Using Data Analytics in Public Procurement: Operational Options and a Guiding Framework
## Contents

<table>
<thead>
<tr>
<th>Acknowledgements</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acronyms</td>
<td>v</td>
</tr>
<tr>
<td>Executive Summary</td>
<td>1</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2. Analytical Framework</td>
<td>6</td>
</tr>
<tr>
<td>2.1. A Performance Measurement Framework for Public Procurement</td>
<td>7</td>
</tr>
<tr>
<td>2.2. Exploring the Links between Drivers and Outcomes</td>
<td>9</td>
</tr>
<tr>
<td>2.3. Using Data Analytical Tools to Analyze Trade-Offs and Incentives in Public Procurement</td>
<td>11</td>
</tr>
<tr>
<td>2.4. Linking Public Procurement Data with Other Dimensions of the Public Sector and Public Administration</td>
<td>13</td>
</tr>
<tr>
<td>3. Data Sources</td>
<td>15</td>
</tr>
<tr>
<td>3.1. Public Procurement Transactions</td>
<td>16</td>
</tr>
<tr>
<td>3.2. Open Data and Data Standards</td>
<td>17</td>
</tr>
<tr>
<td>3.3. Other Relevant Data Sources</td>
<td>17</td>
</tr>
<tr>
<td>4. Operational Options Related to Procurement Data Analytics</td>
<td>21</td>
</tr>
<tr>
<td>4.1. Data-Driven Public Procurement Diagnostics</td>
<td>23</td>
</tr>
<tr>
<td>4.2. Evaluating the Impact of a Procurement Activity or Strategy</td>
<td>25</td>
</tr>
<tr>
<td>4.3. Development of Sustainable Data Monitoring Tools</td>
<td>28</td>
</tr>
<tr>
<td>4.4. Assessment and Improvement of Quality and Openness of Procurement Data</td>
<td>29</td>
</tr>
<tr>
<td>5. Capacity Building</td>
<td>31</td>
</tr>
<tr>
<td>6. Entry Points and Synergies</td>
<td>33</td>
</tr>
<tr>
<td>6.1. Assessments and Diagnostics</td>
<td>34</td>
</tr>
<tr>
<td>6.2. Lending and Technical Assistance Operations</td>
<td>34</td>
</tr>
<tr>
<td>6.3. Building a Continued Country Engagement and Business Line</td>
<td>35</td>
</tr>
<tr>
<td>7. Conclusion</td>
<td>36</td>
</tr>
<tr>
<td>References</td>
<td>38</td>
</tr>
</tbody>
</table>
Acknowledgements

This paper is a collaborative effort across the teams from the Equitable Growth, Finance and Institutions (EFI) Governance Global Practice and the Development Impact Evaluation (DIME) Governance. The authors are Serena Cocciolo (lead author), Sushmita Samaddar, Kristina Aquino, Maria Arnal Canudo, and Stephen Okiya. The authors worked under the guidance of Tracey Lane (Practice Manager, EPSPA), Ed Olowo-Okere (Global Director, EPSDR), Vinay Sharma (Senior Adviser, PA9SS), and Arianna Legovini (Director, DIME0).

This paper is a compilation of experiences and lessons from applying data analytics and research tools in public procurement, and it is drawn from academic papers and World Bank projects. Inputs were extracted from collaborations and discussions with Khalid Bin Anjum (Senior Procurement Specialist, EAERU), Alexandre Borges de Oliveira (Lead Procurement Specialist, OPSPR), Elena Carmen Calín (Procurement Specialist, EECRU), Luiza Cardoso De Andrade (Junior Data Scientist, DIME3), Thiago De Gouvea Scot de Arruda (Economist, DIME3), Vincenzo Di Maro (Senior Economist, DIME0), Manjola Malo (Procurement Specialist, ELCRU), Desislava Enikova Nikolova (Senior Economist, EEEM2), Mihály Fazekas (Assistant Professor, Central European University), Shanker Lal (Lead Procurement Specialist, ESARU), Daniel Ortega Nieto (Senior Governance Specialist, ELCG2), Marcela Rozo Rincon (Senior Public Sector Specialist, ELCG2), Ishtiak Siddique (Senior Procurement Specialist, EAERU), and Antonia G. Viyachka (Procurement Specialist, EECRU). Leandro Justino Pereira Veloso, Ruggero Doino, and Samuel Garoni provided excellent research assistance.

The team is grateful for the comments and guidance provided by the peer reviewers: Ahsan Ali (Lead Procurement Specialist, ESARU), David S. Bernstein (Lead Public Sector Specialist, EMNGU), Shawkat Muhammad Quamrul Hasan (Lead Procurement Specialist, EECRU), Santiago Herrera (Senior Economist, ELCMU), Knut J. Leipold (Lead Procurement Specialist, EAERU), Francesca Recanatini (Lead Economist, ELCG2), Daniel Oliver Rogger (Senior Economist, DIME3), Joanna Alexandra Watkins (Senior Public Sector Specialist, EAEG2), and Luciano Wuerzius (Senior Procurement Specialist, ELCRU).
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANAP</td>
<td>Romanian National Agency for Public Procurement</td>
</tr>
<tr>
<td>CEM</td>
<td>Country Economic Memorandum</td>
</tr>
<tr>
<td>CPA</td>
<td>Central Procurement Authority</td>
</tr>
<tr>
<td>CPF</td>
<td>Country Partnership Framework</td>
</tr>
<tr>
<td>eGP</td>
<td>e-Government Procurement</td>
</tr>
<tr>
<td>GPP</td>
<td>Green Public Procurement</td>
</tr>
<tr>
<td>MAPS</td>
<td>Methodology for Assessing Procurement Systems</td>
</tr>
<tr>
<td>PDA</td>
<td>Procurement Data Analytics</td>
</tr>
<tr>
<td>PER</td>
<td>Public Expenditure Review</td>
</tr>
<tr>
<td>PFM</td>
<td>Public Financial Management</td>
</tr>
<tr>
<td>PPA</td>
<td>Public Procurement Authority</td>
</tr>
<tr>
<td>OCDS</td>
<td>Open Contracting Data Standards</td>
</tr>
<tr>
<td>OCP</td>
<td>Open Contracting Partnership</td>
</tr>
<tr>
<td>SCD</td>
<td>Systematic Country Diagnostic</td>
</tr>
<tr>
<td>SMEs</td>
<td>Small and Medium Enterprises</td>
</tr>
<tr>
<td>WOE</td>
<td>Woman-Owned Enterprise</td>
</tr>
</tbody>
</table>
Executive Summary

The world spent $11 trillion on public procurement in 2018, amounting to 12 percent of global GDP (Bosio et al. 2022). Given these substantial volumes, public procurement can contribute to several objectives: savings, integrity, economic growth, inclusiveness, and sustainability.

Procurement Data Analytics (PDA) can contribute to the achievement of these objectives. It refers to the use of data to generate actionable insights and evidence to monitor outcomes, inform the policy dialogue, guide reform efforts, and assess the impact of reforms and strategies in public procurement.

Despite a growing academic literature and impact evaluations on public procurement, the existing body of evidence is still scarce and limited to a few countries. This impedes drawing generalizable lessons on optimal policies and strategies to achieve the multi-layered objectives of the public procurement function, therefore highlighting the need for a larger adoption of data analytics tools in this area.

With the increasing adoption of electronic government procurement (eGP) systems and the corresponding digitization of transaction records, public procurement has enormous untapped potential for the application of data analytics tools.

This paper highlights the successful approaches and good practices of previous PDA work and provide useful resources to World Bank teams with country engagements relating to public procurement. Possibly interesting to a broader audience, an analytical framework is also discussed to guide the application of data analytics tools in public procurement, data sources, the open government agenda, and data standards.
The PDA products that can be offered to World Bank’s client countries include data-driven public procurement diagnostics, impact evaluations, data monitoring tools and public procurement dashboards, and projects to assess and improve the quality and openness of procurement data. The paper discusses the benefits and expected impact of each of these products, how they can be integrated into other diagnostics or lending operations, and the possible applications for countries without an eGP system.

The long-term goal in the PDA agenda is to empower World Bank’s client countries with the necessary knowledge and tools to effectively conduct these activities independently, so that the use of PDA tools is not just a one-time exercise but rather it becomes sustainably embedded in monitoring, evaluation, and decision-making processes in public procurement.

The Governance Global Units and Development Impact Evaluation (DIME) have been collaborating on a research agenda on various public procurement topics and for the delivery of PDA products in collaboration with task teams in the regions. Support is available for task teams considering the application of PDA tools.
Introduction

Public procurement is a sizeable economic activity, and it can have multi-dimensional objectives. The government’s role in the economy is determined by its overall spending or revenue collection in proportion to GDP. A significant share is government spending on goods and services, which is typically referred to as “public procurement.” The world spent $11 trillion on public procurement in 2018, amounting to 12 percent of global GDP (Bosio et al. 2022). Given these substantial volumes, public procurement can contribute to several objectives.

First, efficient public procurement can generate sizeable fiscal savings. Government spending is typically constrained by existing obligations and expectations from various stakeholders, such as, in the case of the public sector, wage bills, pensions, and social transfers. While public procurement is a relatively maneuverable category of government spending, governments can free up resources for other expenses or to reduce fiscal imbalances by improving value for money in public procurement, rather than curbing investments. For example, Best et al. (2017) show that, in Russia, moving the worst-performing quartile of procurers to 75th percentile-effectiveness would reduce procurement expenditures by around 11 percent, equivalent to $13 billion each year. Globally, it is estimated that halting the waste in public procurement could free up at least $1 trillion a year globally (World Bank 2022b) to put economies on a path toward green, resilient, and inclusive development. Significant savings can be achieved from changes in procurement practices and behaviors, even without structural changes in policies, laws, and regulations, such as through the choice of the auction design (Athey et al. 2011), the use of incentive contracts (Girth and Lopez 2018), or through the centralization of purchases, which

---

1. The procurement and contracting cycle refers to the sequence of related activities, from needs assessment, through competition and award, to payment and contract management, as well as any subsequent monitoring or auditing (OECD 2021). While definitions may vary across countries, in this note, public procurement refers to procurement by public entities as well as by state-owned enterprises.
2. Centralized procurement is a system in which a central procurement body manages the purchase of goods and services for other contracting entities. Framework agreements are an example, and guidance on this instrument is provided by a recent report by the World Bank (2021).
have generated approximately 2 to 8 percent reductions in regional per capita health expenditures in Italy (Ferraresi et al. 2021).

Second, one of the pillars of the anti-corruption agenda is the integrity of public procurement spending. Corruption, collusion, and other uncompetitive practices cause misallocation of public funds, higher expenses, and substandard and insufficient services (OECD 2015). Although it is difficult to measure the exact cost of corruption due to its hidden nature, it has been estimated that 10 to 30 percent of the value of public-funded construction projects may be lost through corruption and mismanagement (OECD 2016). For example, Baltrunaite (2020) shows that corporate donations to political parties and campaigns buy preferential treatment in procurement auctions, leading to approximately 5 percent of misallocated contracts because of corporate donations and estimated costs increases in procurement caused by corporate donations of almost 1 percent of GDP.

Third, public procurement can contribute to economic growth. Public procurement can boost firm growth both in the short and medium-long term, as demonstrated by Ferraz et al. (2015) in the context of Brazil, where firms winning government contracts experience higher employee growth, are more likely to participate and win future auctions and expand their business in new markets. Public procurement is a large sector, and, across 143 countries covered by the World Bank Enterprise Surveys, on average, 13 percent of firms secured or attempted to secure a government contract (World Bank, forthcoming - b). Contracting with the government can represent an economic opportunity for firms to grow and expand in new markets, especially when it is competitive and accessible for all firms, and unnecessary frictions are reduced. For example, accelerated payments have a positive impact on employment growth of suppliers, especially for financially constrained firms (Barrot and Nanda 2020). Stricter regulations can also address payment backlogs and reduce firms’ exit rates, especially in sectors with a large share of small firms (Conti et al. 2021).

Fourth, the inclusiveness and sustainability of public procurement can contribute to the achievement of socioeconomic and environmental aspirations (World Bank 2022c; OECD 2021). When used strategically, public procurement can play a significant role in supporting national policy priorities and addressing local and global challenges, such as in developing local industries, reducing income inequalities, creating local jobs, addressing emergencies, and fighting climate change. Increasingly, countries have been adopting policies and strategies to enhance the inclusiveness and sustainability of public procurement, such as preferential policies for SMEs or women-owned enterprises and green public procurement approaches. For example, 34 percent of World Bank’s client countries globally have a green public procurement strategy or action plan (World Bank 2021h). On the one hand, these policies can align private and public sector incentives and reduce market failures. On the other hand, they can be distortionary, prone to capture, and difficult to monitor (World Bank 2021i). The existing evidence on the impact of these policies is limited and mixed, making it difficult to draw generalizable lessons on how to best use procurement policies to achieve these horizontal objectives.

Procurement Data Analytics (PDA) can contribute to the achievement of these objectives, and, with the increasing adoption of electronic government procurement (eGP) systems and the corresponding digitization of transaction records, public procurement has enormous untapped potential for the application of data analytics tools. The growing potential for data and analytics to drive development is demonstrated in the 2021 World Bank Development Report, which focuses on how to turn data into information and information into insights to support the achievement of development outcomes. PDA refers to the use of data and analytics to create actionable insights and evidence to monitor outcomes, inform the policy dialogue, guide reform efforts, and assess the impact of reforms and strategies in public procurement. PDA can contribute to a problem-driven iterative approach to strengthening and modernizing national public procurement systems through identifying efficiency and

---

3. The World Bank Enterprise Surveys covers 143 countries. The participation rate in public procurement is calculated as weighted average of answers to question j6a “Did this establishment secure or attempted to secure a government contract in the last 12 months?”, which is available for 142 countries. The World Bank Enterprise Survey micro-data and questionnaires are available at https://www.enterprisesurveys.org/en/data.

4. Successful public procurement has been instrumental for the rapid response to the COVID-19 pandemic. See Open Contracting Partnership 2020a; Open Contracting Partnership 2020b; Lal 2021; and World Bank 2021g.

5. There are various tools that can be used to implement preferential treatment programs for SMEs or other groups. One is to employ contract threshold preferences for SMEs, with contracts below a certain threshold value designated specifically for SMEs, as in the case of Brazil, Colombia, Indonesia, Malta, Russia, and Trinidad and Tobago. A second approach is the use of set-asides or quotas, where a reserved portion of the public procurement budget is set aside for certain types of firms, such as SMEs. Quotas may range from as high as 25% in Angola, Kenya, Brazil, and Liberia to as low as 10% in Argentina, Australia, Chad, Egypt, Philippines, and Uruguay (World Bank 2021i). A third way is based on monetary incentives for SMEs or other targeted groups, such as through waivers of participation fees (e.g., India and Honduras) or through preference margins that—for evaluation purposes only—apply a discount to the offered price (e.g., Bolivia).

6. The World Bank (2021b) report provides an overview of international experience in implementing Green Public Procurement.

7. For example, even two studies conducted on the same program for small firms in California draw opposite conclusions. While Marion (2007) found that procurement costs are 3.8 percent higher on auctions using preferential treatment, Krasnokutskaya and Seim (2011) found those distortionary effects are not huge in comparison to benefits to firm growth.

8. An eGP system allows management and implementation of procurement through an electronic portal, such as publication of tender notices, submission of bids, and notification of awards. Specific guidance for successful eGP adoption can be found in other reports, such as by the World Bank 2022c.
integrity gaps, analyzing trade-offs associated to procurement alternative strategies, developing data tools for monitoring the public procurement function, and generating knowledge and evidence on the impact of certain policies.

However, there are challenges due to the lack of common standards and approaches for PDA. This paper aims to fill this gap by reviewing the existing uses of PDA tools and providing a framework for World Bank staff and other potential users on how and when to use them. At this stage, the goal is not to provide a detailed toolkit or off-the-shelf tools, nor to provide generalizable lessons on optimal procurement policies or strategies. Rather, the paper highlights the successful approaches and good practices of previous PDA work and provides useful resources to World Bank teams with country engagements relating to public procurement. It also provides a novel lens to analyze public procurement, such as explicitly assessing the trade-offs between competing policy objectives and measuring the socio-economic and environmental dimensions of public procurement. Given the existing capacity of Public Procurement Authorities (PPAs) and contracting entities in most of World Bank’s client countries, in the near future the World Bank plays a critical role in demonstrating the benefits and opportunities that PDA offers, and in supporting client countries in the delivery of PDA solutions. Indeed, one of the IDA19 commitments included supporting IDA countries in conducting detailed Procurement Data Analytics to increase efficiency of public spending and mitigate corruption risks (World Bank 2020c). However, the goal is to empower client countries in independently and sustainably using data analytical tools on their own administrative data as part of their own standard monitoring, reporting, and evaluation function; thus, this paper also shares the experiences and lessons learned from previous PDA capacity building activities.

As data availability is a critical condition for implementing PDA tools, this paper provides options that can be meaningfully applied in the absence of an eGP system, or in preparation of it, and discusses their implications in terms of resources and sustainability. The widespread adoption of eGP is one of the priorities for improving public procurement and, by design, it generates digitized and structured micro-data and ideal conditions for PDA. However, countries that do not have an eGP system could also benefit from data-based work. For example, surveys of procurement officers or firms could represent alternative data sources to generate empirical insights on the perceived inefficiencies and challenges of public procurement. In countries without an eGP system, a data-oriented perspective could support the preparation of eGP adoption, for example, by mapping the existing data management system and institutional arrangements, evaluating the quality of existing procurement data, and assessing the existing statistical capacity in PPAs and contracting entities.

The paper is organized as follows: The next section provides some advice on standards, principles, and good practices for delivering PDA products. Section 3 describes the data sources that can be leveraged for PDA and the potential significant benefits of data integration. Section 4 describes the options for supporting clients using PDA. Section 5 elaborates on the importance of building statistical and data analytics capacity in national PPAs. Section 6 discusses the entry points and strategies for deploying PDA with other World Bank’s analytical work and operations, and section 7 concludes with some examples of work at the frontier of PDA, and reflections on how this can stimulate public procurement modernization and reforms. While these sections are interrelated, the World Bank’s task teams working on country level diagnostics such as Systematic Country Diagnostics (SCDs) or Public Expenditure Reviews (PERs) may be particularly interested in section 2, section 3, section 4.1, and section 6.1. The task teams working on technical assistance and operations may be particularly interested in section 4.2–4.4, section 5, and section 6.2–6.3. Section 2 and section 3 provide a general discussion on the analytical framework for PDA and data sources, and therefore can be interesting for a broader audience.

---

9. This paper builds on the experience of the GOV-GP and DIME-governance in various projects for which Procurement Data Analytics was a significant component.
10. A Public Procurement Authority (PPA) is an independent body with the responsibility of regulating, monitoring, and evaluating public procurement. A contracting entity is the unit responsible for management and implementation of procurement for its organization.
Analytical Framework

This section elaborates on a common analytical framework and guiding principles for Procurement Data Analytics. Data analytics tools should be designed and used according to the specific country context and objectives of the project, for example, in terms of identifying the necessary data sources or the need for complementary qualitative approaches, such as focus group discussions. However, a standard approach and analytical framework can be adopted to guide the design and execution of PDA and ensure the quality of deliverables. Section 2.1 presents a performance measurement framework for monitoring and evaluating public procurement, following the four key objectives of public procurement conceptualized in section 1: savings, integrity, economic growth, and inclusiveness and sustainability. Section 2.2 discusses an analytical framework for analyzing the drivers of performance. Section 2.3 highlights the potential of PDA in exploring the trade-offs of procurement policies beyond a purely monetary and transaction perspective. Section 2.4 presents the frontier of data integration, and section 1.1 lays out the workflow and principles for Procurement Data Analytics.
2.1 A Performance Measurement Framework for Public Procurement

Traditional procurement indicators identified by the academic literature and practitioners measure the economy, efficiency, and integrity of the public procurement function, and they largely relate to the objectives of achieving value for money and minimizing corruption and collusion risks. Table 1 provides an exemplificative list of indicators (not necessarily exhaustive) measuring these dimensions, as well as indicators related to transparency and competition which are typically considered as critical mediating factors to achieving economy, efficiency, and integrity in public procurement.11 The indicators reported in table 1 cover dimensions of public procurement related to methods and procedures (e.g., use of competitive methods), transparency and integrity (e.g., time for bid submission), competition (e.g., number of bidders), processing times (e.g., time for bid evaluation), prices (e.g., unit prices), and contract implementation (e.g., time overrun). With respect to public procurement integrity, recent academic studies have been proposing and testing innovative ways of using machine learning techniques on public procurement micro-data for the construction of red flags, corruption risk indicators (see Fazekas and Kocsis 2017; Lisciandra et al. 2021), and cartel screens (see Conley and Decarolis 2016; Huber and Imhof 2019), which can be used for monitoring and evaluation, or as a starting point for further investigation, or for courts to authorize inspections (OECD 2013).12

> > >

**TABLE 1 - Public Procurement Indicators Related to Economy, Efficiency, and Integrity**

<table>
<thead>
<tr>
<th>Economy and Efficiency</th>
<th>Transparency and Integrity</th>
<th>Competition and Collusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tender processing</strong></td>
<td>Process</td>
<td>Process</td>
</tr>
<tr>
<td>Total processing time</td>
<td>Time for bid preparation</td>
<td>Open procedure</td>
</tr>
<tr>
<td>Evaluation time</td>
<td>Length of eligibility criteria description</td>
<td>Bid security requirement</td>
</tr>
<tr>
<td>Contracting time</td>
<td>Length of product description</td>
<td>Advertisement of tenders</td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td><strong>Assessment</strong></td>
<td><strong>Participation and competition</strong></td>
</tr>
<tr>
<td>Awarded unit price</td>
<td>Share of excluded bids</td>
<td>Number of bidders</td>
</tr>
<tr>
<td>Final unit price after renegotiation</td>
<td>Weight of non-price evaluation criteria</td>
<td>Share of new bidders</td>
</tr>
<tr>
<td><strong>Contract implementation</strong></td>
<td><strong>Contract implementation</strong></td>
<td>Share of small or medium enterprise bidders</td>
</tr>
<tr>
<td>Cost overruns</td>
<td>Variation orders</td>
<td>Share of woman-owned enterprise bidders</td>
</tr>
<tr>
<td>Time overruns</td>
<td>Renegotiations</td>
<td>Share of international bidders</td>
</tr>
<tr>
<td>Payment time</td>
<td>Disputes</td>
<td>Network of bidders</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Winning rebate</td>
</tr>
</tbody>
</table>

11. Annex 6 provides a more detailed list of procurement indicators, including the necessary data inputs.

12. Red flags can be useful in identifying unusual patterns in certain markets but these patterns are not sufficient evidence of misbehaviors. One reason why red flags cannot provide sufficient proof of corruption or collusion is that, by design, these data-driven methods can produce false positives (flagging cases that do not merit further scrutiny) or false negatives (failing to identify cases that do merit further scrutiny).
Price efficiency is of key interest with respect to the objective of using public procurement to generate fiscal savings, however this indicator is very difficult to measure in a reliable way. An indicator that is often constructed with public procurement data is “public procurement savings,” calculated as the difference between the awarded contract value and the contract value estimated at the time of the tender publication. This indicator is relatively easy to construct, but it can be problematic for various reasons. First, it can be imprecise. The estimated contract value should represent the expected market price for the purchase, but in practice, it is indicated by contracting entities and it may be based on information that is generic, outdated, or disconnected from the actual market conditions of a given locality. Second, if savings indicators are used for performance evaluations of contracting entities or officers, there might be incentives to inflate the estimated contract values which can create an illusion of savings. This is especially the case in competitive markets where, in equilibrium, firms bid very close to their reserve price and therefore the estimated contract values indicated in the tenders are inconsequential. Third, this savings indicator cannot be informative on which strategies could help in improving the efficiency and performance of the public procurement system. Instead, actionable recommendations can be generated from other types of analysis, such as the analysis of drivers of performance that identifies the tender and contract characteristics associated with lower prices for similar purchases (see section 2.3 for more details on the analysis of drivers of performance). For example, Ferraresi et al. (2021) estimate a 2 to 8 percent reductions in regional per capita health expenditures in Italy after the introduction of procurement centralization. The analysis of the price distribution for similar purchases could also provide savings estimates from scenario simulations. For instance, as in Best et al. (2017), considering a scenario where the least performing officers or contracting entities improved their performance at the level of the median officer or entity.

Contract implementation quality is another key dimension of interest with respect to the objective of achieving value for money while ensuring the quality of the procured goods, works, and services, but its measurement requires further development and research. Research studies have tested various approaches to measuring contract implementation quality. However, there has been no agreed upon strategy yet, and this is a dimension where data constraints are particularly binding. One option would be to integrate public procurement data with other data sources in order to create contract implementation and quality indicators, such as from audits data, contract supervision data, and project management reports (section 2.4 and section 3 provide more details on data integration). Ad-hoc data collections could also be considered for specific sectors, for example, through engineering assessments of the materials used for constructing infrastructure projects (Olken 2007) or through visits to hospitals to verify the availability of medicines, their expiration dates, and other proxies of quality. With respect to the construction sector, recent advancements in technologies can be considered for monitoring the progress of construction works (e.g., the use of drones or satellite images) and their quality (e.g., remote sensing technologies to evaluate the thickness and hardness of pavement roads or leakages in water supply or drainage projects), or for efficiently and periodically collecting inputs from citizens (e.g., monitoring activities using mobile apps). More work is needed to assess the pros and cons of different measurement strategies, particularly in terms of their objectivity and scalability.

Beyond the transactional view of public procurement, there is an increasing interest in measuring dimensions related to the strategic role that it can play in promoting inclusive and sustainable growth and in achieving socioeconomic and environmental objectives. Increasingly, countries have been adopting policies and strategies to enhance the inclusiveness and sustainability of public procurement. However, there is no clear evidence yet on the best public procurement strategies and policies to achieve these socioeconomic and environmental outcomes and balance the potential trade-offs with efficiency, nor a definitive measurement framework to monitor the inclusiveness and sustainability of public procurement (section 2.2 discusses the trade-offs in public procurement in...
further detail). With respect to dimensions related to economic growth and inclusiveness, some countries have progressed in connecting public procurement data and firms’ data, which allows for the construction of firm-related procurement indicators, such as employees’ growth in firms working under governments’ contracts, the share of bids submitted by SMEs, and the share of contracts awarded to SMEs. With respect to dimensions related to sustainability, progress in measuring green public procurement (GPP) has been slower, which is complicated by the various GPP approaches in tender design. For example, green criteria can be included at different stages of the procurement process, such as green eligibility criteria and green evaluation criteria (Testa et al., 2012), however, when measurement efforts are done poorly, they can misrepresent the real progress or there could be a risk of “greenwashing” (Open Contracting Partnership 2021). This lack of standardized measurements impedes advancing the empirical literature on the effectiveness of different GPP policy alternatives, and prevents governments and the civil society from monitoring the implementation of GPP legislation.

2.2 Exploring the Links between Drivers and Outcomes

Public procurement outcomes are shaped by a variety of factors. Figure 1 presents a framework to conceptualize these relationships: (i) the policy and regulatory environment, such as whether the country has an eGP system or whether the public procurement law mandates the institutionalization of a complaints mechanism; (ii) the system inputs, such as the number of procurement officers and their skills; (iii) the characteristics of tenders and processes, such as whether a competitive auction or a direct contracting modality is implemented for a given purchases; (iv) the norms and behaviors in the public administration, including tolerance for malpractices, attitudes, and motivations of procurement officers; and (v) the private market dynamics, such as the availability of suppliers in a market, collusion risks, and behavior of suppliers during contract execution. These factors can be interrelated and can influence economy, efficiency, integrity, and social and environmental outcomes. For example, Bauhr et al. (2020) demonstrate that openness and integrity of competitive markets can be improved by ex-ante transparency (i.e., completeness of information in call for tenders); and World Bank (2021c) shows that the use of centralized procurement for common goods and services can increase competition, reduce prices, and maximize procedure efficiency. It is important to consider these multi-layered interconnections and structure public procurement analysis using a logic framework of inputs, outputs, and outcomes.

Understanding the drivers of procurement outcomes can guide policy making on the most important factors that impact on performance and that should be considered when developing new procurement strategies or reforms. It can inform the design of new strategies or reform, prioritize policies, direct further analysis, and correct the course of actions if certain policies do not lead to the expected results. For example, if a country’s processes which started in the last quarter of the year are associated with higher unit prices, it can be recommended to the government to plan better in order to address procurement demands, avoid bunching in the last months of the year, and reduce seasonal price fluctuations. As another example, if a country’s competitive tenders are not associated with higher bidder participation, it can be recommended to the government to further explore other barriers to firm participation in public procurement. The results of this type of analysis can be particularly useful because they help identify improvements that can be achieved through changes in practices and behaviors (e.g., consolidation of procurement demands to achieve economy of scale), without necessarily requiring structural changes in policies, rules, regulations, and IT infrastructure.

17. For example, Latvia calls its procurement “green” when at least one green criterion is applied, and it represents at least 5 percent of the total contract amount. In theory, the country could label its procurement 100 percent green while only 5 percent of its purchase are green or from sustainable sources (Open Contracting Partnership 2021).
18. The Open Contracting Partnership (2021) provides a guidance note on measuring green public procurement.
Regression analysis is a standard tool for identifying the drivers of the procurement outcomes of interest, but typically it is not sufficient for uncovering causal relations and for conducting impact evaluations, unless the study is carried out under specific assumptions or designs. The main objective of regression analysis in public procurement is to assess whether and how procurement outcomes are explained by a set of potential drivers. This is achieved by comparing procurement outcomes among entities or purchases with different characteristics, following the logical framework depicted in Figure 1. The reliability of the results of regression analysis depends on whether the compared purchases are as similar as possible, except for the potential drivers being analyzed. This is typically achieved by estimating regression models that include a set of fixed effects and control variables, and therefore avoiding comparing, for example, purchases of vehicles with purchases of office supplies (see further details in Annex 1.5).\(^{19}\) For this type of analysis, it is necessary to use granular data as much as possible, with purchases disaggregated at product level and product codes\(^{20}\) (e.g., CPV or UNSPSC) associated to purchase items.\(^{21}\) For example, figure 2 compares unit prices and the probability that the contract is awarded to an SME between framework agreements and open methods, but only for purchases for the same product and by the same contracting entity and with contract signature in the same year and quarter.

---

19. Even when the regression model includes a comprehensive set of control variables, a series of potential endogeneity concerns remain, which prevents treating results as definitive indication on causal effects (Angrist and Pischke 2009). This is the reason why this note stresses the importance of considering impact evaluations for these purposes (section 2.2 and annex 3).

20. Products codes are used in procurement to group goods, works, and services into standardized categories. There are various international systems, such as the UNSPCS from the United Nations, and CPV from the European Union.

21. The [Procurement Anticorruption and Transparency (ProACT) platform](https://www.worldbank.org/en/programs/procurement) provides a description of different product classifications. It is based on an algorithm to harmonize different classifications and enable comparisons across countries and purchases.
2.3 Using Data Analytical Tools to Analyze Trade-Offs and Incentives in Public Procurement

A comprehensive approach to Procurement Data Analytics allows a correct assessment of the potential trade-offs associated to procurement policies and the identification of the most feasible policy recommendations. Given the multi-dimensional nature of public procurement, it is natural that there are trade-offs associated to policies and strategies, especially with respect to monetary versus non-monetary outcomes. For example, there is increasing evidence that winning government contracts can be an engine for firm growth, however, the costs and benefits of alternative policy options to incentivize SMEs’ participation, especially in terms of loss of efficiency, need to be defined. This point is illustrated by World Bank (2021c), which shows that in Brazil, the use of framework agreements can reduce unit prices and avoid repetitive processes, but it can also discourage SMEs’ participation and the likelihood that they are awarded a contract (figure 2). As another example, governments considering adopting green public procurement might worry about the risks of decreasing efficiency, for example, due to additional regulations and requirements and reduction of firm participation. These concerns cannot be convincingly addressed from a purely theoretical point of view. Instead, providing robust knowledge on the actual costs and benefits of GPP can effectively support governments in taking informed decisions and might remove some of the concerns that prevent a broader GPP adoption. Empirical evidence and objective facts are ideal elements in understanding the potential trade-offs between competing policy objectives and in informing policy dialogue.

> > >

FIGURE 2 - Regression Analysis, Framework Agreements vs Other Open Methods, Brazil

<table>
<thead>
<tr>
<th>Outcome of interest</th>
<th>Unit price (log)</th>
<th>SME winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framework agreements vs other open methods</td>
<td>-0.0919**</td>
<td>-0.0919***</td>
</tr>
<tr>
<td></td>
<td>(0.0407)</td>
<td>(0.00582)</td>
</tr>
<tr>
<td>Observations</td>
<td>172,605</td>
<td>166,399</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.910</td>
<td>0.566</td>
</tr>
</tbody>
</table>


Note: Model specifications: Comparing FAs and non-FA open methods for the purchase of the same product by the same entity (product - entity FE), with year and quarter FE. Robust standar errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

---

22. Article 33(1) of Directive 2014/24/EU (the European Procurement Directive) states, "A framework agreement means an agreement between one or more contracting authorities and one or more economic operators, the purpose of which is to establish the terms governing contracts to be awarded during a given period, in particular with regard to price and, where appropriate, the quantity envisaged." Guidance on framework agreements is provided by a recent report by the World Bank (2021).
Unintended adverse effects of new policies or strategies can be minimized by systematically analyzing the incentives of relevant procurement actors during the design stage and by conducting impact evaluations during the piloting stage before full roll-out. Procurement regulations and practices may distort incentives and may have unintended adverse effects on procurement outcomes, if not accurately designed, piloted, and evaluated. For example, Palguta and Pertold (2017) show that contract value thresholds, which are often used to balance the need to incentivize more competition for higher value contracts and guarantee more flexibility for smaller contracts, may create incentives for manipulating contract values (figure 3). Bandiera et al. (2021) demonstrate that the introduction of monetary incentives may not be sufficient to improve the performance of procurement officers if the sources of inefficiencies are derived from factors unrelated to officers’ lack of motivation. Gerardino et al. (2017) show that, while audits are necessary to monitor compliance to laws and regulations, their design and targeting can disincentivize the use of competitive methods if procurement officers expect open auctions to be more likely to be audited than single source contracts due to their complexity. Similarly, procurement officers may be less likely to comply to regulations that are difficult for auditors to monitor, such as the application of preferential policies for SMEs or the application of green award criteria. During the design stage of new procurement policies or strategies, studying the incentives of procurement officers, firms, and organizations can be helpful in (i) understanding the root causes of observed outcomes (e.g., why competitive processes are not used as expected); (ii) designing policy interventions that can effectively address them (e.g., reduce the time costs for procurement officers associated with competitive processes); (iii) examining the trade-offs of employing specific practices or regulations; and (iv) customizing “good practice” reforms to specific country contexts. Impact evaluations (described in further detail in section 4.2) can be helpful during the piloting stage of new procurement policies or strategies to assess, before full-scale implementation, whether they are generating the expected results, and to correct the course of action if necessary.

> > >

**FIGURE 3 - Bunching of Contracts Right below Legislative Contract Value Thresholds**


Note: The series shown in bars is a histogram of the anticipated value of construction works, relative to the threshold. The solid distribution beneath the empirical distribution is a seventh-degree polynomial fitted to the empirical distribution, excluding points CZK 750,000 or less below the threshold.

23. Some public procurement legislations establish that a different regulatory framework applies to contracts above or below a given threshold.

24. For example, in some countries, procuring entities are required to reserve a given quote of their spending for SMEs, but it is challenging for auditors to monitor the compliance with this requirement if the public procurement data does not include a tag to identify contracts awarded to SMEs.
2.4 Linking Public Procurement Data with Other Dimensions of the Public Sector and Public Administration

Integration of public sector data is a necessary step for the realization of a “whole-of-government” approach from a data perspective and is a promising area for extensions in governance analytics and research, including in public procurement. Public procurement is multi-dimensional, and it is critically interconnected with other functions of the public sector and public administration. The integration of public procurement with other public sector data—such as public finance management, budgeting, human resources management, justice, tax administration, auditing, and services delivery—can provide a comprehensive picture of the procurement function, increase accountability, holistically explore the environment within which procurement is conducted, and enable the government to develop innovative and impactful procurement strategies. Despite this potential, integrating e-procurement systems into other e-government systems is not yet a common practice. For example, based on a 2016 review of public procurement systems, in OECD countries, eGP systems are most often integrated with business registries (8 countries), tax registries (7 countries), budgeting systems (6 countries), and social security databases (6 countries) (OECD 2017). This is an increasingly important area of work and innovations where there could be demand for diagnostics on the existing data structure and for technical assistance in implementing data integration solutions, especially for World Bank’s client countries that already started digitizing their government information systems but without a “whole-of-government” approach for data integration.

Recent examples from the academic literature and World Bank projects demonstrate that the integration of public sector data can highlight patterns and construct performance or integrity indicators that would not be possible if each dataset is analyzed in isolation, therefore strengthening the precision and depth of the analysis outputs and policy recommendations. One example is linking tax registries and public procurement data. Data on tax filings by firms could be useful in characterizing the firms operating in the public procurement markets such as their size and the link between public procurement and the growth of firms (Ferraz et al. 2015), and in assessing the effectiveness of policies that intend to favor the participation of SMEs in public procurement. Another example is linking public procurement data with complaints data and judicial data. In the context of Italy, Coviello et al. (2018) demonstrated the implications of inefficient courts on procurement outcomes, such as longer delays on the delivery of public works, higher likelihood that contracts are awarded to larger suppliers, and higher shares of postponed payments after delivery. A recent World Bank project for Brazil connected publicly available datasets and public sector datasets provided by the government, including procurement contracts data, sanctioned firms from various jurisdictions, e-invoices, public sector payroll, corporate data and georeferencing, labor relations, government transfer and social benefits, electoral data, and online news (World Bank, 2020d). Figure 4 shows some of the novel governance integrity risk indicators on potential fraud or misuse of public funds that can be constructed from this unique integrated public sector dataset. For example, one of the red flags captures whether any owner of a contract firm in a procurement process is a public official in the same contracting entity that signed that contract, which is constructed from linking procurement contracts data, public sector employees’ data, and firms’ beneficial ownership data. As another example, Bulgaria’s 2022 Country Economic Memorandum (World Bank, forthcoming - a) includes a public procurement chapter which integrates public procurement data, a database of politically connected firms, and firm-level data (Orbis). The chapter analyzes favoritism in public procurement of certain companies, such as politically connected firms and state-owned enterprises, and its consequences on private sector competition and performance of firms that do not enjoy preferential treatments (e.g., unconnected firms).
FIGURE 4 - Governance Risk Indicators Constructed from Integrated Public Sector Data

- **Inconsistency with partners**
  - Partner receiving government cash transfer
  - Partner has a low-salary in another firm or government agency

- **Conflict of interest**
  - Partner is a public official with contract in the same agency s/he works with
  - Firm contributed to political campaign
  - Elected official is partner of a firm that has multiple contracts in the same jurisdiction

- **Negative track record**
  - Banned firm in different jurisdictions has contacts
  - Partner has been judicially convicted

- **Inconsistent corporate characteristics**
  - Multiple economic activities
  - Date of incorporation vs date of first contract
  - Number of registered employees
  - HQ located in poor neighborhood

- **Atypical procurement patterns**
  - Winning rates & high-cost overruns
  - Consistently participating in processes with low average number of bidders
  - High rate of single bidder in competitive tender

- **Bid rigging risk patterns**
  - Bidders belong to same corporate group
  - Partner of bidder is employed by another bidder
  - Identical proposals
  - High rate of victories vs stubborn loser firm
  - Bidders submitted bids using same IP address, HQ, or phone number

- **Atypical spending patterns/e-invoice**
  - High expenditures in specific products
  - High unit price of specific product
  - Outlier/above average per capita number of specific products

This section describes different data sources that can be used for analyzing, monitoring, and generating evidence in public procurement. Section 3.1 presents public procurement transactions data, and the importance of meaningfully connecting data from different stages of the procurement and contracting cycle, and of collecting data from contracting entities into a centralized database. Section 3.2 discusses the advancements in the open government agenda in public procurement and the international Open Contracting Data Standards. Section 3.3 provides other data sources that could also be used to generate evidence on public procurement, such as administrative data extracted from non-procurement government databases or survey data. This should be considered as a “living” section, as additional data sources useful for Procurement Data Analytics can be identified in the future, such as on green public procurement.

The limitations and risks associated with different data sources should be considered for the application of Procurement Data Analytics tools to be successful. For countries without an eGP system, limited available data on public procurement transactions is an obvious constraint. However, there might be challenges even for countries with an eGP system. For example, direct contracting tenders are often managed outside the eGP system and therefore they are, by design, excluded from the data generated by the eGP system. Data accuracy is another typical challenge, as administrative data extracted from government databases often have errors, typos, or missing information. Context-specific knowledge is necessary in minimizing these risks and in using administrative and non-administrative data sources correctly and thoroughly.
3.1 Public Procurement Transactions

Public procurement transactions data typically includes information on tenders, bids, bid evaluations, contracts, and contract amendments.25 The procurement and contracting cycle is typically divided in the following stages: (i) budget planning and tender preparation; (ii) tender process, bidding process, and bid evaluation; (iii) contract award and contract signing; and (iv) contract execution and monitoring. This data can be organized at the tender, lot, item (product), bid, and contract level. Figure 5 provides a visual representation of how the different levels of public procurement data connect with each other. Data analytics tools can be meaningfully applied when the different levels of public procurement data are connected with each other through identifiers. In addition, the findings of the empirical analysis are more precise and comprehensive if the available data is more granular, and data at product level are strictly necessary for certain types of analysis, such as price modelling and analysis of drivers of performance (see section 2.3).

> > >

**FIGURE 5 - Data Map of Traditional Public Procurement Data**

![Data Map of Traditional Public Procurement Data](image)

Source: Authors’ elaboration.

To ensure that public procurement transactions data is effectively used for decision making, it is necessary that the data is homogeneously collected and maintained across contracting entities and connected to a centralized database. Public procurement data is often decentralized or housed by different institutions responsible for managing different parts of the procurement process, and this may complicate the process of data harmonization and centralization. These challenges can be especially relevant in countries without an eGP system and where reporting of procurement activities to a central authority is not mandatory. However, even countries with an eGP system can experience similar challenges, such as in a decentralized public procurement structure. For instance, in Indonesia, there are hundreds of independent eGP platforms with different standards and formats, and there is no centralized database of public procurement transactions that can be used by the central PPA.26

---

25. Open public procurement data can be used for initial feasibility assessment and for additional cross-checking and validation, but it is recommendable to base the PDA work delivered to client countries on the administrative data provided by the government.

26. Data from these different eGP platforms have been harmonized and centralized through ProACT and other private sector applications, but not yet in a centralized database by the PPA.
3.2 Open Data and Data Standards

Public procurement is one of the state functions with the largest amounts of micro-data publicly available, and this is a critical component of the transparency and accountability agenda and an invaluable resource for researchers. One of the key principles in the open government agenda is to increase accountability to citizens and civil society by making government data publicly available. While publicly disclosing certain types of government data can generate privacy concerns—such as for judicial data, beneficial ownership data, or payroll data—public procurement contracts are often among the first set of government documents to be made publicly available in a country. As a result, while this progress has not been ubiquitous across all countries, there is already an incredible amount of public procurement micro-data publicly available in machine readable formats, which represents an invaluable resource especially for independent research studies or for building open platforms for international public procurement indicators. As an example, the Procurement Anticorruption and Transparency (ProACT)27 platform collects open data from national electronic procurement systems from 46 countries and open data on World Bank and IDB financed contracts for over 100 countries.28 The goal of the platform is to help promote transparency, integrity, and efficiency in public expenditures on goods, works, and services. By standardizing open public procurement data collected from these different data sources, the platform uniquely allows cross-country analysis, such as comparing performance of suppliers across contracting entities and jurisdictions and monitoring the market characteristics of sectors across countries.

These initial achievements should motivate continuing efforts in the open government agenda and in strengthening global standards for public disclosure. Despite the progress in the open government agenda in public procurement, some challenges remain, particularly with respect to the divide between publishing information (e.g., government documents in PDF format) vis-à-vis publishing machine readable data (e.g., data in excel, csv or json), and the variety of standards across countries on how to publish public procurement transactions. ProACT effectively demonstrates the benefits of global open public procurement data, as well as the current limitations to fully realizing this potential due to the limited adoption of global publication standards across jurisdictions.29 The Open Contracting Data Standards by the Open Contracting Partnership (OCP) are the internationally recognized and used standards for disclosing public procurement data, and are a useful resource for governments considering opening their public procurement data or updating their current platform. For example, Honduras’s public procurement portal is a good application of Open Contracting Data Standards. The portal’s data are available in json, csv, and xlsx format, and the UNSPSC standard for the classification of goods and services is used. Data on open auctions, direct contracts (at least until 2019), and purchases on the e-catalogue are also accessible from the portal. Paraguay’s public procurement portal is another good example. The portal’s data are clearly organized according to stages of procurement and contracting cycle. Data visualizations are user-friendly, making data more widely accessible. Data can also be downloaded systematically with the use of APIs.

3.3 Other Relevant Data Sources

Beyond public procurement transactions data, other data sources could also be used to generate evidence on public procurement, such as administrative data extracted from non-procurement government databases or survey data. While additional data sources cannot substitute for the information captured in contracts, tenders, and bids records, a holistic approach to Procurement Data Analytics would help in monitoring and evaluating the public procurement function. First, the increasing interest in using public procurement policies as a strategic tool to achieve socio-economic and environmental objectives (World Bank, 2022a) calls for expanding the spectrum of performance indicators that are typically used to monitor and evaluate public procurement (a performance measurement framework is provided in section 2.1). Second, a comprehensive approach to PDA allows a correct assessment of the trade-offs associated with some procurement policies and the identification of the most feasible policy recommendations to client countries (some examples of trade-offs in public procurement policies are presented in section 2.2). Third, surveys with procurement officers and firms can be used in measuring dimensions that cannot be captured in administrative data – such as views and perceptions – and they can be a particularly good option.

---

27. ProACT is the product of a collaboration between the World Bank and the Government Transparency Institute.
28. The platform was in a prototype stage at the time of preparing this paper.
29. For example, different countries use different product codes, which should be standardized into a common product categorization to enable comparisons of purchases for comparable products across countries. As another example, countries use different procurement methods and labels, which need to be standardized across countries to enable international comparisons in the use of competitive and non-competitive methods.
for countries without an eGP system. Fourth, while there is a great potential in integrating data from different administrative sources, this is still a new area for many countries, and support may be needed in assessing the existing data structure and in implementing data integration solutions (the opportunities for and benefits of data integration are described in section 2.4). Table 2 shows a list of non-traditional data sources that can be used to generate evidence on public procurement, their potential applications to policy work, and relevant academic research literature.

> > >

**TABLE 2 - Additional Data Sources for Procurement Data Analytics Work**

<table>
<thead>
<tr>
<th>Non-Traditional Data Sources</th>
<th>Potential Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual procurement plans</td>
<td>The annual procurement plans include planned purchases and annual budget allocations. This type of data can help in assessing the quality of planning by the contracting entities.</td>
</tr>
<tr>
<td>Complaints</td>
<td>Data on complaints can help in connecting procurement processes and contracts to indicators of redressal mechanisms and accountability and in assessing the effectiveness and efficiency of the system in addressing complaints. Example: The World Bank and the Romanian National Agency for Public Procurement (ANAP) co-created an interactive procurement dashboard (described in detail in section 4.3) that uses both public procurement and complaints data.¹</td>
</tr>
<tr>
<td>Contract execution</td>
<td>Contract execution data may include indicators on price and time overruns, default rates, subcontracting practices, and characteristics of products/services delivered, such as quality, durability, and maintenance costs.</td>
</tr>
<tr>
<td>Payments</td>
<td>Payments data may include indicators on the frequency of payments and the time for payment completion, and can help in assessing as to what extent efficiencies in payments can boost the economic impact of working under a government contract for firms and stimulate competition. Example: Two recent research papers by Barrot and Nanda (2020) and Conti et al. (2021) studied the link between firms’ economic performance and, respectively, payments time regulation on payments delays.</td>
</tr>
<tr>
<td>Audits</td>
<td>Audits data can help in identifying which processes are audited, what are the results of the audits, and the times for making a decision. They can also help in determining whether there could be scope for improving targeting, and whether the existing audit system unintentionally creates roadblocks or misaligns incentives. Example: Gerardino et al. (2017) analyzed public procurement and audits data for Chile in an academic research paper.</td>
</tr>
<tr>
<td>Asset declarations data</td>
<td>Asset declarations data include information about income and wealth of public officials, including their shares in firms. Linking public procurement data with asset declarations data provides additional tools to assess integrity risks in procurement contracts (e.g., connecting changes in the wealth of public officials and the integrity risks of procurement contracts awarded by the contracting entity at which an official is employed). Example: The World Bank conducted a corruption risk mapping exercise for Armenia using public procurement data and asset declarations data.⁶</td>
</tr>
<tr>
<td>Non-traditional data sources</td>
<td>Potential use</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Sanctions</td>
<td>Sanctions data includes a list of debarment decisions, specifying the name of the debarred firm, the period of the debarment, and the reasons for debarment. Linking public procurement and sanctions data can help in determining the costs of wrongdoing in public procurement, the pre-debarment volume of activities of sanctioned firms, and whether there are long-term penalties for debarred firms.</td>
</tr>
<tr>
<td>Example: The ProACT platform(^b) integrates public procurement and sanctions data from various open national and international sources.</td>
<td></td>
</tr>
<tr>
<td>Judicial data</td>
<td>Judicial data include data on complaints, disputes, investigations and sanctions, on judges’ impartiality, and on the efficiency of the justice system. Linking public procurement and judicial data can help in determining how these two core government functions influence each other (e.g., whether judicial investigations have an impact on procurement and contract execution times, which types of procurement procedures are more likely to result in complaints or disputes, whether the risks of appeals are a barrier to firm participation, and whether the efficiency of courts has an impact on the decisions of contracting entities to enforce contract penalties).</td>
</tr>
<tr>
<td>Example: Coviello et al. (2018) analyzed public procurement and courts data for Italy.</td>
<td></td>
</tr>
<tr>
<td>Firm registry data and beneficial ownership data</td>
<td>Linking public procurement data to firm registry data can help in studying, for example, the actual market size and universe of potential suppliers for contracting entities, participation rate in public procurement, the type of firms that participate in public procurement, or the impact of policies on encouraging participation of specific types of firms (e.g., preferential policies for SMEs). Linking public procurement data to beneficial ownership data can help in estimating the share of government contracts and public spending awarded to connected firms, and in assessing integrity risks.</td>
</tr>
<tr>
<td>Example: Bulgaria’s 2022 Country Economic Memorandum (World Bank, forthcoming - a) includes a public procurement chapter which uses public procurement data, a database of politically connected firms in Bulgaria, and firm-level data (described in more detail in section 5).</td>
<td></td>
</tr>
<tr>
<td>Employer-employee data and tax data</td>
<td>Linking public procurement to employer-employee data or tax data can help in assessing the impact of winning a public contract on firms’ growth, firm performance, and job creation.</td>
</tr>
<tr>
<td>Example: Ferraz et al. (2015) analyzed public procurement and employer-employee data for Brazil.</td>
<td></td>
</tr>
<tr>
<td>Quality of service delivery</td>
<td>Linking public procurement to service delivery data can help in assessing how the characteristics of public procurement matter to the quality and availability of services and, ultimately, development outcomes. This is most feasible with a focus on specific sectors or projects (e.g., quality of catering services in schools or availability of drugs in hospitals).</td>
</tr>
<tr>
<td>Example: Gelli et al (2016) are conducting a randomized control trial to study alternative methods of implementing a school feeding program (i.e., comparing standard</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2 continued

<table>
<thead>
<tr>
<th>Non-traditional data sources</th>
<th>Potential use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procurement officers surveys</td>
<td>A survey of procurement officers can help in measuring procurement-related information otherwise unavailable in the administrative data, such as time-use, management practices, the complexity of the business process, the administrative costs of purchasing, and attitudes and motivation of procurement officers. Collecting this type of data can help in assessing how managerial and organization dimensions affect procurement outcomes. Example: The World Bank (2021e) collected data from public officials in Brazil on perceptions and experiences of corruption, with relevant insights on public procurement.</td>
</tr>
<tr>
<td>Firm surveys</td>
<td>A survey of firms can help in determining the link between public procurement and the private sector. It can include perceptions of corruption, estimates of the time and monetary costs for participation in public tenders, assessment on the barriers and disincentives to participate in public procurement, and the inefficiencies during contract execution. Example: The World Bank-Enterprise Survey developed a public procurement module on barriers and challenges experienced by firms with respect to public procurement, piloted in May–June 2021 in Croatia, Romania, and Poland (World Bank, forthcoming - b).d</td>
</tr>
</tbody>
</table>

---

While PDA work typically relies on administrative or survey micro-data, international aggregate data and indicators can be used for initial overviews of national public procurement systems and for connecting de jure and de facto practices. There is an increasing effort to build global databases on public procurement. For example, the World Bank’s Global Public Procurement Database reviews the public procurement legislation and the status of eGP implementation across countries. The World Bank’s Contracting with the Government database assesses public procurement processes, de jure and de facto, over the five main phases in the life of a government project to be implemented by a private company: (i) budgeting and needs assessment; (ii) advertisement and bid submission; (iii) bid opening, evaluation, and contract signing; (iv) contract management; and (v) payment.³⁰ The OECD publishes aggregate indicators on the size of public procurement for OECD countries and few other economies. The World Bank Enterprise Survey Unit publishes a country-level indicator of perceived corruption incidence in public procurement and survey firm-level data for 151 countries that can be used to construct public procurement participation rates—overall and by firm types (World Bank, forthcoming - b). These global aggregate data can be particularly useful for cross-country analysis and for initial analysis to identify gaps and motivate projects.

---

³⁰. To ensure comparability across countries, the methodology explores the process, time, and cost of awarding a public contract for road resurfacing to a domestic-owned, medium-sized limited liability company.
Operational Options Related to Procurement Data Analytics

This section reviews the existing Procurement Data Analytics products that can be offered to World Bank’s client countries. In particular, it covers data-driven public procurement diagnostics (section 4.1), impact evaluations (section 4.2), data monitoring tools and public procurement dashboards (section 4.3), and projects to assess and improve the quality and openness of procurement data (section 4.4). This menu of options is meant to guide the initial dialogue with the client government to identify specific needs and policy interests in public procurement. However, the tools and approaches are flexible and should be customized based on the specific government requests and type of engagement. This list of operational options should be considered as a “living” guide, since it may expand as more possibilities become available in terms of comprehensiveness of data and technologies for linking and analyzing public procurement data. Table 3 describes the potential contribution of each PDA product to the four overarching objectives of the public procurement function conceptualized in section 1 (savings, integrity, economic growth, and inclusiveness and sustainability), and the potential entry points and synergies (discussed in further detail in the next sections and in section 6).
## Table 3 - Summary of Procurement Data Analytics Products

<table>
<thead>
<tr>
<th>Procurement Data Analytics Products</th>
<th>Potential Outputs and Benefits</th>
<th>Potential Entry Points and Synergies</th>
</tr>
</thead>
</table>
| Data-driven public procurement diagnostics | • Monitor the public procurement function.  
• Identify efficiency and integrity gaps and their determinants.  
• Identify areas for targeted actions, such as contracting entities and sectors with worst performance.  
• Systematically rationalize alternative strategies and interventions.  
• Provide policy recommendations to inform reform efforts and the design of new policies and strategies. | SCDs, PERs, PFM assessments, CEMs, MAPS, other procurement diagnostics, other thematic assessments (e.g., anti-corruption). |
| Evaluation of the impact of a procurement activity or strategy | • Evaluate the impact of a policy and/or strategy  
• Generate evidence during the pilot stage of a new policy or strategy, and potentially inform their redesign prior to full-scale implementation.  
• Navigate the political economy of complicated reforms through evidence.  
• Contribute to the creation of global knowledge. | Lending and technical assistance operations, such as eGP adoption or capacity building. |
| Development of sustainable data monitoring tools | • Systematically monitor public procurement.  
• Strengthen the monitoring and evaluation function of public procurement authorities.  
• Create tools for evidence-based policy making in public procurement.  
• Build real-time red flag systems.  
• Contribute to transparency and accountability of public administration. | |
| Assessment and improvement of quality and openness of procurement data | • Improve the quality of public procurement data.  
• Increase transparency and accountability by making procurement data accessible to the public.  
• Exploit the opportunities for evidence-based decision making created by eGP solutions.  
• Create the conditions for meaningfully conducting other activities related to Procurement Data Analytics. | Lending and technical assistance operations, such as eGP adoption or capacity building, as well as other PDA products. |

a. The Methodology for Assessing Procurement Systems (MAPS) is used to assess public procurement systems in their entirety. It is built on four pillars: (i) legal, regulatory, and policy framework; (ii) institutional framework and management capacity; (iii) public procurement operations and market practices; and (iv) accountability, integrity, and transparency of the public procurement system.


The main data source for PDA is public procurement data from administrative sources maintained by the government or collected with the assistance of the World Bank. To ensure data accuracy and the government’s ownership of the project, using data sources provided by the government is recommended, rather than those published on open data platforms. Engaging with the clients during demand identification and data availability assessment process can ensure accuracy and completeness of the data used for analysis. It can also bring clarity on data requests for the project and why they are necessary, therefore creating the conditions for full cooperation throughout project implementation.
Data availability is a necessary condition for most of Procurement Data Analytics activities, and this section distinguishes the operational options most appropriate for countries with and without an eGP system. Countries with an existing eGP system are good candidates for PDA since data is already digitized and available from the eGP system. However, an existing eGP system is not a requirement for data-driven diagnostics, impact evaluations, assessments of data quality, or statistical capacity building. Clients without an eGP system might have specific needs that a data-driven approach could help address. In this regard, while eGP adoption at the global level is a priority, the next sub-sections discuss how PDA products can be tailored to these cases.

4.1 Data-Driven Public Procurement Diagnostics

The goal of data-driven public procurement diagnostics is to monitor the effectiveness of the public procurement function, identify efficiency and integrity gaps and their determinants, provide policy recommendations, outline alternative savings strategies, and identify priority actions. Analytics are typically based on a few years of historical micro-data on public procurement transactions and, where possible, on contract execution and quality of delivery. Typically, the analytical report provides a snapshot of the public procurement system and market dynamics and trends, creates a profile of public goods and services expenditures, and identifies efficiency and integrity gaps. For example, in 2020, the World Bank undertook an assessment of the Bangladesh public procurement system (World Bank 2020a), an eGP solution which had been financed by World Bank operations over several years. In the report, the evaluation used data analytics to assess the impact of the Government of Bangladesh’s “10-percent rule.” This rule was introduced by the Government of Bangladesh in December 2016 for the procurement of works using the Open Tendering Method and mandated the rejection of tenders whose offered price is either 10 percent above or below the estimated cost. The descriptive quantitative analysis of time trends revealed that this rule had undesired consequences, such as decreasing competition (figure 6). These empirical insights were instrumental in providing evidence-based recommendations to the government with respect to the continuation of this rule.

**FIGURE 6 - Assessment of Bangladesh’s 10-Percent Rule**

![Graph showing the fall of average number of bidders per package](Data source: All OTM works packages in e-GP (135,179 Nos.))

**Data source:** All OTM works packages in e-GP (135,179 Nos.)

**Average number of bidders per package**

<table>
<thead>
<tr>
<th>Number of bidders</th>
<th>Monthly average</th>
<th>Quarterly average</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data-driven public procurement diagnostics typically use a set of descriptive and diagnostic analytical methods. Descriptive analysis methods (e.g., Pareto Analysis and Spend Analysis) can provide an overview of the public procurement system. For example, Figure 7 shows the Kraljic Matrix for Colombia, which categorizes goods, services, and works with broadly similar characteristics based on their economic relevance and supply risk, and for which similar procurement strategies might apply. For example, “strategic” products should be closely monitored, “bottleneck” products may necessitate supplier security, “leverage” products may allow buyers to use their market power to negotiate on price and other aspects, while “non-critical” products may require little supervision. Descriptive analysis is mostly useful for monitoring purposes and for further directing analyses or investigations into specific entities, sectors, products, or procurement methods. Diagnostic analytics (e.g., regression analysis) typically supplements descriptive analytics to analyze the links between drivers and outcomes, and this helps deepen understanding of the causes of efficiency or integrity gaps (more details on the analysis of drivers of performance are provided in section 2.3). Both approaches can be used in one analytical report, as much as data allows, based on the specific needs of the client and the project objectives. Annex 1 describes in further detail these analytical methods.
For countries without an eGP system, data availability is an obvious constraint that can be addressed through an agile ad-hoc data collection designed for specific project objectives. In these cases, data could be generated from tenders and contracts paper records, from the application of machine learning and artificial intelligence tools to digitize scanned documents, and through surveys with procurement officers or firms (section 3 describes various data sources that can be used for PDA). Data collection from paper records and scanned documents is labor intensive and therefore should be designed with due attention to resource needs. In these cases, it is important to set realistic expectations with the client on what can be achieved in terms of data coverage and depth given the available resources. Nonetheless, a well-designed ad-hoc data collection and analysis can serve specific purposes, and inform the design of policies and strategies. For example, the World Bank is currently planning to support the public procurement regulator in Laos to set up an online webpage where contracting entities could enter data to a central database, which can then be used to conduct a Spend Analysis to inform the design of framework agreements for commonly procured goods and services.

4.2 Evaluating the Impact of a Procurement Activity or Strategy

The goal of an impact evaluation is to generate evidence on the benefits achieved by a particular intervention, such as a given policy change, reform, program, or project. Impact evaluations inform the client country on whether implemented policies, reforms, or programs resulted in the desired effects and, if not, how they could be restructured. Any policy change or reform can be considered for an impact evaluation, such as changes in the regulatory framework for public procurement, capacity building initiatives, the introduction of preference policies for SMEs, or eGP adoption. For example, in 2014, Bandiera et al. (2021) conducted an

---

31. An example is the activity on “Machine Learning Base Text Processing for PPRA Data”, delivered for the Pakistan Federal Public Procurement Authority under the World Bank project “Open Contracting Data Toolkit” (P167154).

32. Ad-hoc data collections could also be useful for countries with an eGP system for gathering information not yet captured in the eGP system. Section 3 describes in detail various data sources that can be used for procurement data analytics projects.
experiment, in collaboration with the Punjab government in Pakistan, to investigate the impact of providing more discretion to procurement vis-à-vis the impact of a policy to incentivize procurement officers through a pay-for-performance scheme. The study finds that the shift in decision rights from monitors from the Accountant General office to procurement officers reduced average prices by 9 percent without reducing quality, while the introduction of the pay-for-performance scheme was found to be ineffective in reducing prices.

Impact evaluations can contribute across the entire cycle of public procurement reforms. Nonetheless, they are still rarely used in public procurement, and therefore there is a knowledge gap on the effective policies and reforms to achieve savings and socioeconomic and environmental objectives through public procurement. First, during the piloting stage of new policies and reforms, impact evaluations can determine the effectiveness of the intervention before implementing it at full-scale and can help in correcting course of actions if necessary. Second, they can be a powerful tool to navigate the political economy of reforms, especially of complicated ones. This was the case for the eGP adoption in Bangladesh, wherein providing credible evidence on the early gains achieved through eGP was a critical factor to overcome initial opposition to the reform and build political consensus (World Bank 2020b). Third, impact evaluations help in understanding which dimensions of the public procurement system can be improved through a given intervention and which ones would require different types of solutions. For example, early results from the impact evaluation on eGP adoption in Bangladesh (World Bank 2020b) demonstrated that the adoption of eGP was successful in making public procurement more accessible and competitive (figure 8), but it was not a panacea and complementary strategies are necessary to address inefficiencies and corruption risks at all phases of the procurement cycle. Despite a growing academic literature and impact evaluations on public procurement, the existing body of evidence is still scarce and limited to a few countries. This impedes drawing generalizable lessons on optimal policies and strategies to achieve the multi-layered objectives of the public procurement function, and it highlights the need for more impact evaluations in this area.

Figure 8 - Procurement Outcomes under Manual System vs e-Government Procurement in Bangladesh

33. Specifically, early results from the impact evaluation demonstrate that in Bangladesh, eGP has helped reduce integrity risks, increase participation of firms, and reduce the prices offered by firms during the bidding process, but it did not affect the final cost of capital investment projects, time overruns, or quality of contract implementation.

34. World Bank blogs summarize some of the recent papers from the economic literature on public procurement. See Arnal Canudo et al. (2021), and Cocciolo et al. (2021a), Cocciolo et al. (2021b).
Two alternative methodologies can be followed for impact evaluations, depending on the country context and timing of the intervention: retrospective impact evaluations of past interventions or experimental impact evaluations (e.g., randomized controlled trials) of pipeline interventions. The former approach is appropriate for studies whose focus is on interventions that were carried out and completed in the past, and where in public procurement data for the periods before and after the intervention are available or can be collected. For example, a retrospective impact evaluation can estimate the benefits of a previous capacity building program already implemented. The latter approach is feasible when the government is considering or planning an intervention and there is scope for embedding an experimental design into the piloting and roll-out plans. For example, an experimental impact evaluation can estimate the benefits of a new capacity building program that the government is planning to implement. For both approaches it is necessary that, at least for a given period of time, a group (e.g., contracting entities) was exposed to the intervention while another group was not. For example, a capacity building program can be empirically evaluated if, at least for a given period, some procurement officers could take part in it while others could not. Annex 3 provides further details on impact evaluation methods.

35. An experimental impact evaluation may not require the collection of historical public procurement data, but it is feasible only when the government is willing to adjust the piloting and roll-out design to follow an experimental protocol.

Countries that are considering adopting an eGP system are ideal cases to empirically evaluate the impact of eGP. Any impact evaluation of eGP must compare procurement outcomes from transactions through the eGP system and outside the eGP system; therefore it must collect public procurement micro-data from the paper-based system. For countries that are considering adopting eGP, the collection of baseline micro-data of procurement transactions can be used for ex-ante data-driven public procurement diagnostics (section 4.1), as well as for creating the baseline for a potential impact evaluation on the benefits achieved by the new eGP system. The empirical evidence on the benefits of eGP adoption and how to best implement it is still limited, and generating more evidence on these aspects would be useful in motivating and providing guidance to countries that are considering adopting an eGP system, as well as in understanding which inefficiencies can be addressed through IT solutions and which inefficiencies require different types of interventions.

4.3 Development of Sustainable Data Monitoring Tools

Interactive data monitoring tools enable public procurement authorities and other users to track, analyze, and display key procurement indicators through customizable and user-friendly visualizations. Interactive dashboards have four main advantages of interactive dashboards. First, they allow efficient monitoring of performance and compliance of the public procurement system by reducing the costs of periodic data analytics exercises and capacity constraints to implementing data analytics. Second, they allow users to visualize the data with great flexibility, such as focusing on specific levels of government, specific contracting entities, specific regions, or specific sectors. Third, they can be programmed sustainably, requiring minimum maintenance. Fourth, they contribute to creating the data system and capacity in PPAs for independently using procurement data analytics tools.

Data monitoring tools can offer a long-term solution to governments on how to extract valuable information from procurement transactions data, but they can be sustainable only if developed in close collaboration with the government and if complemented by capacity building activities. Data monitoring tools are intended to be used by PPAs and contracting entities for monitoring and evaluation purposes. However, their sustainability depends on whether (i) they are designed based on the needs of the clients and (ii) the PPA (or an equivalent entity) has the necessary capacity to maintain it, periodically update it as new data become available, and make necessary changes in the analysis, indicators, and visualizations as new needs arise. A recent project for the Romanian government provides an illustration on how these objectives can be achieved in practice. Between 2020 and 2021, the World Bank worked with the Romanian National Agency for Public Procurement (ANAP) in developing a data monitoring tool to track key performance indicators of the national public procurement system and facilitate reporting to the EU. The World Bank developed a dashboard in close collaboration with ANAP to ensure that it includes indicators, graphs, and analyses that ANAP needs for their internal and external reporting (figure 9). The World Bank also conducted regular workshops with ANAP to transfer the necessary skills and knowledge to replicate, update, and make changes to the dashboard. This approach was empowering and rewarding for the client, and it created the necessary conditions for the long-term sustainability of the dashboard.

---

36. Beyond the purpose of an impact evaluation, data collection from paper-based records before eGP adoption is a useful exercise that can inform how data could be managed and stored in the local offices. These aspects are further described in section 2.4.
37. Annex 4 provides a detailed guidance on the practical steps for the design of data monitoring tools, including a comparison of alternative statistical softwares.
38. See World Bank. Assessment of the Public Procurement System and Further Support to the Implementation of the Public Procurement Strategy (P169141).
39. The workshops were conducted twice a week throughout the duration of the project with a small group of ANAP staff with expertise in statistical analysis.
Data monitoring tools allow the optimization of technology and big data, and are best suited for countries that already have an eGP system. In countries that are currently planning, undertaking, or refining an eGP modernization reform, these data monitoring tools can be integrated with the eGP system and therefore provide (quasi) real-time analytics on public procurement based on the data automatically generated through the eGP system. These applications can be particularly useful in building real-time red flag systems that would enable the early detection of corruption and integrity risks and inform preventive actions. While this solution is ideal, it may not be immediately feasible for all countries with an eGP system. In the case of Romania, at the time of project design, ANAP did not have direct access to the eGP underlying data. Given the existing data sharing arrangements, the interactive dashboard was designed to be, in the short term, periodically updated through data templates submitted by data providers, while in the long-term, it could be adjusted and directly connected to the eGP system if the institutional setting would allow it to do so. Annex 4 discusses in further detail some technical considerations on how to build interactive dashboards that are sustainable, flexible, and adaptable to changing environments or future needs.

4.4 Assessment and Improvement of Quality and Openness of Procurement Data

The goals of this type of assessments are (i) to review the existing structure of procurement data, the data management system, and the data entry process; (ii) identify existing data gaps and inconsistencies; and (iii) assist countries in enhancing data quality and completeness and making their public procurement data publicly available. Since procurement processes typically involve multiple stakeholders and government institutions, this type of assessments can also map out the existing institutional and administrative procurement, contracting, approval, and auditing processes; the data ownership system; and the data sharing arrangements between institutions.

An assessment on the quality of procurement data and existing data management system can be helpful both for countries with and without an eGP system. For countries
that already have an eGP system, it is important to carefully assess the data generated by the system, especially shortly after implementation. The assessment can use routine checks on consistency and completeness of the data, and provide red flags on missing information as well as solutions for the government on how to maintain, manage, and use high quality data from the eGP system. For countries that are preparing for the implementation of a new eGP system, assessing the existing data system and statistical capacity could be a useful preparatory activity, conducted as part of the eGP readiness assessment. These activities are important in ensuring that the full potential of the new eGP system will be realized, such as in using the enormous amount of automatically digitized data for monitoring, evaluation, and data-driven decision making. For example, in 2021, the World Bank team assessed the public procurement systems in St. Lucia, Dominica, St. Vincent, and Grenada. One of the findings was the lack of standards for recording procurement transactions across contracting entities, which strengthened the argument for adopting eGP to standardize processes and reporting mechanisms. More details on the data-driven components that can be included in an eGP readiness assessment are presented in annex 5.

This type of work could also focus on, or include components for, assisting countries in making their public procurement data publicly available, and for assessing and improving the quality and completeness of the procurement data already publicly available. As discussed in section 3.2, open public procurement data is a critical component of the transparency and accountability agenda and an invaluable resource for researchers. Making procurement micro-data publicly available is quite a technical exercise. A recent project for the Honduras government provides an illustration on the type of support that the country may need for this type of work and the available tools.41 Through this project, two readiness assessments were completed – the Open Data Readiness Assessment42 (World Bank 2021j) and the Open Contracting Assessment (World Bank 2021k).42 Technical assistance was provided for the implementation of the Open Data Strategy, the development of Honduras’ public procurement portal, according to the Open Contracting Data Standards, and capacity building on the use of open data.
Capacity Building

The long-term goal in the PDA agenda is that World Bank’s client countries will have in-house capacity and tools for effectively using their own administrative micro-data for monitoring, assessing, and evaluating the public procurement function. The public procurement function is still transitioning toward a more strategic role and more modern systems, which implies that the majority of PPAs and contracting entities are still in the process of expanding the set of competencies available internally, including statistical and analytical skills. At this stage, it is appropriate for the World Bank to support client countries with developing and showcasing the PDA operational options described in sections 4.1–4.4, which can also have a demonstration effect on the usefulness of PDA tools and of acquiring the necessary skills internally. However, the goal is to empower client countries with the necessary knowledge and tools to effectively conduct these activities independently, so that the use of PDA tools is not just a one-time exercise, but rather, it becomes sustainably embedded in monitoring, evaluation, and decision-making processes in public procurement. Capacity building activities on PDA can include workshops and trainings, building efficient data systems and data analytical tools, and supporting the government in identifying the existing skills gaps and the need for acquiring new profiles (e.g., data analysts).

Capacity building activities can be designed based on the existing statistical and analytical skills in the Public Procurement Authority or contracting entities, and can be guided by the ladder of procurement analytical tools at different levels of complexity (figure 11). Climbing the complexity ladder allows PPAs and contracting entities to build the skills, tools, and systems to exploit the full potential of evidence-based decision making. For example, Spend Analysis provides descriptive information on procurement spending; the analysis of key procurement outcomes indicators identifies whether and where there are gaps; the analysis of drivers of procurement outcomes allows the investigation of the factors to be prioritized by new policies and strategies; and a real-time data monitoring tool allows the construction of a red flag system for preventive anti-corruption measures. Climbing the complexity ladder also implies
acquiring the necessary skills that will allow understanding of the significance of impact evaluation, and causal evidence to assess the impact of new policies or strategies, for example, during the pilot stage before full roll-out. A World Bank blog (Goldstein 2022) summarizes recent evidence that training civil servants in advanced econometrics and causal inference changes the attitudes of participants toward the importance of quantitative analysis in policymaking, their appreciation for causal inference, and evaluating policies before full roll-out.

> > >

**FIGURE 11 - Complexity Ladder for Analysis Tools in Procurement Monitoring and Evaluation**

<table>
<thead>
<tr>
<th>Complexity</th>
<th>Real-time procurement monitoring, data integration and surveys</th>
<th>Real-time monitoring through complex data integration and dynamic dashboard. Regular surveys to understand perceptions of actors.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Impact evaluations</td>
<td>Evaluate new policies and strategies before roll-out Evidence-based decision making</td>
</tr>
<tr>
<td></td>
<td>Regular/annual procurement monitoring</td>
<td>Use all analysis tools as a part of a regular procurement monitoring framework</td>
</tr>
<tr>
<td></td>
<td>Ad-hoc analysis of drivers of procurement outcomes</td>
<td>Analyze key drivers of performance indicators through regression analysis</td>
</tr>
<tr>
<td></td>
<td>Ad-hoc review of key public procurement indicators</td>
<td>Analyze key performance indicators like processing time, competition, and prices</td>
</tr>
<tr>
<td></td>
<td>Spend analysis</td>
<td>Overview of the procurement market, the products being bought, and the contracted suppliers</td>
</tr>
</tbody>
</table>

Source: Authors’ elaboration.

**Capacity building activities can be helpful both for countries with and without an eGP system.** For countries that do not have an eGP system, these activities can be especially useful if they already have a system to centralize some tenders and contracts data, or in preparation of eGP implementation. In particular, the eGP readiness assessment is an ideal tool for assessing whether the country has already an authority (e.g., a statistical office within the Public Procurement Authority in charge of monitoring and analyzing the procurement data) and whether that authority has human resource with the necessary statistical and analytical skills (see details in annex 5.2). Preparatory activities of eGP implementation can include capacity building in this authority, as well as their involvement in the design of the new eGP system to ensure that it will capture all necessary indicators for monitoring and evaluation purposes. For countries that already have an eGP system, the full potential of eGP can be achieved if the PPA and contracting entities have the necessary statistical and analytical skills so that the enormous amount of data generated by the eGP system can be used for monitoring, evaluation, and decision making. Capacity building can be achieved in various ways. For example, the co-creation of the interactive procurement dashboard by the World Bank and ANAP (described in section 4.3) does have strong capacity building components, as the co-creation approach had the exact purpose of transferring statistical and analytical knowledge to ANAP. As another example, in 2021, the World Bank delivered a data-driven procurement diagnostic to the Croatian government, complemented by a one-week training for staff from the Ministry of Economy and Sustainable Development and from the Official Gazette to strengthen in-house statistical capacity for measuring the performance of and monitoring the public procurement system in Croatia.43

---

43. See World Bank project “Measuring Performance of Public Procurement in Croatia” (P173682).
Entry Points and Synergies

This section discusses how World Bank operations can benefit from integrating Procurement Data Analytics products, and the opportunities for building a continued country engagement and business lines in public procurement. Following the potential entry points and synergies of each PDA product summarized in table 3, section 6.1 discusses how PDA work can be integrated into other World Bank assessments and diagnostics. Section 6.2 explains how PDA work can contribute to lending and technical assistance operations, and section 6.3 illustrates how PDA work can create appetite for future World Bank operations. These considerations are based on the practices and experiences garnered from the World Bank teams working on this type of projects.

Public procurement is a transversal activity contributing horizontally to the functioning of the overall public sector; therefore, PDA products can also be relevant for sector-specific work or in support of decentralization programs. Public procurement is a component of public service delivery, for example, for the purchase of textbooks, school uniforms, school meals, and school equipment in the education sector, or for the purchase of medicines and medical equipment in the health sector. In these cases, efficiency, timeliness of the delivery, and quality of purchased goods and services are of critical importance; therefore improvements in the public procurement function can contribute to improving service delivery to citizens. With respect to decentralization programs, the transfer of competencies from the central to the local government often includes the transfer of responsibilities for public procurement; and it requires data systems built to centralize information for supporting the oversight, monitoring, and evaluation functions of central authorities. PDA products can contribute to the achievement of efficiency, integrity, economic impact, and inclusiveness and sustainability of public procurement, and therefore can also be relevant for sector-specific work and decentralization programs.
Creating demand and buy-in from the governments on using and implementing Procurement Data Analytics tools requires a gradual and multi-pronged approach. As in most reform efforts, various stakeholders may oppose changes due to vested interests; therefore identifying the right champions in decision-making positions is key. Initial engagements can be created with the Public Procurement Authority, which is typically in charge of setting policy and regulations, as well as monitoring the public procurement function, thus it is naturally in need of data and analytical tools. Other relevant counterparts are institutions that, by virtue of their role in the government, have an inherent interest in public procurement (e.g., the Anti-Corruption Agency on integrity dimension, and the Ministry of Finance on costs and time savings). For countries with an eGP system, another relevant stakeholder is the eGP agency, which is the critical data provider and the direct counterpart for integrating data monitoring tools into the eGP system, for improving the quality of public procurement data, or for disclosing public procurement data. Initial conversations can showcase previous successful applications and benefits of PDA products in other countries (table 1). Having gauged the interest of the government counterpart(s) on PDA, the World Bank’s support can be offered for any of the PDA products presented in section 4. For countries with an eGP system, a simpler starting point might be to conduct a data-driven public procurement diagnostic and develop a data monitoring tool prototype. More advanced or politically sensitive options can be discussed in successive phases, such as disclosing public procurement data, creating a real-time red flag system, integrating data systems, capacity building, and introducing an evidence-based approach to policy making through impact evaluations. This sequence is based on the advice and lessons learned by task team leaders of previous PDA activities. However, the approach for each country should be tailored fit based on the context and the demands expressed by the counterpart(s).

6.1 Assessments and Diagnostics

Procurement Data Analytics can be delivered as a stand-alone technical assistance and diagnostic work to clients, or it can be integrated into other diagnostics such as Systematic Country Diagnostics (SCDs), Public Expenditure Reviews (PERs), Public Financial Management (PFM) assessments, Country Economic Memorandums (CEMs), Methodology for Assessing Procurement Systems (MAPS), or other procurement diagnostics. Public procurement is a sizeable and multi-dimensional government spending category; therefore data-driven public procurement diagnostics (section 4.1) can be included in SCDs, PERs, and CEMs, and inform the preparation of Country Partnership Frameworks (CPFs). For example, the Argentina’s CPF for FY19–FY22 (World Bank 2019a) recognized that inadequate procurement regulatory framework and fiduciary controls to curb corruption were significant institutional constraints to public service delivery; therefore interventions in public procurement were key components in the country’s CPF strategy for better governance and service delivery. As another example, Bulgaria’s 2022 Country Economic Memorandum (World Bank forthcoming - a) includes a chapter for innovative data analytical work to study the risks of capturing government contracts by certain companies, such as politically connected firms. The study provides policy recommendations on improving the public procurement process in the country to safeguard public funds from being used inefficiently and to prevent public procurement from becoming a mechanism of favoritism undermining business dynamics.

6.2 Lending and Technical Assistance Operations

There are great synergies between PDA products and lending operations with a procurement component. For countries that are considering adopting or strengthening an eGP system, there are great opportunities for strengthening the design, impact, and continuity of eGP operations by integrating Procurement Data Analytics components. One of the greatest benefits of eGP adoption is the automatic generation of digitized data that can be used for monitoring the public procurement function and strengthening it through evidence-based policy making. For example, the eGP readiness assessment can include an evaluation of the existing data management system and statistical capacity (section 4.4); a data system solution can be integrated into the...
eGP infrastructure (section 4.3); and an impact evaluation can be designed to estimate the impact of the newly introduced eGP system on procurement efficiency and integrity (section 4.2). The initial piloting and roll-out stages of a new eGP system can include an assessment of the eGP data (section 4.4) to verify the quality and completeness of the data generated by the eGP system and, if necessary, adjust the eGP system design. The project design for introducing the new eGP system can include specific components on statistical capacity building (section 5), on disclosing the public procurement data to citizens and civil society (section 4.4), and for future data-driven public procurement diagnostics (section 4.1). Beyond eGP operations, the benefits of PDA products could also apply to other lending operations with a procurement component. For example, as in the case of the new World Bank project to support the devolution process in Zambia,\textsuperscript{47} decentralization programs can consider assessing and strengthening the data system infrastructure and capacity for oversight, auditing, monitoring, and evaluation of the public procurement at the local level.

Integration of public sector data can be instrumental for a “whole-of-government” approach, and procurement data related activities can be included in any public sector operation whose objective is to strengthen government information systems and capacity for governance analytics. As an increasing number of countries and systems are digitizing their public sector processes and information, there is growing evidence on the benefits of data integration solutions. This is still a novel area of work, and the World Bank can play a crucial role in supporting client countries to create the necessary data infrastructure, integrate information from various segments of the public sector, and build statistical capacity for effective and efficient big data analytics. Opportunities can arise from broader public sector reforms or from the demand of specific public sector agencies; and their realization requires the coordination and collaboration of units and teams from various sectors, both on the client side and on the World Bank side.

6.3 Building a Continued Country Engagement and Business Line

The various Procurement Data Analytics products should be treated as a comprehensive and integrated menu of options that can be offered to clients and help the World Bank in developing a continued and long-term country engagement. For example, when a data-driven public procurement diagnostic generates a series of policy recommendations, the World Bank can then continue supporting the government in prioritizing a few interventions and designing impact evaluations to learn from the most effective approaches and enhance the efficiency and effectiveness of public procurement. Similarly, data-driven public procurement diagnostics typically include recommendations for improving public procurement data quality and building statistical capacity of government counterparts, which can translate into successive technical assistance support by the World Bank. The work with the Romanian National Agency for Public Procurement (ANAP) provides an illustration of this sequence. In 2019, the World Bank delivered a data-driven public procurement diagnostic, which demonstrated the benefits of using data analytical tools in public procurement and included guidelines for data improvements and statistical capacity building. Between 2020 and 2021, the World Bank continued working with ANAP for the co-creation of a monitoring dashboard, providing an efficient and sustainable tool for monitoring the Romanian procurement system, building capacity in ANAP on data analytics, and fostering a culture of data-driven and evidence-based policy making.\textsuperscript{48}

\textsuperscript{47} See World Bank. Zambia Devolution Support Program (P178492). The project will include a capacity building component for strengthening the core functional areas of local governments, including public procurement. Capacity building will cover human resources; skills; infrastructure; and data for oversight, monitoring, and evaluation of local government activities.

\textsuperscript{48} See World Bank. Assessment of the Public Procurement System and Further Support to the Implementation of the Public Procurement Strategy (P169141).
Conclusion

This paper highlights the different operational options for supporting World Bank’s client countries through Procurement Data Analytics. While some of these tools have already been implemented in several countries, there are various areas of work that are being developed, such as measurement of green public procurement and integration of public sector data. The World Bank has a key leading role to play in these areas, and country engagement in the regions is essential in identifying opportunities for innovations and experimentations.

Procurement Data Analytics mainly seeks to strengthen the efficiency of public procurement and government spending to achieve national objectives, and a stronger focus on these more comprehensive objectives can encourage government take-up. The application of data analytical tools is often associated with the use of corruption flags to uncover malpractices, but this narrative may discourage governments from using and disclosing public procurement data. Procurement Data Analytics can contribute to the anti-corruption agenda, but this is only one of the benefits of evidence-based policy making in public procurement. It can also increase fiscal space and support the achievement of socioeconomic and environmental aspirations. A comprehensive presentation of the potential benefits of Procurement Data Analytics can encourage governments to adopt such analytical tools.49

49. Requirements from international organizations or international treaties could also be another strategy to incentivize governments in opening public procurement data and adopting transparent monitoring and reporting mechanisms. For example, EU member states are mandated to monitor and report key procurement indicators under the directives 2014/23/EU, 2014/24/EU, and 2014/25/EU.
Procurement Data Analytics tools can best contribute to strengthening the national public procurement system when they are part of a broader country engagement and dialogue with government counterparts. A data-driven public procurement diagnostic can determine the existing efficiency and integrity gaps and their drivers. A public procurement data dashboard can strengthen the ability of the Public Procurement Authority to monitor and evaluate the national public procurement system. All the operational options presented in section 4 are most effective when they are not treated in isolation, but rather when they become part of the ongoing dialogue with government counterparts to inform policymaking and reforms, which can also include a series of testing, piloting, and evaluating new policies or strategies.

The ongoing collaboration between Governance Global Units and DIME-Governance intends to provide support and resources to World Bank task teams for the implementation of Procurement Data Analytics tools. The Governance Global Units and DIME-Governance have been collaborating on a research agenda on various public procurement topics, such as on the link between public procurement and private sector growth,50 and for the delivery of PDA products in collaboration with task teams in the regions. Additional support is available for task teams, such as additional references and lessons from previous experiences, assistance in the design and implementation of projects or project components that use Procurement Data Analytics tools, and connections with data analysts from the World Bank or from external research institutions.51

50. For example, the Governance Global Units and DIME-Governance are working on a series of research studies under the World Bank project Public Procurement and Firm Behavior (P177551).
51. A course on Data Analytics in Public Procurement is available at World Bank’s Open Learning Campus website at https://olc.worldbank.org/content/data-analytics-public-procurement-self-paced.
References


Annex 1: Approaches to Data Analytics in Public Procurement
A1.1 Pareto Analysis and ABC Analysis of Suppliers

**What it is:** Pareto Analysis and ABC Analysis of suppliers are tools to identify the products and suppliers that represent the largest spending. These analyses can be done at different levels, such as region, entity, or sector.

Pareto Analysis is named after Vilfredo Pareto, who established that, in the 19th century, roughly 20 percent of the people controlled or owned 80 percent of the wealth in Italy. It was found that this principle, known as the 80/20 rule, could be applied to almost all other distribution scenarios including spending. Applying this principle to the total expenditure of a company on its purchases identifies the items and suppliers that the procurement process can focus on to reduce their spend. A normal distribution will find that 80 percent of spend is made on 20 percent of the purchases made. Alongside other analyses, this information can be used to create procurement strategies to meet their organizational goals. This tool is appropriate in identifying the products and suppliers that represent the largest spending. It can be done at different levels, such as region, entity, or sector.

Pareto Analysis provides the basis for the ABC Analysis of suppliers, sometimes referred to as the 80/20 or the 90/10 rule. The ABC method categorizes suppliers based on their importance in each context, for instance, the evaluation of the cost of items based on how much money is spent to purchase them.

**How it is useful:**

- Pareto Analysis and ABC Analysis of suppliers can be used to identify the products and suppliers with largest budget implications. Improving efficiencies in these areas is expected to have the largest budget implications. Thus, this information is useful in targeting interventions and efforts in areas with the largest budget implications. ABC analysis of suppliers is also helpful in adapting relationship management to a particular type of supplier.
- Applying Pareto Analysis helps organizations in documenting the total expenditure from their purchases within the procurement process. The procurement team can analyze the materials, products, and services being acquired to categorize those that can be approached to lower expenses and prices. Procurement can then use this data, along with other analysis, to develop procurement plans and fulfill the needs of the organization.

**Limitations:**

- Pareto Analysis focuses on the largest cost purchases and may lead to overlooking the smaller ones. Therefore, Pareto Analysis can lead to overlooking small problems and complications, which if left unresolved, can lead to critical problems.
- Pareto Analysis and ABC Analysis of suppliers cannot inform on the performance gaps of the system nor on how its efficiency can be improved. However, they are useful for monitoring purposes and for conducting a more in-depth analysis.

A1.2 Spend Analysis

**What it is:** Spend Analysis consists of various tools to describe and analyze historical purchasing data, for example, what is being purchased, by whom, and from what suppliers. It can be done at different levels, such as region, entity, or sector.

**How it is useful:**

- Spend Analysis can provide an initial overview of the procurement system. It is a useful first step for organizations to assess whether expenditures and distributions of purchases are in line with their objectives, and to develop a robust procurement strategy. It can enhance organizational performance through prioritization, resource deployment, category management, and economies of scales.
- Spend Analysis allows understanding of the purchase and expenditure patterns and provides useful insights to inform the budgeting process and forecast future expenditures based on the historical purchasing data.

**Limitations:**

- Spend Analysis cannot inform on the system’s performance gaps or how to improve its efficiency. However, it is useful for monitoring purposes and for conducting a more in-depth analysis.
A1.3 Kraljic Portfolio Matrix

**What it is:** The Kraljic Portfolio Matrix is one of the tools for Spend Analysis in public procurement. The Kraljic Portfolio Matrix classifies goods and services into four categories: leverage, strategic, non-critical, and bottleneck (figure A1.1). The categories are based on two dimensions—profit impact and supply risk. Profit impact considers the criticality of a product for a given buyer or a given economy and its budget impact. Supply risk considers the difficulties in procuring items in the market and the likelihood of disruptions in the supply chain. The Kraljic Portfolio Matrix identifies groups of products with broadly similar characteristics, and for which similar procurement strategies might apply. For example, “strategic” products should be closely monitored, “bottleneck” products necessitate supplier security, “leverage” products allow buyers to use their market power to negotiate on price and other aspects, while “non-critical” products require little supervision.

> > >
**FIGURE A1.1 - The Kraljic Portfolio Matrix**

Profit impact and supply risk can be defined based on qualitative assessments conducted by experts through surveys, in-depth interviews, or case studies on specific sectors or buyers; on the other hand, proxies can be defined from public procurement transactions data. Table A1.1 provides an example on how profit impact and supply risk can be defined through an entirely data-driven methodology based only in public procurement transactional micro-data.
TABLE A1.1 - Data-Driven Definition of Supply Risk and Profit Impact Indexes

<table>
<thead>
<tr>
<th>Supply risk</th>
<th>Profit impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Market concentration and entry market</td>
<td>1. Importance of purchasing</td>
</tr>
<tr>
<td>2. Number of potential suppliers</td>
<td>2. Frequency of the purchasing</td>
</tr>
<tr>
<td>3. High/low value item</td>
<td>3. Value/cost of purchase</td>
</tr>
<tr>
<td>4. Contract implementation risk</td>
<td></td>
</tr>
<tr>
<td>Market share of the main supplier for each product/sector</td>
<td>Share of total contracted amount in each product/sector over total contracted</td>
</tr>
<tr>
<td>Number of firms that submitted a bid for a given product/sector</td>
<td>Average number of times in a given time-period that items in each product/sector are purchased</td>
</tr>
<tr>
<td>Average unit price of items for each product/sector</td>
<td>Total/average of value contract per type of product</td>
</tr>
<tr>
<td>Average duration of contract execution within each product/sector; Average time overrun (%) within each product/sector; Average cost overrun (%) for each product/sector</td>
<td></td>
</tr>
</tbody>
</table>

How it is useful:

- The Kraljic Portfolio Matrix is helpful in identifying products with similar characteristics and which similar purchase strategies may be applied. It provides broad principles and purchasing strategies for broad groups of products and services.
- The matrix helps in identifying supply chain weaknesses, including the supply chain’s upstream segment, that is, all activities that occur prior to the product or service entering the purchasing organization. It provides useful information to supply chain professionals to build the right type of supplier relationships and to manage supply disruption risks.

Limitations:

1. The Kraljic Portfolio Matrix categorizes hundreds of products and services into four broad groups. It is a useful initial categorization, but, by design, the categories are broad, as well as the recommended purchasing principles and strategies. More solid recommendations and strategies can be provided through more in-depth analysis.

2. The Kraljic Portfolio Matrix is a momentary picture of purchases. It is not designed to capture current or rapidly changing market conditions, nor the power dynamics between the buyers and suppliers. Alternative approaches may be preferable to monitor and capture present or rapidly changing market conditions, for example, if the country procurement system keeps separate records of unit estimated prices derived from regular and frequent market condition reviews. The COVID-19 pandemic is a clear example of this limitation, since it led to products that typically have low supply risk and low profit impact (e.g., personal protective equipment) which are critically essential for buyers on an international scale, while uncertainties in the market and the high demands drove an increase in the supply risk.
A1.4 Analysis of Performance and Compliance

**What it is:** Analysis of performance and compliance is a descriptive analysis on performance indicators and compliance to public procurement regulation. It can be done at different levels, such as by region, entity, or sector. Compliance is an integral part of risk management within an organization, and analysis of compliance assesses whether the organization’s practices align with public procurement laws and regulations. Analysis of compliance considers dimensions such as which processes are used, whether entities opt out from centralized procurement options, or data transparency indicators. Analysis of performance considers indicators such as competition, processing time, unit prices, cost overruns, or time overruns.

**How it is useful:**

- Analysis of compliance in the procurement sector monitors to what extent contracting entities comply to existing laws and regulations; therefore it can guide ex-ante and ex-post audits and help prioritize investigations. It can also inform reform efforts to increase enforcement and actions to improve control processes and prevent fraud and corruption.
- Analysis of performance and compliance can identify entities, sectors, or procedures with the most relevant efficiency gaps or integrity risks. This information can be useful in conducting more in-depth analysis and in focusing reform efforts into specific entities, sectors, or procedures.

**Limitations:**

- The analysis of performance and compliance identifies efficiency gaps and risks, but it cannot explain the origins and causes of these gaps and risks; therefore it cannot provide definitive information on how to best address them. For example, the analysis of compliance might identify a group of contracting entities that are more likely to fail to comply to procurement rules, but it cannot explain why there might be some compliance gaps and how to improve the integrity of procurement procedures in these entities.

A1.5 Regression Analysis

**What it is:** Regression analysis estimates the empirical relationship (correlation) between outcome variables and input variables.

**Input variables** are considered as the “drivers” of the outcomes of interest. In statistical perspective, they are referred to as the “independent variables” or “explanatory variables.” For analysis of public procurement, input variables represent all choices that are taken by the contracting entity before the publication of the tender, such as the procurement method, the selection method, bid submission time, and month of publication. **Outcome variables** are the outcomes of interest for the analysis, such as unit prices, time overruns, or cost overruns.

In public procurement, regression analysis can be used to empirically estimate how unit prices (outcome variable) vary with the submission time, or whether they are affected by the choice of the procurement method (input variables).

**Intermediate outcomes** can be perceived as the link between input variables and final outcomes (figure 1). For example, it is often observed that contract unit prices are affected by the procurement method, and one channel that can explain this relationship is the higher number of bidders observed when competitive procurement methods are used. In this example, the number of bidders is an intermediate outcome, affected by the procurement method (input variable) and influencing the relationship between procurement method and unit prices (final outcome). From an econometric point of view, it is advisable to estimate the relationships between input variables and intermediate outcomes independently. The regression models that link final outcomes and input variables should not include intermediate outcomes. For example, a regression model with unit prices as outcome variable and both submission period and number of bidders as explanatory variables are not well identified; therefore the estimates from this regression model should be considered as suggestive and non-conclusive.

The objective of regression analysis is to understand how procurement outcomes are affected by process characteristics and the decisions made by procurement officers. Regression analysis is an empirical technique and it fully relies on the
variables captured in the data. However, each procurement purchase is unique, and its characteristics and dynamics cannot be fully observed and measured. Regression analysis allows comparison of purchases and understanding of patterns and drivers of procurement outcomes, but it should be clear that the estimates cannot represent causal relationships, unless the analysis is based on an impact evaluation design. Nonetheless, the results of regression analysis can be very informative if purchases are closely or similarly compared to each other, and for this, the choice of the regression model is critical. In case of public procurement analysis, one important model choice is the exclusion of intermediate outcomes (e.g., number of bidders) from the main analysis. Another important model choice is the inclusion of product fixed effects (e.g., using CPV codes), which ensures that only purchases for similar products are compared between each other and therefore the estimates are not driven by the intrinsic differences of products. Finally, a regression model can also include entity fixed effects, and this would lead to the comparison of purchases by the same entity, therefore ensuring that the estimates are not driven by any systematic difference across entities.

**How it is useful:** This analysis allows identification of the drivers of performance and provides insights on potential areas of interventions.

**Limitations:** Except in very specific circumstances and under some strict assumptions, regression analysis on observational data cannot be used to draw conclusion on causality. For example, estimates from regression analysis cannot be used to conclude that “using competitive processes has a causal effect on competition and prices.” Causal relationships between the decisions made by procurement officers and procurement outcomes can be estimated through well-designed impact evaluations.
Annex 2: Impact Evaluations
Methods for causal inference and impact evaluation of public policy is a booming field in economics, as exemplified by two recent Nobel prizes given to (i) Abhijit Banerjee, Esther Duflo, and Michael Kremer in 2019, for their experimental method, and to (ii) David Card, Joshua D. Angrist, and Guido W. Imbens in 2021, for their non-experimental methods to rigorously evaluate policy. While estimating the impact of policies require careful understanding of underlying assumption, the World Bank, in general, and DIME specifically, have the knowledge to help operational colleagues and government partners to pursue these rigorous evaluations.

**What it is:** An impact evaluation provides information about the positive and negative, and the primary and secondary effects of an intervention or a program—intended and unintended, direct and indirect, as long as they are measurable, and data are available. Impact evaluations are based on solid research designs that allow the establishment of the causal impact of the intervention or program on the outcomes of interest.

**How it is useful:** It generates solid evidence on the impact of an intervention or a program on key outcomes of interest. The evidence can be used to monitor, review, and re-design the program or intervention and, ultimately, make the organization more effective. An impact evaluation can be undertaken to improve or reorient an intervention (i.e., for formative purposes) or to inform decisions about whether to continue, discontinue, replicate, or scale up an intervention. For example, an impact evaluation can be built during the piloting stage of a new program or intervention, ensuring its effectiveness before full-scale implementation. An impact evaluation can also help in navigating the political economy of difficult and controversial reforms.

The next sections describe the different types of impact evaluation designs. Section A2.1 discusses experimental impact evaluations (e.g., randomized controlled trials) of new interventions and section A2.2 discusses methods for retrospective impact evaluations of past interventions.

**A2.1 Experimental Designs**

Many frameworks prioritize randomized controlled trials (RCTs) as the most credible and robust evidence of the effectiveness of an intervention. By randomly assigning potential participants to either control or treatment groups, RCTs allow the creation of groups that are equivalent based on initial conditions, both observable and non-observable or unknown factors, and therefore perfectly comparable. Therefore, any ex-post differences between groups can be considered a result of the intervention.

For example, an experimental evaluation of a procurement training program could consist of the random assignment of eligible procurement officers to the training group or the non-training group. Baseline data, though not a strict requirement, are useful in determining whether the random assignment was implemented correctly. The evaluation only requires end-line data (collected after the training is completed) on procurement outcomes of interest for both the training and the non-training group.

RCTs can provide meaningful information on the impact of an intervention and are powerful tools to inform policies and reforms. The following design considerations should be considered when planning for an RCT:

- The random assignment into the treatment and the control group(s) should be feasible and enforceable. An RCT requires strong ownership and commitment from the client in following the impact evaluation design and in gathering the necessary data for analysis.
- The sample size should be sufficient to accurately detect differences between treatment and control group(s) given the expected strength of the intervention and the length of the causal chain.
- The design of an RCT should be established before the implementation of the treatment, at the time of the design of the intervention that is being evaluated. Therefore, the duration of an RCT spans the entire duration of the intervention being evaluated.

**A2.2 Quasi-Experimental Designs**

Quasi-experimental designs seek to mimic what can be achieved in terms of causal inference by a well-constructed RCT. These types of impact evaluations are useful when random allocation to treatment and control groups is not possible; for example, because the program being evaluated has already been implemented. Instead of using random allocation to construct treatment and control groups, quasi-experimental designs rely on other methods for constructing a counterfactual or comparison group.

The key objective of these methods is reducing selection bias, such that differences between the treatment and the
comparison group can reasonably be attributed to the impact of the intervention. The main challenge is that these quasi-experimental designs can provide reliable estimates only under certain conditions which cannot be fully controlled by the researchers.

**Regression discontinuity** is useful where a program, intervention, or a policy applies above (or below) a given cut-off point, but not below (or above it). In the case of public procurement, certain publicity requirements apply only for contracts above a certain threshold, but not for smaller contracts. This method is particularly useful with large populations or samples, so that there is a large number of observations around the cut-off point. To provide reliable estimates, this method requires that there is no sorting around the thresholds. This is a condition that cannot always be guaranteed, but it can be empirically tested; and there are recent methodological advancements for explicitly modeling the selection process and controlling for that in the empirical estimation. In the case of public procurement, there might be thresholds that define the likelihood that contracts are subject to audits, and that are unknown to procurement officers, and this was exactly the estimation strategy used by Gerardino et al. (2017) in order to study the impact of previous audits on future uses of open competitive methods.

**Propensity score matching** allows the creation of a comparison group by matching treated and untreated observations based on a set of relevant observable variables. This method attempts to create two comparable groups, with the only difference that one group was subject to a program or an intervention while the other group was not. For example, in the case of public procurement, this method could use observable process, and product and entity characteristics to match similar purchases. However, this method is based on similar assumptions as with a regression analysis, primarily that after controlling for observable characteristics, the treatment assignment is as good as random. The difference between propensity score matching and regression analysis is that the former relies on a data-driven algorithm to define optimal comparable groups and can increase precision, but it is more demanding in terms of richness of data and numerosity of observations.

**Difference in difference** measures the difference in outcomes between the treatment and the comparison group, and how it changed from the pre-intervention period to the post-intervention period. This method requires that there is a control group that did not receive the program in both periods, before and after the implementation of the program. It also requires that the treatment and the comparison group experienced similar time trends in the period before the implementation of the program. In the case of public procurement, a difference in difference strategy can be used to evaluate a new legislation or regulation, as long as this new legislation or regulation only applies to certain products but not to other similar ones, or this new legislation or regulation was piloted first in certain regions, districts, or provinces.

### A2.3 Non-Experimental Design

#### PRE-POST DESIGN

This method is also known as pretest-posttest design, or before-and-after design, where a measure of an outcome of interest is taken prior to an intervention (pretest), followed by a measure of the same outcome indicator after the intervention (posttest). In the case of public procurement, this method would compare procurement outcomes before and after the introduction of a new legislation.

This pre-post analysis cannot inform on the impact of an intervention unless it can be reasonably assumed that the outcome of interest would have remained the same over time even in absence of the intervention. This is a very strict assumption to meet, since in most cases, it is reasonable to expect that things change over time. In the case of public procurement, even in the absence of policy changes, the private sector market is expected to continue evolving over time and therefore competition levels will increase over time.
Annex 3: Practical Steps for Developing Data Monitoring Tools
A3.1 Considerations When Planning to Develop a Data Monitoring Tool

**Fitness to purpose.** The development of an interactive dashboard may be in response to varied needs—either short-term, mid-term, or long-term goals. It is critical to have this in mind during the design phase to ensure that the goals shape the product life cycle. For instance, sustainability would be critical over the long term, hence there is a need for capacity building. Short-term goals could be regular and timely monitoring of public procurement during emergencies, like during the COVID-19 pandemic.

**Capacity building.** It is important to reflect on the existing statistical and programming capacity of the client since these elements are critical in ensuring the use of the platform and its sustainability. There are two broad aspects related to capacity building. The first aspect is about the client’s capacity to use the dashboard, since the goal is to empower the client to use data-driven decision making. The second aspect is about the capacity to maintain and update the dashboard when new data become available or to make any future changes to the dashboard to introduce new visualizations or new analysis depending on the business needs. Both these aspects may require training in cases where there are existing knowledge gaps.

A3.2 Steps in Developing a Monitoring Dashboard

**Assessing the existing data system.** The existing data system should be assessed for data quality and availability. Data should be accessed from the eGP system and can be complemented by other available data sources. A detailed analysis of the available public procurement data should be done to identify any existing data gaps. Additionally, it is necessary to assess the completeness and quality of the available data. These assessments are necessary in developing the dashboard consistently with the existing data system and in designing data validation steps to prevent and correct data errors.

**Assessing the existing statistical knowledge in the entity.** It is important that the existing statistical capacity in the public procurement authority (and other relevant users) is assessed. Procurement Data Analytics is founded on statistical concepts, hence there is a need for basic statistical knowledge such as on mean, median, and regression analysis. In addition, statistical knowledge is necessary in order to interpret the empirical results and draw actionable insights from the data.

**Assessing the existing coding knowledge in the entity.** It is essential to assess the existing coding capacity in the public procurement authority, for example, the programming language or software currently used and at which level. Different types of software can be recommended depending on the existing capacity of the institution. For instance, R and Python are more flexible tools, but they have higher barriers to entry for users with little background in programming languages.

**Discussion on alternatives to the software for the dashboard.** It is advisable to compare all available alternative softwares that can be used to build the dashboard. This will aid in identifying the best solution to the client’s need, which will also depend on the existing coding knowledge in the public procurement authority. For example, R has a wide array of functions, it is open source, and it is very flexible, but packages with fewer users may not have high stability and quality. Python is easy to use, it is open source, and is fast and dependable, but it may present higher barriers to entry for users with little background in programming languages. Tableau does offer an excellent service for data visualization and can use multiple types of data files, but it is expensive and not well-designed for replication and reproduction. These aspects are important to consider when designing the dashboard. Table A3.1 presents in detail the pros and cons of the various types of software.

**Capacity building on developing the dashboard.** There is a need to develop strong internal capacity in the public procurement authority in order to ensure sustainability of the dashboard. Staff will need to be trained on the methodologies and instruments used for building the dashboard, ensuring that they will have the necessary capacity to maintain it. Workshops to train staff should cover topics such as data analysis workflows and construction of indicators.

**Capacity building on how to use the dashboard.** To ensure proper and efficient use of the dashboard, workshops could be offered to inform on the effective use of the dashboard, especially on the interpretation of the indicators and the results of the analysis, and how to use the empirical results for decision making.
**TABLE A3.1 - Comparison of the Various Types of Software**

<table>
<thead>
<tr>
<th></th>
<th>R and Python</th>
<th>SAS and SPSS</th>
<th>STATA</th>
<th>Tableau</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data extraction</td>
<td>✔️</td>
<td>✗</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>Data processing</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>Dashboard creation</td>
<td>✔️</td>
<td>✗</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Cost of data processing</td>
<td>Free</td>
<td>License</td>
<td>License</td>
<td>License</td>
</tr>
<tr>
<td>Cost of dashboard hosting and server</td>
<td>License</td>
<td>Not applicable</td>
<td>Not applicable</td>
<td>License</td>
</tr>
<tr>
<td>Level of coding experience required</td>
<td>Some basic exposure to coding</td>
<td>None required for basic introduction</td>
<td>None required for basic introduction</td>
<td>None required</td>
</tr>
<tr>
<td>Open source coding; Resources available</td>
<td>Well-documented online courses and learning resources</td>
<td>Troubleshooting resources but few online learning courses</td>
<td>Troubleshooting resources but few online learning courses</td>
<td>Well-documented online courses and learning resources</td>
</tr>
<tr>
<td>Additional tools required</td>
<td>None</td>
<td>For dashboard</td>
<td>For dashboard</td>
<td>For data extraction and processing</td>
</tr>
</tbody>
</table>
Annex 4: Data-Driven Components of an eGP Readiness Assessment
A4.1 Activity 1: Data Map Based on Existing Institutional and Administrative Arrangements

Output for this activity: This activity will map the existing institutional and administrative procurement and contracting process (who does what), the different data fields that are owned and managed by different institutions, and the existing institutional arrangements for data sharing.

Rationale:

Procurement processes often involve multiple stakeholders and government institutions. While the procurement process may be initiated by the local contracting entity, the evaluation and approval process may be handled by the approving authority; the payments for the contract may be handled by the treasury; and the contract implementation may be handled by the project implementation team. Thus, the data for each part of the entire procurement and contracting process may be owned and managed by different institutions in the country. The adoption of an eGP system implies that the procurement and contracting processes will need to be streamlined and centralized into the eGP platform, and this has consequences in terms of data creation and data access. Assessing the existing institutional and administrative setting of the procurement and contracting processes, including the data ownership system, is a useful step for designing the new eGP, which will need to identify who produces the data, who owns the data, and who has the rights to access and use the data. This will set the stage for easy adoption of eGP and the need for this activity is imminent in cases where parts of the procurement and contracting process is owned by different government institutions.

A4.2 Activity 2: Assess Whether There Is an Authority in Charge of Procurement Data Analytics

Rationale:

The new eGP system will automatically create an enormous amount of digitized data that can be used for efficient monitoring, evaluation, and reporting of the public procurement function. The first step for realizing this great potential is to assess whether there is an authority with this specific mandate and, eventually, build capacity in this authority on data analytics and data-driven decision making.

Identifying the authority in charge of Procurement Data Analytics will also set the stage for two subsequent elements of the eGP design. First, the data sharing agreement between this authority and the eGP solution provider should also be defined, and ideally this authority in charge of data analytics will be granted direct access to the primary data as stored in the eGP system. Second, there are often concerns about the actual use of newly implemented eGP systems. This authority in charge of data analytics could be responsible for monitoring the actual take-up of the new eGP system, which would require collecting and analyzing data on the tenders and contracts managed through the eGP system as well as on the tenders and contracts managed outside of it.

A4.3 Activity 3: Assess the Data That Will Be Generated by the New eGP System

Outputs for this activity: The new eGP system will generate machine-readable data from the various stages of the procurement cycle. This activity will carefully assess: (i) which data fields are going to be generated by the new eGP system, (ii) at which level of granularity, (iii) whether it will be possible to connect data from different stages of the procurement cycle, and (iv) which entity will be in charge of periodically monitoring eGP data quality and completeness.

Rationale:

Typically, tenders and contracts data are captured in both PDF files (e.g., contract documents) and machine-readable formats. In most cases, the PDF documents are not generated from the machine-readable data, but rather the machine-readable data are inputted manually from the PDF documents. This implies that the information captured in machine-readable formats can vary (e.g., will there be records for unit prices?), as well as the
level of granularity (e.g., will there be records of the number of bidders at tender level or tender-product level?).

Some complications can be derived from the various stages of the procurement cycles (e.g., tender process, bidding process, bid evaluation, contract award, and contract signing) and the multi-layered structure of the process. For example, each tender can be divided into multiple lots, each tender or lot can receive multiple bids, and multiple lots can be aggregated into one contract. In order to meaningfully use this data, it is necessary that the tenders data, lots data, bids data, and contracts data are consistently organized and can be connected.

For these reasons, it should be ensured that the new eGP system will create the necessary data fields, with the right granularity, and with the right links between different datasets to allow the construction of the relevant key performance indicators (KPIs). This assessment should be done based on the relevant KPIs for monitoring and reporting purposes chosen by (or in cooperation with) the authority identified or created in Activity 2 (see section A4.2). This assessment should also support the identification of which entity will be in charge of periodically monitoring eGP data quality and completeness (e.g., the authority identified or created in Activity 2 or the eGP solution provider).
Annex 5: Workflow and Principles for a Data-Driven Public Procurement Diagnostic
A data analytical project consists of various stages, from design to publication, as visualized in figure A5.1.

**FIGURE A5.1 - Data Analysis Process**

1. **Design**
   - Identify data needs
   - Create a data map
   - Conduct sampling and randomization

2. **Acquisition**
   - Do they data really exist?
     - Yes
       - Prepare data license agreement
       - Transfer data securely
     - No
       - Prepare data collection instruments and protocols
       - Collect and transfer data securely

3. **Processing**
   - Remove direct identifiers
   - Tidy data
   - Monitor data quality
   - Clean data

4. **Analysis**
   - Construct indicators
   - Conduct exploratory analysis
   - Polish results
   - Export outputs

5. **Publication**
   - Prepare final research output
   - Assess risk of disclosure
   - Publish data
   - Release reproducibility package

**Source:** World Bank. 2021f. “Report on training provided on the application of the proposed procurement and contract data analytics in Stata.” Delivered for the World Bank project Measuring Performance of Public Procurement in Croatia (P173682).
Principle 1: Define the objectives of the analysis, develop the main questions to answer, identify the indicators of interest and the most appropriate analytical tools, and verify data availability (figure A5.2). This process should be driven by the client’s demands and, ideally, it should be undertaken in collaboration with all government counterparts interested in the findings of the analysis and with the relevant data providers. In practice, this may not be a linear process, as the project objectives may be refined based on the availability and accessibility of data. This stage is important in designing a product that the government will own and that will have concrete impact, clarifying with the clients the exact data requests, and setting common expectations on what the project will be able to achieve. During this planning stage, it is recommendable to write the pre-analysis plan to document the objectives of the analysis, the needs and availability of data, the indicators of interest (a performance measurement framework is provided in section 2.1), and the empirical methodology. Transparency and replicability in the analytical workflow should also be increased.

\[ \text{Figure A5.2 - Process of Defining the Analysis Plan} \]

Source: Authors’ elaboration.

Principle 2: Prepare a collaborative data work environment and set up the workflow structure in such a way that all steps of the analysis are clear, logically organized, transparent, and fully replicable. Annex A5.1 describes in detail some good practices for structuring the data work, such as setting up a standardized and scalable structure for the data files and the script files, separation of the scripts by tasks (e.g., importing, cleaning, and analysis), and the use of master script files. To increase transparency and replicability, all script files and data work documentation can be stored in a GitHub repository, which can be made accessible publicly or to selected users. This is a convenient and increasingly common way to organize a collaborative data work environment—to easily share the data work with team members, peer reviewers, and the client; to build project memory that can be the starting point for future data analytics work for the same country; and to facilitate learning and sharing of good practices between teams. For Procurement

---

1. Client countries have specific demands and needs, and analytical data tools can be adapted to appropriately reflect them. While some governments are interested in price modelling in order to identify potential areas for savings, others are interested in Spend Analysis in order to identify the products that can be suitable for framework agreements, and others in building a data monitoring tool.
Data Analytics, given the confidentiality of most data sharing agreements with the government, the micro-data cannot be published or shared outside the project team, but it is always recommendable to agree in advance with the government on the publication of the analysis output (e.g., diagnostic report) and the script files for the data work.

**Principle 3: Enable a constructive review process.** The ex-ante review process during the preparation of the Concept Note should focus on the pre-analysis plan and assess whether the data sources and analysis methodology are appropriate to achieving the project objectives. While final reviews at the time of the Decision Review Meeting are typically provided on the almost final product and close to the delivery, mid-term reviews (or Quality Enhancement Reviews) are relatively under-utilized processes that can provide constructive inputs for teams and ensure that there is enough time to adjust or expand the analysis as suggested by the reviewers. For Procurement Data Analytics products, it is recommendable that, in addition to public procurement specialists, some of the reviewers have expertise in economics, statistics or data analytics, and that the revision process covers the script files for the data work. Transferring the data work to the client should also be planned accordingly.

**A5.1 Good Practices for Structuring the Data Work**

The structure of the data work folder should give all documents an unambiguous location, and the location and naming of files should describe their purpose and function and be intuitive to all team members. As represented in figure A5.3, a viable folder structure could be one where the data work is organized by the following subfolders:

- **Data,** where all the datasets being used and produced for the analysis is stored;
- **Code,** where all the scripts (e.g., Stata dofiles or R scripts) created as a part of the coding are stored;
- **Documentation,** where, the data descriptions, analysis plan, and any other documentation of the analysis process is stored; and
- **Output,** where all the tables, graphs, and charts produced as a part of the analysis are stored.

The code should be divided into subfolders corresponding to different stages of the data work, such as importing, cleaning, constructing, and analyzing the data (figure A5.4). The data folder should follow the same structure, and therefore be organized in raw datasets and cleaned datasets. Organizing the data and code in this manner can help ensure that the team can find and revise parts of the data process separately and easily.

A master script (in Stata, a master do-file) is a single code script that can be used to execute all the data work for a project. Any team member should be able to run this script and all the data work scripts executed by it by changing only the directory file path in one line of code in the master script. These master files can be used to structure and execute the code. This ensures that all data work is coherently organized, all steps of the code follow a logical sequence, and the full data work can be easily replicated from start to end. This can be achievable only if all steps of the data work are coded in scripts (ideally into the same statistical software, but that’s not strictly required) and no action is taken manually, for example in Excel.

The Documentation folder should contain all the data explanations and analysis notes related to the project, such as:

- data dictionaries (files listing the contents of each dataset);
- codebooks (files describing the values found on all the variables);
- data quality issues (records on the issues and inconsistencies found in the data, and which observations had issues); and
- explanations and decisions (records of explanations and clarifications received by the data providers and decisions taken about data cleaning and analysis).
FIGURE A5.3 - Folder Structure


FIGURE A5.4 - Code Folder Structure

Annex 6: Procurement Indicators and Associated Data Sources
<table>
<thead>
<tr>
<th>Indicators</th>
<th>Data inputs</th>
<th>Notes/Description</th>
</tr>
</thead>
</table>
| **Dimension 1: Time effectiveness of the tendering process** | | Data source: Tender or contract data  
Periods need not all be covered, however, some key periods should be included:  
1. Preparation period  
2. Tender evaluation  
3. Contract signing period  
4. Contract duration |
| Preparation period (2-1) | 1. Initiation date  
2. Publishing date  
3. Submission deadline  
4. Evaluation signing date  
5. Evaluation approval date  
6. Contract signing date  
7. Contract effective date  
8. Contract validity/End date  
9. Contract end/Revised date | |
| Bid submission period (3-2) | | |
| Tender evaluation period (4-3) | | |
| Approval period (5-4) | | |
| Contract signing period (6-5) | | |
| Contract effective period (7-6) | | |
| Contract validity period (8-7) | | |
| Contract duration (9-8) | | |
| **Dimension 2: Accessibility and openness of the tendering process** | | Data source: Bids and tender data |
| Accessibility during submission | 1. Number of documents sold  
2. Number of bid submissions  
3. Number of responsive bids  
4. Number of bids withdrawn  
5. Tender security amount and validity | |
| Accessibility through eligibility | 1. Technical qualifications needed  
2. Past experience needed  
3. Bid security requirement  
4. Firm turnover requirements  
5. Length of eligibility criteria  
6. Length of assessment criteria | Data source: Tender requirements and qualifications |
| Accessibility through location | 1. Location of procurement  
2. Regional eligibility for contract | Data source: Tender data |
| Accessibility through publication | 1. Advertisement of tender on newspapers  
2. Newspaper publication date  
3. Advertisement of tender on website  
4. Portal publication date | Data source: Tender data |
| **Dimension 3: Efficiency in procurement size and consolidation** | | |
| Procurement size | 1. Estimated value  
2. Lot value  
3. Number and type of lots  
4. Quantity of items in lots  
5. Unit price | |
<p>| <strong>Dimension 4: Procurement planning efficiency</strong> | | Data source: Annual procurement plans |
| Expected preparation period | 1. Planned initiation date | |
| Expected bid submission period | 2. Planned publishing date | |
| Expected tender evaluation period | 3. Planned submission deadline | |
| Expected approval period | 4. Planned evaluation signing date | |</p>
<table>
<thead>
<tr>
<th>Indicators</th>
<th>Data inputs</th>
<th>Notes/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected contract signing period</td>
<td>5. Planned evaluation approval date 6. Planned contract signing date</td>
<td></td>
</tr>
<tr>
<td>Expected contract effective period</td>
<td>7. Planned contract effective date 8. Planned contract validity/end date</td>
<td></td>
</tr>
<tr>
<td>Expected contract validity period</td>
<td>9. Planned contract end/revised date</td>
<td>Data source: Budget</td>
</tr>
<tr>
<td>Expected contract duration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budget allocation</td>
<td>1. Total budget for procuring entity 2. Total budget for procurement</td>
<td>Data source: Tender data</td>
</tr>
<tr>
<td></td>
<td>3. Planned value of procurement by item</td>
<td></td>
</tr>
</tbody>
</table>

**Dimension 5: Transparency and integrity of bidding, evaluation, and approval**

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparency in evaluations</td>
<td>1. Share of excluded bids 2. Time for handling complaints 3. Weight of non-price evaluation criteria</td>
</tr>
<tr>
<td>Transparency in awards</td>
<td>1. Notification of award date 2. Where the notification was published</td>
</tr>
</tbody>
</table>

**Dimension 6: Competition and collusion**

<table>
<thead>
<tr>
<th>Process</th>
<th>1. Open procedure</th>
<th>Data source: Tender data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competition</td>
<td>1. Winning rebate 2. Supplier is an incumbent firm 3. Supplier is a small and medium enterprise 4. Supplier is a woman-owned enterprise 5. Supplier is a local firm 6. Supplier is an international firm</td>
<td>Data source: Contract data</td>
</tr>
<tr>
<td>Indicators</td>
<td>Data inputs</td>
<td>Notes/Description</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
</tbody>
</table>
| **Market** | 1. Winning advantage of incumbent  
2. Winning advantage of local firms  
3. Market share of successful bidders  
4. Market concentration | Data source: Bids and contract data |
| **Dimension 7: Contract implementation** | | |
| Number of variations by type | 1. Contract variation date  
2. Nature of variations | |
| Number of Modifications/cancellations | 1. Modification date  
2. Cancellation date  
3. Type of modification/cancellation | |
| Time overruns | 1. Planned contract end date  
2. Revised contract end date | |
| Cost overruns | 1. Planned contract amount  
2. Revised contract amount | |
| **Data that should show heterogeneity in indicators/outcomes** | | |
| Purchaser | 1. Contracting authority name  
2. Location of contracting authority  
3. Contracting authority officer name and designation on tender | |
| Procurement | 1. Category of procurement process  
(open, direct contracting, request for proposal)  
2. Type of procurement  
(goods, works, services)  
3. Procurement product:  
   a) Description/Name of works  
   b) Product code (e.g., CPV) if any | |
| Evaluation committee and approving authority | 1. Name, designation, office of evaluation committee member and chair  
2. Name, designation, office of approving authority | |
| Firms | 1. Name and address of firm  
2. Proprietor name, if any  
3. History and status of debarment  
4. Tax and registration details | |