets to share information on their aid projects, especially the major donors in statistics such as the World Bank, UNDP, USAID, FCDO, IDRC, etc. These datasets usually have a similar density of information as the CRS data and are usually updated more frequently than CRS. However, the majority of donors still lack appropriate portals and public datasets. Furthermore, merging these different datasets is possible, but time intensive.

PARIS21 has been exploring these data sources since 2019 and has accumulated knowledge over this period. For example, the USAID dataset helped PRESS 2019 to identify the USA's support to statistics for the first time; in particular, its effort with the Demographic Health Survey (DHS). PARIS21 has also established a methodology for merging and harmonising the aforementioned datasets. The methodology maps variables in different datasets against each other and uses internal project identifiers to avoid duplications.

### ***Multilateral donors' prospective reporting to the PRESS survey***

The online PRESS survey (introduced in section 2<sup>4</sup>) includes a feature that allows donors to report on future projects. Since many donors have a biannual programme of work, in each year's survey, they are encouraged to provide information on the project they have planned or committed to in the near future. In the survey, donors can verify, edit, and cancel future projects in the next round of reporting. However, previous editions of PRESS did not fully reflect future projects due to the report's focus on accuracy. However, these future projects could still contain valuable information to assist in the projection of aid flows. The methodology developed by PARIS21 leverages these projects as an

underutilised existing data source that may not reflect completely on the activities from donors. Nevertheless, it is useful for nowcasting funding.

The comparison of the above data sources can be found in [Table 2](#)~~Table 1~~. As another important data source in the area of data and statistics, the Eurostat's donor survey is analysed in [Box 2](#)~~Box 2~~.

**Table 2: Comparison of data sources**

Data source	No. of observations	Advantages	Disadvantages
<b>Conventional data source</b>			
<b>CRS</b>	More than 2.5 million	<ul style="list-style-type: none"> <li>• Full coverage of DAC donors' portfolios</li> <li>• Data quality assured by the WP-STAT<sup>12</sup> standards</li> </ul>	<ul style="list-style-type: none"> <li>• 12-month lag</li> <li>• Relatively moderate coverage of multilateral organisations</li> </ul>
<b>Alternative data sources</b>			
<b>PRESS survey</b>	4000 to 5000	<ul style="list-style-type: none"> <li>• Only statistics- or data-related projects are included</li> <li>• Allows reporters to enter future planned projects</li> </ul>	<ul style="list-style-type: none"> <li>• Only covers key multilateral donors in statistics</li> <li>• Detailed information may be omitted by reporters during the reporting process</li> </ul>
<b>IATI<sup>13</sup></b>	More than 500,000	<ul style="list-style-type: none"> <li>• A shorter lag in reporting than CRS</li> <li>• Reporting from NGOs and philanthropic foundations</li> <li>• Relatively wider coverage on multilateral organisations</li> <li>• More activities reported over the years</li> </ul>	<ul style="list-style-type: none"> <li>• Lack of quality assurance</li> <li>• Incomplete portfolio</li> </ul>
<b>Donor transparency portals</b>	The full portfolio for each donor	<ul style="list-style-type: none"> <li>• Full coverage of donors' portfolios</li> <li>• Well-maintained and updated</li> </ul>	<ul style="list-style-type: none"> <li>• Fewer than ten major donors provide accessible databases</li> </ul>

<sup>12</sup> DAC Working Party on Development Finance Statistics, see <http://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/wp-stat.htm>

<sup>13</sup> See <https://iatistandard.org/>

### **Box 2: The Eurostat's Donor Survey – and why it is not included in the PRESS database**

The annual Donor Survey by Eurostat aims to provide an overview of statistical projects in the respective countries to allow for better planning of assistance in the field of statistics, to benefit from acquired experience, and to avoid overlap. The survey is distributed to donor countries, international organisations, and the recipients (mostly Eurostat member states or partner states) of support to data and statistics.

According to the report of the 2020 round, the Donor Survey's main objectives included:

- facilitating better planning and prioritisation of assistance, especially amongst the donors;
- increasing transparency/visibility for donors and implementing agencies and potentially serving as an “audit” document for beneficiaries;
- providing the beneficiaries with a broad overview of the areas in which they are receiving support, also at regional level, and prioritising their future needs;
- benefitting and learning from the experiences, good practices, and shortcomings of other projects;
- facilitating a dialogue between beneficiaries, donors, and implementing agencies.

At donor country level and on the donors' side, respondents to the Donor Survey overlap with CRS. However, the CRS reporting is coordinated by one national focal point, while the Donor Survey questionnaire is distributed to both the aid agencies and national statistics offices who provide technical support. However, the Donor Survey dataset does not have a good system to harmonise the data from the two sides, nor does it contain the programme identifiers, which are essential to avoid duplication when combining information with other sources. Unlike the development of financing databases, the Donor Survey also lacks a mechanism to update previous years' reported projects. Coverage of the survey is also not comprehensive enough to provide additional values to the alternative data sources identified in this document.

### ***3.3.2. Addressing gaps in the alternative data sources***

To take advantage of the alternative data sources, PARIS21 combined the information from CRS and online surveys with these alternative data sources to create a more up-to-date, harmonised database on funding for data and statistics. However, this means that two important problems must be addressed, described below:

**1) The completeness problem:** A common weakness of the alternative data sources, compared with CRS, is their incomplete coverage. This is especially apparent in IATI where, unlike CRS, donors may not report their full portfolio, leading to a lack of comprehensive information (a “vertical” information gap i.e. multiple donors but incomplete project information for individual donor). Furthermore, the CRS uses a more “centralized” reporting



system for each donor country, whereby information from different agencies providing ODA is gathered under one entity before reporting to CRS as a whole. The IATI, on the other hand, allows different agencies in a single donor country to report their data separately. This implies different reporting patterns for different agencies based on their capacity to do so regularly, and a lack of overall coordination.

**2) The coverage problem:** CRS contains information from over 100 donors. On the other hand, donor transparency portals suffer from a “horizontal” problem, i.e., they usually have complete coverage and contain the full portfolio, but fewer than ten major donors provide access to such open and easy-to-use databases.

PARIS21 has mitigated the horizontal and vertical problems in alternative data sources by harmonising and linking these databases. The alternative sources provide a wider coverage, while the combination of CRS, IATI and donors’ databases enhance the completeness of the data. The final dataset used in the analyses combines the alternative data sources and PRESS data using project IDs and other identifiers. As a caveat, however, the total number of projects included in the alternative sources still only represents 40% of all number of projects reported in the CRS.

The advantage of using such data sources, such as the timeliness and inclusion of philanthropic foundations, makes them a useful extension of PRESS data, especially when trying to solve the lag issue. However, their weaknesses imply that they are not a substitute for CRS or, by extension, conventional PRESS analysis.

### ***3.3.3. Linking the alternative sources: the new harmonised database***

The next step in leveraging the independent alternative data sources described above is to link them with PRESS data and create a new harmonised database of disbursements for data and statistics **at project level**.

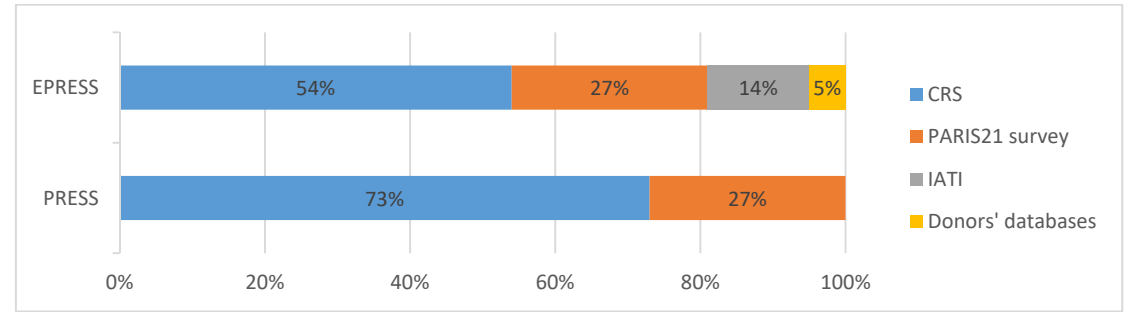
Using disbursements as the primary variable to determine the support to statistics is even more beneficial at this stage, and data on annual disbursements is adequately available in most of the sources considered above. The activity-based CRS data, for example, contain nowcasting/forecasting regressions on this new data set by Simon etc. to see value add and/or consistency of our findings! Should we present this? I think you did it and found little to no change right? disbursement information for more than 98% of the projects. Similarly, the PRESS survey for multilateral donors provides specific information on disbursement plans (though it is project-based). The donors’ transparency portals are also expenditure/disbursement based.

However, the IATI database is weaker in this regard: it contains an unusually high percentage of negative commitments or disbursements. For example, data downloaded from the IATI database in 2018 contained negative value commitments or disbursements in 18% of projects reported<sup>14</sup>. In comparison, less than 1% of activities reported in CRS has negative value, mostly associated with loan repayments and correction of previous entries, i.e. not a database error.

For the harmonised database, the missing disbursement values were imputed based on the assumption that the commitments were distributed evenly from the start date to the end date of the project. The negative projects in IATI were corrected or removed by cross-validating against CRS. The duplicated projects were removed based on project identifiers. Consultations with several multilateral donors were conducted to ensure the validity of the final data.

Figure 5 presents an example of how the new harmonised dataset looks after linking the different sources. For the 15,312 projects in the new datasets, 54% of disbursement activities come from the CRS data, compared with over 73% in the earlier dataset used for PRESS. In the new dataset, the PRESS survey accounts for 27% of projects, while IATI data and donor databases account for 19% of projects. By filtering through the data using recipient country and year, donors can already observe the upcoming funding received by a country for statistical development. It is then easier for them to identify funding gaps in prioritised areas. Hence, the new harmonised database can achieve better diversification of data sources and reduce the dependence on CRS.

**Figure 5: Comparison of PRESS and the new harmonised database for the share of projects in the final datasets, by sources of data, 2016-2018**



<sup>14</sup> Based on data collected in 2019, the future improvements of the IATI database have resolved this issue.

### **3.4. Bringing them together – nowcasting and forecasting with the new harmonised database**

PARIS21 applied the same model on nowcasting and forecasting support to statistics (disbursements) but based on the new harmonised database, linking PRESS data with the aforementioned alternative sources. The main findings of this approach were similar to those based on CRS estimations: there is no indication of a systematic increase in funding to statistics in the current or coming years.

## **4. Conclusion**

The nowcast and forecast analysis aims to address the specific reporting lag issue faced by the PRESS, rather than to substitute it. Even with longer lag, PRESS still reports the most reliable and comprehensive data on funding to statistics. The PRESS database continues to serve as the source data for SDG Indicator 17.19.1.

The outputs from nowcast and forecast also vary in their accuracy. While the merged database is as robust as PRESS, and the nowcast results are relatively reliable, the forecast analysis is based on several assumptions. Results produced through the nowcast and forecast analyses can therefore be used for different products to serve different purposes:

- The nowcast on disbursements and some information provided by the harmonised dataset can be directly presented in PRESS going forward, as a natural extension of its current content, based on existing data sources such as CRS.
- The forecast for funding gaps could be presented in separate policy briefs due to its speculative nature.
- The complete, harmonised database can also be deployed on a platform such as Bern Network's 'Clearinghouse for Financing Development Data'<sup>15</sup> to be used as a dynamic planning tool by development partners.

The immediate next step in improving the new methodology is to strengthen the communication and consultation between PARIS21, donors, and recipients. Benefits from such consultations would lead to further robust results by correcting erroneous information contained in IATI data, validating assumptions on the termination or continuation of projects in assumptions for forecasting, and

---

<sup>15</sup> <https://bernnetwork.org/>

enhancing data sharing in general. Consultations for this product will also help PARIS21 to shape its work in order to better meet the demand of its stakeholders.

As the COVID-19 pandemic will affect donors' GNI and consequently their ODA, fluctuations in development financing for data and statistics can be expected. Despite the difficulty these anomalies will bring to the analyses, the transparency and timeliness of the information on funding to data and statistics have also become more relevant, as identified in Part I of PRESS 2020. The results from these analyses can play a crucial role in informing and supporting the statistics and development co-operation communities in their response to these challenges.

## References

- CCSA. (2020). *How COVID-19 is changing the world: a statistical perspective. Volume II*. CCSA.  
Récupéré sur [https://unstats.un.org/unsd/ccsa/documents/covid19-report-ccsa\\_vol2.pdf](https://unstats.un.org/unsd/ccsa/documents/covid19-report-ccsa_vol2.pdf)
- Hudson, J. (2013). Promises kept, promises broken? The relationship between aid commitments and disbursements. *Review of Development Finance*, 109-120.  
doi:<https://doi.org/10.1016/j.rdf.2013.08.001>
- PARIS21. (2020a). *Combating COVID-19 with data: what role for National statistical systems?* Paris: PARIS21. Récupéré sur [https://paris21.org/sites/default/files/inline-files/COVID\\_Policybrief\\_Full.pdf?v=2.0](https://paris21.org/sites/default/files/inline-files/COVID_Policybrief_Full.pdf?v=2.0)
- PARIS21. (2020b). *Partner Report on Support to Statistics 2020*. Paris: PARIS21.
- UNFPA. (2020). *Technical Brief on the Implications of COVID-19 on Census*. New York: UNFPA.  
Récupéré sur [https://www.unfpa.org/sites/default/files/resource-pdf/Census\\_COVID19\\_digital.pdf](https://www.unfpa.org/sites/default/files/resource-pdf/Census_COVID19_digital.pdf)