Improving Productivity Measurement in World Bank Group Interventions

Besart Avdiiu
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>5</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>6</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>7</td>
</tr>
<tr>
<td>2. Productivity Measures</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Overview</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Key Measures to Be Considered in Projects</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Comparison of Productivity Measures to Each Other and across Firms</td>
<td>11</td>
</tr>
<tr>
<td>3. The Way Project Interventions May Be Linked to Effects on Firm Growth and Productivity: The Theory of Change</td>
<td>13</td>
</tr>
<tr>
<td>3.1 Firm Capabilities</td>
<td>14</td>
</tr>
<tr>
<td>3.2 Access to Finance</td>
<td>15</td>
</tr>
<tr>
<td>3.3 Market Access</td>
<td>15</td>
</tr>
<tr>
<td>3.4 Business Environment</td>
<td>16</td>
</tr>
<tr>
<td>4. Productivity Estimation</td>
<td>18</td>
</tr>
<tr>
<td>4.1 TFP Estimation Issues</td>
<td>18</td>
</tr>
<tr>
<td>4.1.1 Functional Form</td>
<td>18</td>
</tr>
<tr>
<td>4.1.2 Aggregation</td>
<td>19</td>
</tr>
<tr>
<td>4.1.3 Endogeneity</td>
<td>19</td>
</tr>
<tr>
<td>4.2 Addressing Endogeneity</td>
<td>20</td>
</tr>
<tr>
<td>4.3 Guidance on Data Collection of Underlying Variables</td>
<td>21</td>
</tr>
</tbody>
</table>
## Contents

5. Attribution of Changes in Productivity to Project Interventions 25

5.1 The Issue of Attribution 25

5.1.1 Randomized Control Trials 27

5.1.2 Difference-in-Difference (DID) 27

5.1.3 Regression Discontinuity Design (RDD) 27

5.1.4 Matching and Synthetic Control 28

5.2 Impact Measurement Challenges 29

5.3 Examples of How Others Have Measured Effects on Productivity 29

5.3.1 Providing Consulting and Securing Orders in the Arab Republic of Egypt 30

5.3.2 Providing Grants to Microenterprises in Ghana and Sri-Lanka 31

6. Key Dos and Don’ts 32

7. Appendix: Technical Details 33

7.1 The General Estimation Approach for TFP 33

7.1.1 Simultaneity Bias 33

7.1.2 Selection Bias 34

7.2 Control Function Approach Details 34

7.3 TFP Estimation in Stata 35

References 37
Abstract

This note aims to help teams implementing interventions to employ firm productivity measures in their projects. It does so by providing guidance on (a) measuring firm productivity, (b) strengthening theories of change as they relate to firm productivity, and (c) attributing productivity changes to projects. The note begins by first introducing the various productivity measures that can be used. It then discusses how project interventions may be linked to effects on firms and any measure of productivity, through a theory of change. Further, it provides guidance on productivity estimation in practice, before discussing how to attribute actual changes in productivity to project interventions. Finally, the note summarizes the key “dos and don’ts” for teams working on productivity.
Acknowledgments

The author, Besart Avdiu (bavdiu@worldbank.org), is an economist in the World Bank’s Finance, Competitiveness and Innovation Global Practice. The author would like to thank Asya Akhlaque and Denis Medvedev for their guidance as well as the peer reviewers: Alvaro Gonzalez, Arti Grover, Elwyn Davies, and Leonardo Iacovone for their insightful comments. The views expressed in this paper do not necessarily represent the views of the World Bank Group, its affiliated organizations, the executive directors of the World Bank Group, or the governments they represent.
Introduction

Productivity is a principal driver of economic growth, accounting for more than half of the differences in GDP per capita across countries (Cusolito and Maloney 2018; Jones 2016). In recent years, however, the productivity engine appears to be sputtering. Global productivity growth is slowing, and the difference in productivity between developing and developed economies is widening (Cusolito and Maloney 2018)—threatening progress on income convergence and poverty reduction. Recent shocks such as the COVID-19 crisis may have lasting negative effects on productivity growth (Apedo-Amah et al. 2020; Cirera et al. 2021), while the growing threat from climate change and policy responses to this threat may also change the productivity dynamics of firms. Therefore, there is a renewed interest in measuring and understanding the effects of public policies and project interventions on firm productivity (Andrews, Criscuolo, and Gal 2016; OECD 2020).

Productivity measures have been considerably refined in recent years. Granular firm-level data are more available, providing opportunities for improved measurement and a better understanding of the limitations of more traditional measures.

The objective of this note is to help teams implementing interventions, at the World Bank Group (WBG) or beyond, strengthen their theories of change and improve the measurement of firm productivity, including the attribution of productivity changes to interventions.

The primary audience of this note is operational teams. Its objective is therefore to present technical concepts based on frontier literature in plain language and provide practical recommendations to task team leaders on how to deal with common methodological and measurement challenges.

The rest of this note is organized as follows. Section 2 introduces the various productivity measures that can be used. Section 3 discusses the way project interventions may be linked to effects on firm growth and any measure of productivity, through a theory of change. Section 4 provides guidance on productivity estimation, including data collection. Section 5 discusses the way to attribute actual changes in productivity to project interventions and gives two examples from the literature. Finally, section 6 summarizes the key “dos and don’ts” for teams working on productivity.
Productivity Measures

2.1 Overview

Productivity is the efficiency with which firms (or industries, whole economies, and so on) convert inputs, such as labor and capital, into output. Mathematically, it is the residual component (or explainer) of output, after accounting for inputs. Hence, productivity measures can differ on the basis of considered outputs (for example, quantities, revenue, or value added) or inputs (for example, just labor or both labor and capital). They can also differ on the basis of methodology used to account for the effect of inputs on output.

The two main categories of productivity measures are labor productivity and total factor productivity (TFP). Labor productivity measures the output or value added per unit of labor (for example, per worker, hours worked, or sometimes even personnel costs). TFP measures how efficiently a firm is able to transform all production inputs (such as labor, capital, materials, and energy) into its products. Depending on the output measure considered, TFP can be divided into revenue-based TFP (TFPR) or quantity-based TFP (TFPQ). Further, based on methodology, one can distinguish the TFP index (which can be revenue or quantity based) from more sophisticated econometric TFP estimates, which this note will refer to as TFPQ and TFPR for simplicity. The next subsection provides more detail on each of these measures, which can serve as project indicators for monitoring the level of productivity.

Any of these measures for projects can be calculated ex ante or ex post if the necessary data are available. However, when one considers productivity measures ex ante, one can more easily ensure that all relevant data will be collected, for example, by administering a (baseline) survey ahead of a project.

2.2 Key Measures to Be Considered in Projects

This section discusses, in more detail, select measures of productivity that can be considered in projects, covering (a) their definition, (b) the dimensions of firm performance that each one does and does not capture, (c) the key assumptions built into each measure, and (d) the data required for each measure. It serves as guidance on which measure to choose under different scenarios.
These measures can also be considered as productivity indicators for projects and can be used to monitor the level of productivity in projects.

**Labor productivity** is typically defined as the value added per unit of labor, usually per worker or hours worked. Instead of value added, revenues or (more rarely) physical output are sometimes used.

Labor productivity has the advantage of requiring simpler data collection. It requires only two variables, which are often widely and readily available: (a) revenues or value added; and (b) a measure of labor, such as number of workers or hours worked.

However, using revenues for labor productivity confounds the effect of increasing inputs and materials with productivity. Therefore, using value added, which omits intermediate inputs, can be more precise than simply using revenues. For example, certain industries, like retail, may have higher average revenues per worker than other industries owing to their more extensive use of intermediates, but that does not mean they are more productive. Nevertheless, the practical difference between using revenues or value added is not trivial, because value added additionally requires data on intermediates or inputs. For labor, hours worked can be a more precise measure than the number of workers, owing to possible differences in the hours worked by each employee. However, data on hours worked are often unavailable.

Importantly, labor productivity focuses on labor inputs only and hence does not consider differences in capital (such as machines and computers). However, capital typically increases the productivity of labor. For example, a worker with a machine or tool may be more efficient than a worker without one. Hence, when one relies on labor productivity only, firms using a lot of capital will appear more productive overall, compared to those that use less, ceteris paribus. Therefore, unlike TFP, labor productivity is not the most accurate measure of overall productivity when production inputs include other factors beyond labor.¹

**Revenue-based total factor productivity (TFPR)** is the traditional measure of productivity (often just called TFP). It is defined “as the part of firm-level revenue or sales that cannot be explained by the contribution of capital, labor, energy, and other [input] factors” (Cusolito and Maloney 2018, 27).

TFPR is based on revenues or value added as a measure of output, which includes both prices and quantities. This can be a disadvantage because prices reflect market power, demand, and quality (see figure 1).

---

**Figure 1: Decomposing Firm Performance**

![Figure 1: Decomposing Firm Performance](image)

Note: K = capital; L = labour; M = materials; TFP = total factor productivity.
Source: Cusolito and Maloney 2018.

---

¹ An option to reduce this issue of accuracy is to include controls for capital intensity in the calculation. However, if data on capital are available, then teams can simply go on to estimate TFP instead of labor productivity.
When markets are not competitive, revenues will reflect not only the efficiency or capacity to turn inputs into output but also markups that are incorporated into the prices. As a result, firms in less competitive markets (such as monopolists) may appear to be more productive than in reality, simply because they can charge higher prices. Similarly, other demand factors can also pollute TFPR as a measure of efficiency, through their effect on prices. For example, when firms sell to both foreign and domestic markets, the price effects embedded in TFPR will also simply reflect demand-side conditions. However, the prices reflected in TFPR are an advantage when it is crucial to consider quality differences that cannot otherwise be captured in the data.

TFPR will require data on revenues (sales) and all inputs. The inputs will typically include labor and the capital stock at the very least but may also include energy and other inputs. Further, proper TFPR estimation will also require data on a proxy variable, typically investment or intermediate inputs (such as the use of materials or energy). Finally, it typically requires panel data with at least two periods for all variables.²

**Quantity-based total factor productivity (TFPQ)** is conceptually similar to TFPR, except it is based on physical outputs instead of revenue, which leads to unique practical challenges. It does not include prices and is therefore a so-called purer measure of technical efficiency.

However, if (vertical) quality differentiation matters, then the lack of prices can be a drawback, because quality is reflected in prices.³ Producing a higher-quality product will require more inputs (or more expensive inputs) than producing a low-quality product, which will make a high-quality firm appear less productive according to TFPQ. For example, consider two firms that use the exact same inputs to produce rugs with the same size. Firm A can transform these into one high-quality rug, while Firm B can produce only one low-quality rug. One cannot fairly say that both firms are equally productive. Nevertheless, they would be equal according to TFPQ because each can produce one rug (output) for a given level of inputs. In reality, the higher-quality rug would command a higher price, making Firm A more productive according to TFPR. However, in rare cases, one can try to overcome this drawback by controlling for quality when estimating TFPR, as done, for example, in Atkin, Khandelwal, and Osman (2017).⁴

Another challenge of TFPQ is how to define output and quantities, especially in a way that is comparable across firms. This can be particularly difficult for services. For example, there is no single clear way to quantify entertainment services. Further, units can vary widely across firms, from number of products to kilograms, liters, hours, and so on. A clear comparison of differently defined units may not be possible, even when they appear similar. For example, spending two hours in a restaurant may not be directly comparable to spending two hours in a cinema. Further, quantities may not be well defined and comparable even within a single firm if it produces multiple products.⁵ To avoid these issues, researchers often use TFPQ among narrowly defined industries with homogenous, perfectly comparable goods, such as in the ready-mix concrete industry (Syverson 2008). TFPR does not face these challenges, because revenues are in currency terms (for example, dollars) and perfectly comparable.

TFPQ requires the same data as TFPR, except it uses output quantities instead of revenues. Alternatively, if data on prices are available in addition to revenue, then quantities can be calculated. Ideally, prices should be available at the firm level, but industry-level prices can be used when there is perfect competition. Without perfect competition, industry-level prices will cause bias because different firms would face different prices. Importantly, TFPQ (like TFPR) requires panel data with at least two time periods.⁶

In practice, firm-level quantities or prices are often not reported or are too difficult to define. Hence, many studies resort to calculating TFPQ with only the necessary data for TFPR by making (strong) assumptions on consumer demand. One approach is to assume a particular demand function. For example, Hsieh and Klenow (2009) calculate TFPQ using a model of monopolistic competition, in which products are imperfect substitutes of each other, and therefore consumers will pay a price above marginal cost. However, these approaches have the disadvantage that their required assumptions are not fully testable empirically and may not hold in reality.

**TFP index** is a special case of TFP estimation, based on stronger assumptions and an easier estimation method requiring less data than that for econometrically estimated

---

2. Although proper econometric TFP estimation requires panel data, in principle it can be estimated without panel data using the ordinary least squares (OLS) residual method. However, this has many issues and is not recommended.

3. Vertical quality means that products can be objectively ranked from better to worse. Horizontal quality differences, such as color differences, cannot be ranked and should therefore not be reflected in prices.

4. This approach requires detailed data on quality and is most easily applied among firms producing a single, well-defined product.

5. Multiproduct firms also pose other challenges, discussed in the estimation section.

6. For TFPQ, like TFPR, panel data are required for proper econometric estimation. However, in principle, TFPQ can be estimated without panel data using the ordinary least squares (OLS) residual method, which has many issues and is not recommended.
TFPR or TFPQ. In particular, panel data over time are not absolutely necessary, though they would improve the estimation significantly. Further, data on a proxy, such as investment, is not necessary. Nevertheless, the TFP index can be revenue or quantity based, thereby corresponding to TFPR or TFPQ but without being fully econometrically estimated.

However, one must be willing to accept the assumptions of perfect competition and constant returns to scale (CRS). These may not be warranted, especially in developing countries. In particular, researchers are often interested in the effects of limited competition on productivity. Further, CRS means that a firm is infinitely scalable at a constant rate, that is, any increase in inputs results in a proportional increase in outputs. However, decreasing returns to scale are often more realistic, because the marginal (additional) benefit of each additional input tends to fall. As a result of these major limitations, the TFP index has been losing popularity in practice.

2.3 Comparison of Productivity Measures to Each Other and across Firms

As discussed, if firms use only labor inputs, then labor productivity is a good measure of firm productivity. However, this is unlikely to often be the case across an entire data set. This condition might apply in rare cases, for example, among some micro informal firms or certain narrowly defined services. Alternatively, labor productivity can be used if the researcher is not interested in overall firm productivity, but only in the productivity of labor.

When firms use inputs beyond labor, overall firm productivity is best captured by TFPR or TFPQ. In theory, TFPQ is a purer measure of technical efficiency. When (vertical) quality differentiation does not matter and output units are easily defined and directly comparable, TFPQ is the most precise productivity measure. Conversely, TFPR is more appropriate when vertical quality differentiation is high and market power is low, such that output prices reflect relevant quality differences instead of markups. In practice, the decision will often not be clear cut and estimation of both is useful. However, proper econometric TFPR and TFPQ estimates require repeated observations per firm (panel data) and a large number of firms, making these measures more difficult to apply in studies with smaller sample sizes.

If the data to estimate TFPR or TFPQ are unavailable, one can alternatively consider labor productivity or the TFP index, depending on the data. If there is perfect competition and firms exhibit constant returns to scale, then the TFP index will be more appropriate. However, as discussed, perfect competition rarely holds in practice, especially in developing countries. Finally, if only data on output (that is, revenue, sales, and value added) and labor (number of workers or hours worked) are available, then labor productivity will be the only possible measure. A summary of the comparisons is given in table 1.

### Table 1: Comparison of Productivity Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor productivity using revenues</td>
<td>• Requires only revenue and labor (for example, number of workers or hours worked)</td>
<td>• Confounds the effect of increasing intermediate inputs and materials with productivity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Confounds the effect of increasing capital and investment with productivity</td>
</tr>
<tr>
<td>Labor productivity using value added</td>
<td>• Requires information on only revenues, materials and intermediate inputs, and labor</td>
<td>• In addition to revenues, requires information on materials and inputs</td>
</tr>
<tr>
<td></td>
<td>• Does not confound increasing intermediate inputs and materials with productivity</td>
<td>• Confounds the effect of increasing capital and investment with productivity</td>
</tr>
</tbody>
</table>
Importantly, one should be careful when comparing the TFP levels of firms across sectors or countries. TFP estimates depend on the output elasticities of capital and labor, which, in turn, depend on the underlying production functions (that is, fundamental technologies) that can differ substantially across sectors and countries. For example, some sectors or countries may be more capital intensive than others. In this case, comparing the TFP values of two firms with different production functions leads to an "apples and oranges" problem, as coined by Bernard and Jones (1996). TFP comparisons are then misleading, because two firms with the same TFP and input levels, but with different output elasticities, will likely see different output values. Therefore, TFP comparisons between firms within the same sector and country are more straightforward, because they are more likely to have similar production functions.

There are several ways to minimize the issue of TFP comparisons between (firms in) different sectors or countries. First, the TFP estimates used for such comparisons may rely on fixed output elasticities. The disadvantage of this approach is that it is ad hoc and the chosen elasticities may not correspond to the true ones for each sector or country. Second, one can control for sector-level productivity and, hence, look at only within-sector differences in productivity. However, this approach precludes explicit cross-industry comparisons, but allows cross-country ones. Third, one can look at TFP growth instead of levels, ideally at the firm level, which filters out some of the fundamental differences across sectors or countries. This approach is likely to be most appropriate for WBG operations, because they are often interested in changes of productivity over time, especially following an intervention. Finally, one can rely on labor productivity, which does not have the problem of differing input elasticities for capital and labor, because only labor is considered. Hence, labor productivity levels can often be compared more clearly between sectors or countries, while TFP can make more sense between firms within a sector, if one assumes a similar production function or technology exists within sectors. However, labor productivity still has the usual caveat of ignoring capital in this case.

Note: TFP = total factor productivity; TFPQ = quantity-based total factor productivity; TFPR = revenue-based total factor productivity.

---

7. As an example of such differences, consider comparing a firm in the hairdressing industry with a mining company, or a country whose economy is mainly reliant on the mining industry compared to one mainly reliant on simple services.

8. These elasticities are often set at two-thirds for employment and one-third for physical capital, as in, for example, Bloom et al. (2020) or Nayyar, Hallward-Driemeier, and Davies (2021).
The Way Project Interventions May Be Linked to Effects on Firm Growth and Productivity: The Theory of Change

This section presents the theories of change that describe how typical project interventions targeted at firms may affect intermediate and long-term outcomes to improve firm growth and productivity. The relevant areas are typically (a) firm capabilities, (b) access to finance, (c) market access, and (d) the business environment. A theory of change will be developed for each area separately, and a schematic overview is shown in figure 2.

Figure 2: Schematic Generalized Theory of Change

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Interventions</th>
</tr>
</thead>
</table>
| Firm Capabilities: Lack of managerial and organizational skills | • Training or mentoring  
• Labor incentives  
• Learning of skills from demand  
• Support to innovate |
| Access to finance: Constrained access to capital and credit | • Application support  
• Grants or Vouchers  
• Guarantees or Collateral  
• Crowdfunding or Fintech |
| Market Access: Lack of local or external demand | • Finding of buyers  
• Securing of orders  
• Marketing/Export support  
• Public procurement  
• Networks/Alliances |
| Business environment: High regulatory compliance cost | • Information  
• Monitoring  
• Feedback channels |

Intermediate outcomes
• Changed practices
• Improved worker’s skills
• Technology adoption
• Investment in innovation
• Capability gained from buyers and competitors
• Secured contracts/sales
• Access to new markets

Long-term outcomes
• Greater sales, profits
• More and better jobs
• Higher productivity
• Improved products (quality, variety)

Note: Fintech = financial technology.
Of course, these areas can interact with each other. For example, a firm’s low capabilities can lead to more difficulty in obtaining financing and vice versa. Similarly, access to new markets can have learning effects that improve capabilities, while better capabilities allow for easier access to new markets. Finally, the business environment can affect all other areas. Hence, interventions across multiple areas can sometimes yield positive synergies. Importantly, when choosing an intervention area and instrument, one must very carefully consider the underlying market failures, instead of focusing on outward firm characteristics, such as size (for example, simply supporting small and medium enterprises).

Although improved productivity should normally improve firm performance and growth, the reverse is not necessarily true, because factors beyond productivity can affect firms. Hence, the two should not be conflated (De Loecker and Syverson 2021). In this context, differentiating the effects on productivity from reducing the costs of inputs is also important. Cost reductions, for example, on regulation compliance, can improve overall firm performance. However, this approach does not necessarily improve productivity in the sense of increased output for given inputs. Such effects depend on what the firm does with the savings.

Finally, a differentiation of productivity effects at the level of the firm versus the whole market is also important. For example, bringing down transport costs may not improve firm productivity but may improve productivity at the sector- or economy-wide level. Similarly, policies that allow for efficient allocation of resources across firms can improve total productivity, because the most productive firms grow faster than the least productive ones. Although the total economy’s inputs are being used more efficiently in this case, the individual productivity of any one firm, or the unweighted average across firms, need not change.9

### 3.1 Firm Capabilities

Firm capabilities cover a wide range of interventions, which should directly increase the productivity within a firm. First, such interventions often address the lack of managerial, organizational, and entrepreneurial10 skills. Further, these interventions can also focus on improving technology adoption and innovation. Finally, such projects could also improve access to capability-enhancing soft infrastructure, such as labor, talent, business development services, and entrepreneurship ecosystems (including incubators and accelerators).

Better firm capabilities are expected to improve productivity, and therefore firm performance, almost by definition. Better technology and innovative processes will lead to a more efficient use of resources. Similarly, successful new products will improve sales, and more skilled employees can do better work. Further, there is evidence that better management acts as a form of technology (for example, by reducing waste), which improves productivity (for example, Bloom et al. 2013).

Information constraints are the most common market failure motivating interventions in this area. Often firms do not know the value of management skills and training, or how poorly the firm is actually run (see, for example, Bloom et al. 2013). In other words, firms tend to be overconfident of their management skills and also “don’t know what they don’t know.” As a result, capabilities can be inefficiently low without intervention. Further, the typical market failures constraining access to finance (discussed later in this note) can also limit investments in firm capabilities.

For innovation and technology adoption11 in particular, further factors exacerbate market failures. Innovation returns are especially uncertain and produce an intangible asset, which typically is not accepted as collateral, thereby increasing financial constraints. Second, there are positive externalities from innovation, which create spillover advantages for other firms. The innovating firm can rarely appropriate the spillovers fully. Because the firm considers only their own benefits from innovation and not those to other firms, externalities lead to too little investment in innovation from a social perspective. Finally, there can be coordination failures as innovation occurs within ecosystems composed of individual actors, networks, underlying infrastructure, and institutions.

Interventions can deploy a great variety of both financial and nonfinancial instruments to address market failures and close capability gaps. Nonfinancial instruments can involve different

---

9. In other words, each firm’s productivity can remain the same, but a country’s productivity can increase if the most productive firms get bigger (receive more of the scarce labor and capital), while the least productive firms become smaller.

10. This lack can also include the entrepreneurial mindset or personal initiative (see Campos et al. 2017).

11. Technology adoption is often seen as a form of innovation, especially in developing countries.
forms of training or consulting, which can be delivered individually or through group-based approaches (Iacovone, Maloney, and McKenzie 2022). There is extensive evidence that nonfinancial instruments can improve management (for training, see, for example, McKenzie 2021) as well as innovation and technology adoption (Cirera et al. 2020) by alleviating information constraints or coordination failures. The financial instruments often consist of (matching) grants or soft loans, which improve productivity by financing and incentivizing the appropriate investments in firm capabilities. A critical assumption for firm capability interventions is that there will be sufficient initial interest from firms to participate and then comply with the programs. Owing to the aforementioned information constraints, this assumption does not always hold.

Typical intermediate outputs can include the number of firms supported with the various instruments; the extent of relevant financing provided; or improvement to the availability and cost of soft infrastructure, such as incubators or business development services. Outputs at the firm level can also include the changed behavior and management practices of firms, improved workforce skills, and increased adoption of technologies or investments in innovation.

3.2 Access to Finance

A lack of access to finance can inhibit the proper investments to increase firm productivity and growth. Hence, interventions in this area aim to ease access to capital and credit, thereby enabling productivity-enhancing investments.

Two broad categories of market failures are typically relevant for access to finance: (a) information asymmetries; and (b) missing markets, especially for nontraditional lending. Information constraints can increase (perceived) risk, which curtails financing. One issue is that the firm will typically have more accurate information than a lender about the true prospects of a project as well as control over the resources and decisions required for its success. When one cannot distinguish between high- and low-risk firms, a vicious cycle of adverse selection can occur whereby loans are priced according to the average risk, thereby pricing out stronger-than-average firms. Information constraints can also cause moral hazard when lenders cannot perfectly monitor the activities of the firm. Although information constraints exist everywhere, a lack of required infrastructure and legal frameworks can make them much more binding in developing countries. For missing markets, the required financial services may not be sufficiently available in a country, leaving viable projects without financing.

Common instruments include providing firms with grants, vouchers, loans, new forms of collateral, guarantees, application support, and so on. Naturally, providing financing in the form of grants or loans will directly increase access to finance. Nevertheless, policies are most effective and sustainable when the focus shifts from subsidizing inputs to creating new (financial) markets. In this sense, providing access to capital markets or early-stage risk financing is one possibility. Another option is to de-risk investments for local investors by providing guarantees or collateral. Technical assistance and application support for firms can also improve the chances of successful financing. Finally, financial or technical assistance on the side of financial institutions can help them better serve firms. In particular, improvements to crowdfunding solutions, fintech (financial technology), digital payments, financial infrastructure, and improved legal protections for creditors can be beneficial. One critical assumption for the success of these interventions is that firms use loans in a productive way for business purposes. To help fulfill this assumption, policies can couple access to finance with improvements in financial literacy and other firm capabilities, which can also improve longer-term access to finance after the project closes.

3.3 Market Access

Market access interventions aim to address the lack of local and external demand by increasing access to domestic and foreign markets. These interventions primarily improve productivity through increasing demand and learning effects.

For firm growth, there must be adequate demand. However, access to richer markets can also increase revenues for
given products, thereby boosting TFPR directly. Furthermore, increased market access can also boost TFPQ, owing to learning effects. For example, Atkin, Khandelwal, and Osman (2017) find strong evidence of learning-by-exporting effects, which improve technical efficiency and quality. In turn, the higher productivity can help firms grow and access further markets, thereby creating a virtuous cycle.

A common market failure preventing market access is the presence of prohibitively high search costs and contract frictions that obstruct buyer-seller relationships. In this sense, sellers need to find appropriate buyers (and vice versa), which can cost time and money (for example, through marketing). These costs can occur naturally, but they can also be caused by policy, for example, in the form of export tariffs. Further, the required financing to invest in market access can be lacking. This is exacerbated by the fact that such investments, including quality certification and compliance with necessary policies, can have uncertain returns. Similarly, when contracts are not perfectly enforceable, otherwise profitable transactions may not occur. This issue is especially important in international trade, because enforcement can be complicated when different legal systems are involved.

Market access interventions often provide firms with market platforms, access to alliances, networks, connections between buyers and suppliers, procurement opportunities, support for marketing and exports, secured orders, and so on. The most common programs to achieve these goals include supplier development programs, programs for global value chain integration, building the national quality infrastructure, facilitating compliance with standards and quality certifications, public procurement programs, and export development programs. A critical assumption for the success of such interventions is often that firms will be able to fulfill received orders and potentially scale up.

Typical intermediate outcomes can include the access to new markets, secured contracts or sales, exports, and learning effects from buyers or competitors.

### 3.4 Business Environment

The business environment is often defined as comprising all factors external to the firm that greatly influence its functioning. Therefore, interventions in this area can be very broad. Improving the business environment can improve productivity within firms by enabling productivity-enhancing investments and incentives to improve firm capabilities. Furthermore, a well-functioning business environment ensures productivity growth between firms in aggregate, by ensuring that more productive firms can enter and grow faster, while less productive firms shrink and exit as they are outcompeted.

Unlike the other areas that arise due to market failures, an inadequate business environment itself creates market failures or fails to correct them properly. The most common challenges in this context are either that firms do not know how to comply with regulations and tax obligations or that compliance is too costly. Further, policies that create or allow for excessive market concentration will inhibit competition, thereby reducing private sector growth and productivity.

In this sense, a better business environment is expected to facilitate improvements in firm capabilities and access to finance and markets. Further, reducing compliance costs will generally benefit firms and may (but not necessarily) translate into productivity growth.

Moreover, ensuring adequate competition can improve productivity both across firms in aggregate as well as within firms (Akcigit et al. 2020). At the aggregate level, competition ensures that resources should flow to the most productive firms, thereby raising the overall productivity level. It will also ensure that the weakest firms exit, and strong firms can enter. At firm level, competition can incentivize firms to improve their capabilities, by investing in better management, technology, and innovation.

The relevant instruments can typically be applied at the industry- or economy-wide level. These include, trade, competition and tax policy, contract enforcement, regulation along the firm’s life cycle, digitally enabled government services, sector-specific regulation, and various forms of
simplified administrative procedures. Such interventions can also provide firms or governments with information, improve monitoring schemes, and facilitate feedback channels.

However, interventions can also simply remove those existing policies that are harmful but may not appear so at first glance. For example, incentives for small firms, such as tax breaks, can discourage firm growth in order to maintain those benefits, while not increasing firm-level productivity (see, for example, Aghion et al. 2017). Further, such policies create distortions, discourage compliance, and enable the survival of less productive firms, thereby reducing aggregate productivity. A critical assumption for the success of business environment interventions is usually a high level of government commitment and ownership to ensure proper implementation.

Intermediate outputs typically include changes in laws; the operationalization of new policies; or improvements to relevant public services, contract enforcement, bureaucratic processes, processing times, the number of required administrative steps, and compliance costs. They can also include various measures of competition.
Productivity Estimation

This section provides practical guidance on estimating the aforementioned productivity measures, while more technical details can be found in the appendix. Note that labor productivity is straightforward to calculate by simply dividing the output measure (for example, value added or revenues) by a measure of labor (for example, the number of workers or hours worked). Hence, this section focuses on TFP in its various forms.

As discussed earlier, TFP captures how efficiently inputs are transformed into outputs. Mathematically, this process is determined by a production function. Hence, to calculate TFPQ and TFPR, one must first estimate the production function. Nevertheless, this procedure faces several econometric challenges.

However, researchers are often simply interested in productivity changes over time (for example, following an intervention). Then, examining productivity growth can reduce the importance of biases in TFP estimates if estimates are not biased differently across time. As a result, the full refinement of TFP estimates to avoid all sources of bias may be a secondary concern for practical purposes. Therefore, possible small biases need not always hinder teams from analyzing productivity. Further, as already discussed, looking at TFP growth also helps minimize issues stemming from comparisons of firms across different sectors or countries. Hence, teams should focus on productivity growth instead of productivity levels whenever possible.  

4.1 TFP Estimation Issues

The main estimation issues involve functional form, aggregation, and endogeneity.

4.1.1 FUNCTIONAL FORM

First, a functional form for the production function must be chosen before estimation, because this cannot be determined by the data. The two most commonly used candidates are the Cobb-
Douglas and the Translog production functions. The estimation will be limited to estimating the parameters of these functions. For practitioners, either can be chosen, and examining results in each case, for robustness, may be worthwhile.

4.1.2 AGGREGATION
Second, a level of aggregation for applying the production function needs to be established, because estimation requires multiple observations. Different industries use different technologies (for example, some will be more labor intensive), and therefore production functions will vary among industries. A choice needs to be made at what level of aggregation firms are similar enough in their technology to group them together in estimating productivity. At the same time, productivity regressions are more accurate when there are more observations, so these groups cannot be too small. A common approach is to use the industrial classification at the 2- or 3-digit level to group firms. This approach does not mean that the sample would be restricted in this way. Rather, the total sample of firms would be divided into a few groups (for example, based on 2- or 3-digit industry classification) simply for TFP estimation, which is done separately for each group.

4.1.3 ENDOGENEITY
Third, the production function estimation needs to correct for various causes of endogeneity, which introduce bias. Fundamentally, endogeneity occurs here when unobserved factors, such as productivity, are correlated with both the inputs and the outputs. This note considers the following five sources of endogeneity:

- **Simultaneity bias.** This bias arises because firm-level productivity directly affects both outputs and inputs. Firms know their productivity and choose inputs accordingly. Hence, through its effect on input choices, productivity also indirectly affects output. The simultaneous direct and indirect effects on output can make isolation of productivity difficult. This is because estimation uses output and input data to determine productivity, but productivity also determines both inputs and outputs. Therefore, one cannot straightforwardly disentangle these two simultaneous effects in the data.

- **Selection bias.** As firms enter and exit, one’s data may be polluted by observing only the survivors, because lower-productivity firms are likelier to exit and therefore become unobservable. This condition is also known as survivorship bias. Again, both the inputs and the outputs one observes are dependent on productivity simply through the types of firms that survive, because one cannot observe the weaker, exiting firms.

- **Measurement error in inputs.** Labor is usually measured in hours worked or number of employees, whereas it would be more appropriate to control for the type of labor, education, experience, and specific skills. For materials, information on discounts or quality differences in inputs may be lacking. For capital, one usually needs to aggregate over various categories such as equipment, machinery, land, and buildings and correct for the appropriate depreciation, which is not straightforward. Further, the two ways of measuring capital—that is, directly via book value or through the perpetual inventory method—are not free from problems and assumptions (Cusolito and Maloney 2018). Hence, if such important elements of inputs that influence outputs are missing from one’s observable data, that absence will also cause endogeneity.

- **Omitted price bias for outputs and inputs with imperfect markets.** Productivity estimates pick up all the unobserved factors determining output, beyond inputs. Therefore, when firm-specific quantities or prices for outputs and inputs are unavailable and these markets are not perfectly competitive, then price-cost margins are also absorbed in the productivity estimate. Then, using industry-level prices to deflate sales would bias TFP. For example, there would be downward bias for efficient firms that were able to pass on efficiency gains into sales prices. Similarly, there would be upward bias for firms that were able to negotiate lower input prices.

- **Multiproduct bias.** When estimating firm-level TFP with multiproduct firms, one (implicitly) assumes the same technology across several types of goods within a firm. This assumption is not realistic for multiproduct firms, which will bias their input coefficients. Furthermore, if there are synergies of the production process across good types,
firm-level TFP can be biased because the production function will be incorrectly specified.

### 4.2 Addressing Endogeneity

Endogeneity is the most complex problem in productivity estimation. Bias from omitted input and output prices can be resolved only by collecting firm-specific quantity or price data. Further, measurement error in inputs remains unresolved in the absence of attempts to improve data collection (Cusolito and Maloney 2018).

For multiproduct bias, there are three broad approaches (Cusolito and Maloney 2018). First, one can focus on only single-product firms and eliminate multiproduct firms from the sample. This approach is problematic because multiproduct firms account for an important share of total output in many sectors (Cusolito and Maloney 2018). Second, one can aggregate product prices to the firm level and estimate TFPR at this level. However, this approach requires the restrictive assumption that markups are common across products within a firm. Finally, one can conduct the analysis at product level, using a mechanism to allocate firm input expenditures to individual products. While this approach is beyond the scope of this note, details can be found in De Loecker et al. (2016).

The most important endogeneity sources, however, are simultaneity and selection. A number of methods have been developed to address them, as will be discussed. Notably, the approaches for TFPQ and TFPR differ fundamentally from that for the TFP index, which uses stronger assumptions in favor of a simpler method.

Historically, two methods have been used traditionally to address endogeneity for TFP. These are the instrumental variable and the fixed effects approaches. However, these two methods have significant shortcomings and have not performed well compared to more recent methods. As a result, they are no longer commonly used. Therefore, this note does not provide further details on them and instead focuses on the control function method. For a more in-depth comparison of different estimation methods and their robustness to specific data problems, see Van Biesebroeck (2007).

Currently, the most popular approach for TFPQ and TFPR is the control function method, which requires certain assumptions and (at least) one additional proxy variable. Recall that simultaneity arises because one cannot disentangle the direct and indirect effects of productivity on output, owing to productivity influencing input choices. Control function approaches use economic theory to introduce an additional equation, which allows one to disentangle these effects. This theory also explicitly considers firm entry and exit, thereby helping with selection bias. Effectively, economic theory shows that productivity can also be a function of the capital stock and an appropriate proxy variable. Hence, by having two relationships to estimate productivity, it becomes possible to disentangle it in the data. Inevitably, for the added equation to be helpful, it must require an additional new variable. Otherwise, there would not truly be additional information in the estimation, which is why the proxy is required. Typically, investment or intermediate inputs (such as materials or energy) are used as proxies.

The control function approach relies on two main sets of assumptions: (a) one about optimal firm behavior in which firms maximize profits and (b) one about timing. The first assumption provides the additional equation relating the proxy, productivity, and capital from theory. The timing assumptions ensure the proxy (and other inputs) do not simultaneously affect current production, which can similarly induce simultaneity. Importantly, capital is assumed to face a time-to-build adjustment cost. Hence, the current level of capital is decided in the last period because investments need time to translate into usable capital stock. In other words, capital investments require time until they can be shipped, installed, and used.

Effectively, one can then first estimate a type of so-called “labor productivity”. As discussed, this does not require a complicated estimation. Then, one can control for the capital effects on productivity through estimating the control function, which relates the capital stock and a proxy, such as investment, to productivity. This second estimation can circumvent simultaneity, because current investment does not affect current production and the current capital stock was predetermined, owing to the time-to-build assumption.

---

15. For example, today’s usable capital stock is determined by investment decisions made yesterday.
16. Current investment translates into capital stock only in the next period, so it affects only future, and not current, production.
The most commonly applied versions of the control function approach are (a) Olley and Pakes (1996) (OP); (b) Levinsohn and Petrin (2003) (LP); (c) Ackerberg, Caves, and Frazer (2015) (ACF); and (d) Wooldridge (2009) (WRDG). Their main differences relate to the chosen proxy variable and assumptions around the timing of inputs. The OP method uses investment as a proxy, while the LP method uses intermediate inputs, such as material inputs or energy. The ACF and WRDG methods can use either proxy, provide a correction to certain issues with the OP and LP approaches, and estimate both productivity equations in one step, making them more efficient. More details are provided in the appendix.

A caveat of the control function methods, as highlighted by the crucial role of timing assumptions, is that they require repeated observations over time per firm (panel data). Owing to the demanding estimation process, a large number of firms is necessary. Finally, data on an appropriate proxy variable are also required. Therefore, these measures are more difficult to apply, especially in studies with smaller sample sizes. When the data do not allow control function methods to be used, the TFP index can be considered.

The TFP index simplifies the issues of getting to the production function largely by assumption instead of estimation, which is both its advantage and disadvantage. To this end, one must accept the assumptions of perfect competition in the input markets\(^{17}\) and constant returns to scale (CRS). Then, the input elasticities (that control function methods try to estimate) are theoretically equal to the shares of revenue paid to each input (for example, the wage bill, rental costs, and material costs). CRS also implies that the sum of all input factor elasticities is equal to one. Hence, capital’s share is normally constructed as a residual (that is, 1 minus the sum of all other shares), so data on the cost of capital are not needed. The TFP index is then calculated by dividing output (or value added) by the sum of inputs weighted (multiplied) by the expenditure shares. The TFP index’s advantage is that panel data, a proxy variable, and a large sample are not strictly necessary. Nevertheless, panel data and a larger sample would also improve the TFP index estimate. Further, as discussed in section 2.2., the required TFP index assumptions are very strong.

Wherever possible, estimating TFPR or TFPQ by using a control function method is preferable to calculating the TFP index. However, TFP estimation should be avoided if the underlying necessary data are too inaccurate. Further, there is no golden rule as to which proxy to use in which application. In fact, carrying out robustness with multiple proxy variables (variable intermediates, investment, or both) is the preferred strategy.

### 4.3 Guidance on Data Collection of Underlying Variables

This section provides guidance on data collection for important variables on productivity analysis.

As discussed earlier, labor productivity requires only output (for example, physical output, revenues, sales, or value added) and labor data (for example, number of employees or hours worked). All TFP measures additionally require capital stock data (for example, firm assets). Capital should ideally be at replacement rather than book value. TFPR and TFPR require data on a proxy, which is typically investment, material, or electricity.\(^{18}\) Finally, TFPQ ideally requires output price data if physical output data are not available, unless one is willing to make strong assumptions on the demand function (as, for example, in Hsieh and Klenow 2009).

Some of the essential variables to be included, either as a means to calculate productivity or as correlates for analysis, are employment (if possible, disaggregated, for example, into production and nonproduction employees, by skill level, or both), salaries, assets (vehicles, machineries, and equipment), expenditures for intermediate inputs (if possible, disaggregated between domestic and imported), investment, value added, revenue, and firm age. Some of the good-to-have variables to be included are exports, imports, hours worked, energy, other expenditures (for example, marketing and packaging), prices of inputs and outputs, and products information.

Hours worked can be a more precise measure of labor than the number of employees, owing to possible differences in the hours worked by each employee. Hence, when one uses employment rather than hours, construction of full-time equivalent (FTE) measures of the number of employees is important. When data on hours are available, FTE can be

\(^{17}\) In this sense, there should also be no input factor adjustment costs.

\(^{18}\) Ideally, material should be included for the TFP index as well when it is a relevant input and value added is not used as a measure of output.
calculated by dividing the total hours by the full-time workload (for example, 40 hours). However, data on hours are often unavailable, highlighting the importance of obtaining data on the number of full-time and part-time employees. In practice, the data will often not distinguish between part-time employees with different hours. In that case, an appropriate approximation for the FTE is simply to divide the number of part-time employees by 2 and add that to the number of full-time employees. This approach assumes each part-time employee works exactly half the time of a full-time employee. When one works with yearly data, a similar adjustment for seasonal employees can be constructed. For example, a full-time seasonal employee working only three months a year is equal to 0.25 FTE employee for that year.

The main sources for productivity data are (a) administrative data and (b) survey data. Administrative data often come from tax or customs offices and should contain (all or parts of) the firms’ balance sheets. In practice, establishing connections with the statistical agencies and other government counterparts early on is important, because administrative data can be available even in some of the poorest countries but can take significant time to acquire. There are also private providers of similar data, such as the Orbis or Amadeus databases. Balance sheet data are straightforward to use for TFPR, because data for revenue, employees, capital, wage bill, investment, and cost of intermediate inputs are readily available. Such data also allow the calculation of value added if desired. Note that capital here corresponds to the book value of the firms’ fixed assets. However, one rarely finds data on quantities or prices in balance sheets, making TFPQ difficult to calculate. A further disadvantage of administrative or balance sheet data is that they will typically include only formal firms. Many informal firms may not have proper balance sheets in the first place. Similarly, not all (formal) firms are required to file taxes or to make their balance sheets public or otherwise be included in administrative data. As a result, smaller firms will typically be missing. Finally, some data may be misreported, for example, owing to tax avoidance.

Surveys can also be used to obtain the necessary variables for productivity. An advantage is that one can obtain detailed information and potentially cover a broad range of firms. Nevertheless, surveys can also face challenges such as misreporting or nonresponse bias. Furthermore, survey sampling design must be careful to ensure representativeness. This approach requires an appropriate sampling frame, such as an enterprise census, which may not always be available. In practice, a partnership with local statistical agencies on survey data collection is useful. Note also that TFPQ and TFPR require panel data, which means that surveys must ask about multiple years or quarters or be implemented multiple times among the same firms. Whenever possible, survey data should also be compared (through distributions by size or other variables) or matched to administrative data. This method serves to adapt or contextualize survey results and can allow for better productivity measures thanks to the expanded data set. To calculate different productivity measures, one can use the questionnaire in table 2 as a basis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Question</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output: Revenue (for TFPR)</td>
<td>What were this establishment’s total annual sales for all products and services [in the last fiscal year]?</td>
<td>Often, surveys ask about the last completed fiscal year, but other time frames are possible, for example, for quarterly surveys or for cases when earlier data are also needed.</td>
</tr>
<tr>
<td>Output: General</td>
<td>Describe the top three products sold (most important, 2nd most important, 3rd most important) [in year X].</td>
<td>This variable is not directly used in the estimation equation but can be helpful when working with multiproduct firms.</td>
</tr>
<tr>
<td>Output: Quantity units, such as kilograms and liters (for TFPQ)</td>
<td>Specify the units of each of the top three products sold [in year X].</td>
<td>The units are required to allow for quantity comparisons or aggregation.</td>
</tr>
<tr>
<td></td>
<td>Question</td>
<td>Additional Information</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Output: Quantity (for TFPQ)</td>
<td>What was the quantity of each of the top three products sold [in year X]?</td>
<td>For multiproduct firms, the top three products tend to be sufficient.</td>
</tr>
<tr>
<td>Output: Revenue by product (for TFPR)</td>
<td>What was the share of total sales for each of the top one, top two, and top three products [in year X]?</td>
<td>This approach will allow researchers to back out revenue for each product, when dealing with multiproduct firms.</td>
</tr>
<tr>
<td>Materials and intermediate goods: General</td>
<td>Describe the top two raw materials and intermediate goods [in year X].</td>
<td>In practice, the top two are often sufficient. More than two can be asked. The questionnaire should at least ask about the top one.</td>
</tr>
<tr>
<td>Materials and intermediate goods: Quantity units</td>
<td>Specify the units of the top two raw materials and intermediate goods [in year X].</td>
<td></td>
</tr>
<tr>
<td>Materials and intermediate goods: Quantities</td>
<td>What was the quantity of the top two raw materials and intermediate goods [in year X]?</td>
<td></td>
</tr>
<tr>
<td>Materials and intermediate goods: Cost</td>
<td>What was the annual cost of raw materials and intermediate goods used in production [in year X]?</td>
<td></td>
</tr>
<tr>
<td>Materials and intermediate goods: Cost distribution</td>
<td>What was the share of annual costs for each of the top one and top two raw materials and intermediate goods [in year X]?</td>
<td></td>
</tr>
<tr>
<td>Materials and intermediate goods: Optional</td>
<td>For the year [X], please provide the following information: (a) Total annual cost of electricity (b) Total cost of production or service operations (including costs of marketing, distribution, and similar costs)</td>
<td>This question can provide useful variables on materials or intermediates, but it is not absolutely necessary.</td>
</tr>
<tr>
<td>Labor: Number of full-time employees</td>
<td>[At the end of fiscal year X], how many (permanent), full-time individuals worked in this establishment? Please include all employees and managers.</td>
<td></td>
</tr>
<tr>
<td>Labor: Number of part-time employees</td>
<td>[At the end of fiscal year X], how many (permanent) part-time employees—employees who work less than a full shift per day—were employed by this establishment?</td>
<td>If absolutely necessary, this question can be excluded in favor of asking about only full-time employees.</td>
</tr>
<tr>
<td>Labor: Number of full-time seasonal employees (optional)</td>
<td>How many full-time temporary or seasonal employees were employed at the end of the year [X]?</td>
<td>This question can improve the precision of the FTE number of employees, but it can be excluded if necessary. It should be accompanied by questions on the average length of employment.</td>
</tr>
<tr>
<td>Labor: Number of part-time seasonal employees (optional)</td>
<td>How many part-time temporary or seasonal employees were employed at the end of the year [X]?</td>
<td>This question can improve the precision of the FTE number of employees, but it can be excluded if necessary. It should be accompanied by questions on the average length of employment.</td>
</tr>
</tbody>
</table>
| Labor: Length of employment of seasonal employees (optional) | What was the average length of employment of seasonal or temporary employees at the end of the year [X]?
This question is required only when asking about seasonal employees. |
|---|---|
| Labor: Wages | For the year [X], what was the total annual cost of labor including wages, salaries, bonuses, and social security payments?
This question can improve the labor measure by differentiating the type of labor, but it is not absolutely necessary. Additionally, a question about wages for each category can be asked. |
| Labor: Type of employees (optional) | At the end of the year [X], how many individuals were working in this establishment according to their main occupation? [Insert sub-questions on occupations. For example, production and nonproduction]
This question can improve the labor measure by differentiating the type of labor, but it is not absolutely necessary. Additionally, a question about wages for each category can be asked. |
| Labor: Skills (optional) | What is the percentage of full-time employees who completed the following degrees as the maximum level of schooling [at the end of year X]? [Insert sub-questions for relevant levels of schooling]
This question can improve the labor measure by gauging the skill level, but it is not absolutely necessary. Additionally, a question about wages for each category can be asked. |
| Capital | What was the value of total fixed assets for this establishment in year [X]?
Alternatively (for current capital):
Hypothetically, if this establishment were to purchase all the machinery, vehicles, equipment, and buildings it uses now, in their current condition and regardless of whether the establishment owns them, how much would they cost? |
| Investment | What was the total value of investment, including equipment, machines, software, and buildings, in year [X]?
This question is not required when investment will not be used as a proxy in TFP estimation. |

Note: FTE = full-time equivalent; TFP = total factor productivity; TFPQ = quantity-based total factor productivity; TFPR = revenue-based total factor productivity.
Attribution of Changes in Productivity to Project Interventions

5.1 The Issue of Attribution

This section discusses the issue of attributing effects to projects, for example, in the context of project evaluation of productivity effects. Observing an improvement of productivity among firms supported by a project does not mean that this effect can be attributed to that project. A simple before-and-after comparison among benefiting firms could be misleading, because differences can be due to multiple other reasons beyond the project. Hence, to measure the effect on productivity that is attributable, one needs a counterfactual. An attributable effect is also known as a *causal* effect.

In practice, the changes in productivity need to be compared between the recipients of project support and a comparison group representing what would have happened to the recipients without the project. This comparison group needs to be as similar as possible to the recipients’ group. Following the literature, this note refers to participating firms as “treated” firms, with the project constituting the “treatment.”

A main difficulty in constructing the counterfactual is that firms participating in a project can be systematically different from those not participating. This is due to self-selection; that is, firms voluntarily choose to participate or not. For example, firms that are already more productive may be more likely to participate and remain with a program that will increase productivity. Hence, one cannot simply compare participating firms to nonparticipating firms because the participating ones will appear more productive regardless of the program. In this case, a simple comparison would overestimate the effects of the program. Similarly, in the opposite case, a simple comparison would underestimate the effects.

The most rigorous way to build an empirical counterfactual is through randomized control trials (RCTs). RCTs can be considered a type of experiment that randomly assigns program participation. When RCTs are not possible, researchers must look for other sources creating random assignment or assignment that is as good as random (quasi-random). Three popular
approaches are (a) the difference-in-difference (DiD) method; (b) regression discontinuity design (RDD); and (c) matching, which includes the synthetic control method. Crucially, RCT work should begin before the project starts (ex ante), while DiD, RDD, and matching analysis can start after a project has concluded (ex post).

Alternatively, the effects of a policy change or an intervention can be simulated on the basis of economic theory, in a partial or general equilibrium model (for example, in a computable general equilibrium model). This approach can provide estimates of productivity changes at the sector- or economy-wide level but only via the channels that are explicitly included in the model. For example, in the case of computable general equilibrium (CGE) reforms, the productivity effects normally would come from a reallocation of factors, rather than within-firm productivity growth, because the latter usually are not modeled. Although simulations can use data to calibrate the model’s parameters, these are typically only a few general data points. Such data can come from pre-intervention data, the literature, or data from other contexts, and they are not based on data capturing changes from an intervention. Because simulations rely mainly on theory and not on a comparison of data before or after an intervention for different groups, they can be implemented before (ex ante) an intervention.

Table 3 provides an overview of the three classes of attribution approaches: experimental, quasi-experimental, and simulations. Owing to the vast theoretical modeling possibilities for simulations, the rest of this section will focus on empirical experimental and nonexperimental approaches.

Table 3: Approaches to Attribute Productivity Effects to Interventions

<table>
<thead>
<tr>
<th>Approach</th>
<th>Description</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
</table>
| Experimental  | RCTs are the main form of experiment for productivity interventions. Researchers randomly assign firms into treated (affected) and nontreated groups. Owing to controlled random assignment, RCT design must begin before the intervention and data collection continue afterward, thereby requiring a mixture of ex post and ex ante approaches. | • Has clear attribution of effects, with the highest standard  
• 2. Allows researcher the most control over whether the required assumptions will hold | • Has high cost in terms of data requirements  
• 2. Must be considered before beginning project implementation |
| Quasi-experimental | Quasi-experimental approaches attempt to mimic experiments, such as RCTs, when these are not possible. The main methods for productivity interventions are DiD, RDD, and matching. These approaches require data before and after an intervention, but can be implemented ex post, because the researcher need not be involved during project design. | • Is less costly than RCTs in terms of data requirements  
• Can estimate effects for large representative samples  
• 3. Can be implemented after a project closes if data are available | • Has less strong attribution compared to that of RCTs  
• Is not always applicable, because the fulfillment of necessary assumptions depends on circumstance |
| Simulation    | Expected productivity effects can be simulated on the basis of theoretical models. Because data from the intervention are not required, these approaches can be conducted ex ante. | • Can estimate the effect at any level of aggregation the model allows  
• Typically, requires little to no new data collection  
• 3. Can be implemented before a project starts | • Does not directly relate attribution to observed changes after the project  
• Results driven by model parameters and assumptions  
• 3. Has many equally valid modeling approaches, making modeling choices difficult |

Note: DiD = difference-in-difference; RCT = randomized control trial; RDD = regression discontinuity design.
5.1.1 RANDOMIZED CONTROL TRIALS

RCTs solve the self-selection problem by randomly assigning participation in the project, giving the researcher direct control over the control group. Hence, RCT design must begin before the intervention starts implementation (ex ante). If one uses a large enough random sample, unobserved differences across firms should not systematically matter for participation in the project, because such differences would be random. As a result, the control group is truly comparable to the treated group in the absence of treatment.

Although RCTs directly solve the issue of self-selection regarding firms assigned to a project, issues can still arise owing to noncompliance. For example, some firms may be randomly assigned to receive training sessions, but then do not attend them. If, for example, those less likely to attend are less productive, there will be bias. Three solutions to this issue include measuring (a) the intent-to-treat (ITT) effect, (b) the local average treatment effect (LATE), or (c) partial identification.

The ITT effect basically redefines the objective of the analysis when noncompliance is an issue. Comparing firms that were assigned to participate in the program (regardless of whether they actually did) to those that were not assigned still provides the causal effects of offering the program. Hence, one cannot estimate the effect of treatment itself, but can still capture the effects of intending the treatment, which is also useful for policy. LATE is the ITT measure divided by the fraction of compliers. This approach gives the causal effect of a treatment for the complier subpopulation, under certain assumptions. Finally, partial identification attempts to estimate the possible range of values of the causal effect from a perfect RCT (see, for example, Siddique 2014).

5.1.2 DIFFERENCE-IN-DIFFERENCE (DID)

With the DID method, the researcher defines a control group (typically after the project has ended) that is expected to have evolved similarly to how the treatment group would have evolved in the absence of treatment. This control group can have different initial levels of the outcome (that is, productivity) than the treatment group. However, one must assume that changes in the outcome across time are the same in the two groups. Because one does not have an RCT, this assumption is not within one’s control and cannot be guaranteed. Nevertheless, with available data on an appropriate control group, DID analysis can begin after a project closes (ex post).

To provide supporting evidence for the main assumption, one should examine the past evolution of the outcome (that is, productivity) between the groups. These trends should be effectively parallel, meaning the rates of change over time are the same. Then, if the assumption of parallel trends was true in the past (before treatment), one may feel more confident that the assumption holds in the post-treatment period.

If so, one simply compares the difference between the two groups after the treatment with their difference before treatment, that is, by taking a difference-in-difference, which gives the DID approach its name.

For example, suppose the difference in productivity after treatment was 10 percent and before treatment it was 2 percent. The assumption is that the counterfactual difference in the absence of treatment would have continued to be 2 percent, because the changes across groups are assumed to be the same. Then, a change of 8 percent can be attributed to the project. In practice, these differences would be calculated in a simple regression framework.

A case when such an approach may apply could be a project implemented among only certain firms. For example, consider a project available to only large firms in one region of a country. A natural candidate for a control group could then potentially be large firms in another region of the country. Similarly, if a policy targets firms in one sector, then firms in unaffected sectors of the same country may be an appropriate control group.

5.1.3 REGRESSION DISCONTINUITY DESIGN (RDD)

RDD exploits the fact that some rules are quite arbitrary and therefore provide good quasi-experiments when one compares firms that are barely affected by the rule with firms
that are barely not affected. Because the research design is given by these arbitrary circumstances and not controlled by the researcher, this method can start being used after a project closes (ex post). The rules and corresponding arbitrary cutoffs can come from multiple sources, such as eligibility requirements for a program, geographical borders, or even time.

For example, consider an intervention that administers a test to firms, with a threshold number of points for eligibility to participate. Firms that are slightly above the threshold are likely very similar to those just below the threshold. Hence, the firms just below the threshold are a good control group for firms just above. Then the eventual productivity differences between these two groups can be attributed to the project. These differences can be formally calculated in a regression framework. A similar argument applies if the rule affects only the probability of participating in the project.

Another common example would be a project that is offered in only one state. Then firms close to the border in a nonparticipating state may be similar enough to firms just across the border in the participating state, thereby yielding an appropriate control group. Such a case is known as a spatial RDD. Similarly, one can also use regression kink design (RKD). This is a variant of RDD, where instead of a jump in the likelihood of treatment owing to the rule, one would expect a kink (a jump in the rate of change instead of the level).

Nevertheless, firms’ inability to manipulate their position with respect to the rule is important; that is, the rule must be truly arbitrary, thereby providing truly random differences between treatment and control. For example, a project may target only small firms below 20 employees. In principle, one could compare firms slightly above and below this employee threshold. However, one again has a selection problem if firms purposefully choose to remain, or misreport that they are, just below the threshold, so that they can benefit from the program or other incentives (such as tax breaks). Such a case is known as bunching.

Several tests can help ascertain the validity of the RDD approach. However, even when the approach is valid, one caveat is that the researcher is limited to firms close to the threshold. Hence, this method requires large data sets. Further, the found effects may differ for firms further away from the threshold.

5.1.4 MATCHING AND SYNTHETIC CONTROL

Matching methods use statistical techniques to construct an artificial control group by matching each treated unit (for example, firm or sector) with a nontreated unit of similar characteristics. Given appropriate data, this method can begin implementation after a project closes (ex post). Two popular versions are the propensity score matching (PSM) method and the synthetic control method.

Using data on observable characteristics from before a project began, a researcher can use matching methods to identify nonaffected units (for example, project nonparticipants) most similar to affected ones (for example, the participants). When this matching is done on the basis of a propensity score estimating a probability of treatment (for example, participation in a project), it corresponds to PSM.

The synthetic control method is commonly applied to estimate the treatment effects on aggregate outcomes experienced by a single unit (that is, country, state, or sector) through simulation. Therefore, it can be especially useful for broad interventions, such as those for the business environment. The method constructs the control group as a weighted combination of untreated units to simulate what would have happened to the treatment group if it had not received the treatment. As a result, researchers do not need to find any single unit in the untreated group, similar to the treated group. Instead, they can create their own as a combination of multiple untreated units.

For example, consider an intervention in one country (or region, or sector, and so on), denoted A. Synthetic control would construct a control group as a weighted average of other unaffected countries, denoted B, C, and D, based on observable data. This weighted combination of unaffected countries should yield a synthetic (fake) country that perfectly mimics the treated country before treatment. For example, the synthetic control group could consist of 20 percent country
B, 50 percent country C, and 30 percent country D. The evolution of the outcome variable (for example, productivity) for the resulting synthetic control group (for example, synthetic country) is an estimate of the counterfactual, that is, what would have been observed for the affected country in the absence of the treatment. The difference between the treatment and synthetic control groups corresponds to the effect.

An important caveat is that matching methods rely on the assumption that there are no systematic differences in unobserved characteristics between the treatment units and the matched control units. This assumption is a very strong one, which arises because matching uses observable data to determine the control group. However, matching can also be combined with the DiD method to also control for any unobserved, time-invariant characteristics between the two groups. Further, matching requires extensive panel data sets that include characteristics of treatment and control groups before and after the treatment.

5.2 Impact Measurement Challenges

Even when an appropriate method to attribute effects is chosen, further measurement challenges may arise. The most common ones in the context of operations, include the following:

- **Timing.** The channels of effects from intervention to productivity take time to materialize. Projects need to decide how long they will wait to measure effects after project completion. Waiting too long has the downside of potential bias from other events affecting firms that may occur in the meantime.

- **Sample size.** Researchers can estimate the minimum sample size needed based on assumptions about the magnitude of expected effects. One must pay attention to the sample size in order to ensure that one can find the effect of the intervention in case there is one. If the sample size is not sufficient, then the estimation can find zero effect when the true effect is larger. A larger sample size will always improve measurement but can be more costly and time consuming to collect.

- **Number of data periods.** The more data periods one uses, the better the estimation of effects, but the more costly the data collection in most cases. Also, some productivity measures, such as TFPR and TFPQ, require at least two time periods.

- **Scaling.** A generalization of effects from the firm level up to the sector- or economy-wide level is not always straightforward. As discussed earlier, interventions can have economy-wide effects, without productivity effects at the individual firm level. Further, because firms have different sizes, the simple average among all firms would not necessarily correspond to the average productivity at a more aggregated level. One solution is to use a weighted sum, that is, summing the productivity of each firm multiplied by a firm-specific weight corresponding to its importance. Weights can be based on employment, output, or value added.

- **Statistical efficiency.** Normally, productivity is measured through regressions and then the effect is also measured through regressions. If these regressions are run sequentially, some efficiency (and, potentially, statistical power) is lost. An alternative approach is to measure productivity and effect simultaneously in the same regression, which can provide lower standard errors and more precise estimates.

5.3 Examples of How Others Have Measured Effects on Productivity

This section briefly presents two relevant examples of how researchers have measured the effect of interventions on productivity in contexts that are relevant for operations. The section focuses on explaining the project intervention, the counterfactual, the measurement of productivity, the specific data collected, and the effects found on productivity.
5.3.1 PROVIDING CONSULTING AND SECURING ORDERS IN THE ARAB REPUBLIC OF EGYPT

PROJECT INTERVENTION AND COUNTERFACTUAL

In a study of Egyptian rug manufacturers, Atkin, Khandelwal, and Osman (2017) implemented a project to causally identify the effect of exporting on firm profits and productivity. Treatment firms (or project beneficiaries) were a randomly selected sample of small rug producers who were offered the opportunity to produce rugs for the export market through a local intermediary. The intermediary also provided treated firms with inputs such as thread and reed and explained the technical aspects of the rug order. Randomly selected control firms did not receive the export opportunity. This intervention falls under the category of market access.

MEASURE OF PRODUCTIVITY

The authors use two measures of productivity: output per labor hour and TFP, both adjusted and unadjusted for changes in product specifications or quality.

Unadjusted output per labor hour comes from firms' responses to the question: "How long does it take you to make 1 [square meter]?" Their unadjusted TFP measure additionally accounts for capital, which in this case is simply the number of active looms. Unadjusted TFP is then equal to the residual from an estimated Cobb-Douglas production function that includes both labor and capital. It corresponds to TFPQ because output is measured in quantities (square meters). Issues from multiproduct firms do not arise here, because all firms produce only rugs. In their adjusted measure, the authors additionally include control variables on rug quality specifications. This approach corrects TFPQ to reflect quality differences.

To estimate TFP, the authors use the control function approach with the one-step estimator developed by Wooldridge (2009). They use labor hours lagged by one period as an instrument for current labor and use warp thread count as the proxy variable. Further, they estimate the production function using only control firms, thereby avoiding the need for assumptions on how treatment changes the TFP process. However, this approach assumes that the parameters of the production function are the same for treatment and control firms. The authors argue that this assumption is reasonable because all firms produce a narrowly defined product using the same technology in all periods. Finally, using the literature and their data, the authors assume that the costs incurred to find foreign buyers should not be included in the production function here.

DATA COLLECTED

The authors collected data through surveys given every four months on firm production, rug quality, and household and demographic characteristics. The firm data included profits, revenues, expenses, output quantity and prices, input quantity and prices, total labor hours worked, and the specifications of the rugs produced that month. They measured product quality along 11 dimensions from a skilled quality assessor to capture specifications and difficult-to-codify attributes that depend on the technical skill of the firms. Additionally, they collected data on information flows among buyers, the intermediary, and producers that include transcripts of buyer feedback and the content of discussions between the intermediary and producers.

EFFECTS ON PRODUCTIVITY

Treatment firms reported profits 16–20 percent higher than those of control firms. Unadjusted output per labor hour and unadjusted TFP decreased, but there was also a large improvement in quality for treatment firms compared to control firms. Specifically, if one looks at the intent-to-treat effects, unadjusted output per hour is 24 percent lower and the unadjusted TFP is 28 percent lower in treatment firms relative to control firms. The treatment-on-the-treated effects are even larger for both measures of productivity. This result suggests that treatment firms switched to producing higher-priced, higher-quality products that take longer to produce. The authors also found evidence of a learning-by-exporting effect that improved technical efficiency. Finally, after controlling for rug quality specifications, treatment firms show higher output per labor hour and TFP, as well as higher-quality rugs when compared to control firms. The decrease in unadjusted TFP, but increase when adjusting for quality, highlights the need to carefully consider the differences between TFPQ and TFPR in cases where quality is especially important, as previously discussed.
5.3.2 PROVIDING GRANTS TO MICROENTERPRISES IN GHANA AND SRI-LANKA

PROJECT INTERVENTION AND COUNTERFACTUAL
Laurin, Koelle, and Quinn (2019) estimate a production function using the data from two previous randomized controlled trials in Sri Lanka (De Mel, McKenzie, and Woodruff 2008) and Ghana (Fafchamps et al. 2014). In both studies, the authors built a counterfactual by randomly assigning microenterprises to receive either a cash transfer or an in-kind transfer, which were offered in different waves. The sample in Sri Lanka included microenterprises in retail, sales, and manufacturing sectors, and in Ghana, it comprised microenterprises that did not have any paid employees or motorized vehicles. This intervention falls under the category of access to finance.

MEASURE OF PRODUCTIVITY
The authors use the standard definition of TFP as the residual from a Cobb-Douglas production function. They consider gross output instead of value added to avoid netting out the effect of intermediate inputs. Output is specified in terms of revenue, meaning they estimate TFPR. As inputs, they consider capital, labor, and materials. Revenue and capital are measured in real (inflation-adjusted) national currency, while labor is measured in weekly hours worked.

They use mainly three methods to estimate TFP: (a) a linear panel estimator, following Blundell and Bond (1998); (b) a control function estimator, following Wooldridge (2009), using materials as the proxy; and (c) an estimator exploiting the firm’s first-order condition from its optimization problem, following Gandhi, Navarro, and Rivers (2020). Nevertheless, all three measures give similar estimates.

A novelty in this case is demonstrating that with high-quality panel data, these sophisticated methods can be applied to informal microenterprises. Most of the previous literature has applied such productivity estimates to only large firms with detailed, sophisticated accounting practices and financial records. Nevertheless, this approach requires a few technical adjustments. First, to ensure enough statistical power, the authors pool data from treatment and control firms. To account for the effects of treatment on the variables in the production function, they partial out treatment and time effects from output and inputs before they enter the production function estimation. Second, to reduce the influence of outliers that are due to measurement error, they winsorize each input at the 1 percent and 99 percent levels. Third, they restrict the sample to firms with strictly positive amounts of all inputs.

DATA COLLECTED
The authors use the sample and survey data from the studies in Sri Lanka and Ghana. Both surveys include data on capital, labor, and intermediate goods purchases.

EFFECTS ON PRODUCTIVITY
Capital grants have significant heterogeneous treatment effects on TFP along the distribution. TFP of the median firm increases 5–6 percent and 7–9 percent for the 80th percentile for firms in Sri Lanka and in Ghana, respectively. Between 19 percent and 29 percent of the increase in revenue resulting from capital grants can be attributed directly to an increase in productivity in Sri Lanka, and between 21 percent and 35 percent in Ghana. One possible mechanism through which capital grants increase TFP is that treated firms are more likely than control firms to invest in assets that have a higher technology component. This action results in more efficient means of production and helps firms reach different market segments. The results were sustained over time (for example, over 6 years in Sri Lanka).
Key Dos and Don’ts

This final section draws on the examples and guidance presented earlier to outline a series of key principles for project task team leaders to rely upon when making decisions about how to capture the potential effect of project interventions on productivity.

**DO**
- carefully consider differences among labor productivity, TFPQ, and TFPR;
- recognize the limitations of any productivity measure and the implications for interpretation;
- try different productivity measures and estimation techniques;
- continuously collect as much high-quality data as possible, including on output and input quantities;
- try to establish connections with the statistical agency early on, because administrative data can be available even in some of the poorest countries;
- try to partner with the statistical agency on survey data collection;
- if non-administrative data are used, compare their (size) distribution with administrative data to adapt or contextualize results;
- try to match survey data with administrative data to help get better measures of productivity;
- be upfront about the drawbacks of available data and methods;
- use the most appropriate method for the available data, even if it requires a less sophisticated approach (for example, use revenues per employee instead of value added per employee, if data for the value added calculation are too inaccurate);
- consider how to attribute effects to projects (for example, an RCT) early on, ideally in the design stage;
- consider synergies between interventions, such as combining access to finance or markets with support to firm capabilities; and
- focus on addressing underlying market failures instead of outward firm characteristics (for example, size).

**DON’T**
- misinterpret labor productivity as overall firm productivity (when labor is not the only input);
- try to do more with the data than is warranted, for example by using sophisticated methods (such as econometric TFP) instead of simpler approaches if the underlying data are too inaccurate;
- attribute productivity changes to a project without addressing endogeneity (for example, through an RCT); or
- conflate effects on firm growth with effects on productivity, even though the two are related.
Appendix: Technical Details

This appendix will provide some details on total factor productivity (TFP) estimation, in a semitechnical matter, including implementation in Stata.

7.1 The General Estimation Approach for TFP

As discussed in the main text, TFP requires estimating the production function. Then, productivity is extracted as the residual of output that is not explained by the inputs. A basic equation would be\(^\text{19}\)

\[
\text{output} = F(\text{inputs}) + \text{productivity}
\]  

(1)

where \(F(\text{inputs})\) is the production function dependent on inputs. Hence, having input and output data, as well as a proper estimation of the function \(F\), allows one to measure productivity. Here, one can see why a firm that produces more outputs given the same inputs would be considered more productive.

After assuming a functional form for the production function, one can attempt to estimate its parameters, that is, the elasticities of the input factors. To this end, simultaneity and selection bias are the main challenges that the various estimation methods attempt to address. For a better understanding of how estimation deals with these problems, a slightly deeper dive into these sources of endogeneity is useful.

7.1.1 Simultaneity Bias

One sees from equation (1) that one must estimate the unknown (or not fully known) production function \(F(\text{inputs})\) based on observed data for inputs and outputs. However, productivity is

---

19. Note that this equation can represent a Cobb-Douglas (that is, a multiplicative) function, if all variables are in logs.
also unknown, unobservable to researchers, and directly influences output. Further, productivity in a given period is also correlated with the level of used inputs (capital, labor, and others). This is because firms simultaneously know their productivity and choose their inputs according to their productivity. This action leads to a typical omitted bias or simultaneity bias. In other words, one wants to use input and output data to determine productivity, but productivity also determines inputs and output. In practice, a simple estimate of $F(\text{inputs})$ would calculate the following equation based on observable data:

$$\text{estimated output} = F(\text{inputs}) \quad (2)$$

where $\text{estimated output}$ should capture the part of output explained by inputs. Because the inputs data are influenced by productivity, this calculation will not provide $F(\text{inputs})$, but rather

$$\text{estimated output} = F(\text{inputs( productivity)}) \quad (3)$$

where productivity is the portion of output not explained by inputs, that is, not captured by estimated output. However, naively calculating productivity as a residual of equation (1) using equation (3) will result in the following estimate:

$$\text{productivity} = \text{output} - F(\text{inputs( productivity)}) \quad (4)$$

Because the only data are on inputs and output, equation (4) makes clear why bias arises in a simple estimate of productivity. The unknown productivity is on both sides of the equation and cannot be disentangled. Fundamentally, both the true productivity and the true production function $F$ are unknown. Hence, one has two unknowns but only one equation for them, which cannot have a mathematical solution.

### 7.1.2 SELECTION BIAS

The endogeneity problem illustrated in equation (4) can also come from selection (or survivorship) bias, that is, the entry and exit of firms. Again, both the inputs and the output observed are systematically dependent on productivity through the types of firms that survive, with lower productivity firms likelier to exit. The capital coefficient is particularly affected by this bias owing to its higher adjustment cost or time-to-build aspect. For example, firms with a higher capital stock can absorb lower-productivity shocks, because their value of remaining active is higher. This leads to downward bias for the capital coefficient (Cusolito and Maloney 2018).

### 7.2 Control Function Approach Details

The essence of the control function approach is as follows. Recall from equation (4) the issue of two unknowns and one equation. Under certain assumptions on firm behavior, economic theory shows that productivity can also be a function of the capital stock and an appropriate proxy variable. If one uses further timing assumptions (and data for the proxy), this additional function can also be estimated. One then has two equations for two unknowns, thereby helping solve the endogeneity issue. Effectively, the new equation controls for correlation between input levels and the unobserved productivity shock. An alternative interpretation is that control functions help model the simultaneity of TFP and input choices.

As seen, this method broadly relies on a behavioral and a timing assumption. The behavioral assumption is that firms maximize profits, which generates an optimal input demand equation from theory. This equation directly relates the proxy variable to the firm’s productivity and some variables about the state of the firm (typically capital), while considering firm entry and exit. An embedded technical assumption here is that this function is monotonic.\(^{20}\) Hence, economic theory (based on these assumptions) introduces one more productivity equation into the mix, allowing one to disentangle equation (4).

The timing assumption ensures the timing of variables is not an issue. In technical terms, one often assumes an exogenous Markov process\(^ {21}\) for productivity. Importantly, capital is assumed to face a time-to-build adjustment cost, meaning the current available capital is decided by investments in the last period. The level of other variables, including labor, can be determined as usual, after productivity is observed,\(^ {22}\) or at least more flexibly than capital. This approach is quite

---

20. A monotonic function is either entirely nonincreasing or nondecreasing. If the function would alternate between increasing and decreasing and one does not know exactly how, then this equation would not be helpful. However, this assumption is not a strong one in this case.

21. A Markov process is not a strong assumption here. It means that tomorrow’s productivity is influenced by today’s productivity, but not directly by yesterday’s productivity, except to the extent that yesterday’s productivity influenced that of today.

22. Some methods, like the ACF (Ackerberg, Caves, and Frazer 2015) correction, also adjust this assumption.
realistic, because capital has a stronger time lag before becoming productive. On any given day, a firm will know its productivity and can (more easily) determine how many labor hours to use. It will also determine how much to invest. However, capital investments require time before they can be shipped, installed, and used. Therefore, each day, the firm will use the capital (machines, buildings, computers, and so on) that it already has, and its investment decisions will affect only future output. As a result, the additional equation from theory, relating productivity, capital, and a proxy such as investment, can avoid simultaneity. Therefore, it can help disentangle productivity, including from other inputs such as labor, when coupled with a version of equation (3) based on more flexible inputs.

The appendix now turns to the details of the main control function applications: (a) Olley and Pakes (1996) (OP); (b) Levinsohn and Petrin (2003) (LP); (c) Ackerberg, Caves, and Frazer (2015) (ACF); and (d) Wooldridge (2009) (WRDG).

- **OP.** The OP method uses investment as a proxy. However, investment is often not an ideal variable. First, investment is lumpy, meaning that most firms do not invest at all in many years and then make one large investment. In the data, this aspect implies many zero-value observations which are dropped. Similarly, certain types of adjustment costs (so-called nonconvex costs) can imply that investment does not perfectly react to productivity. Therefore, the required relationship between productivity and investment may be weak. Finally, investment can be correlated with other relevant unobserved factors, such as research and development spending or foreign direct investment.

- **LP.** By contrast, the LP method uses intermediate inputs, such as material inputs or energy, as a proxy. The advantage is that such intermediate inputs are more variable. Hence, they are typically used in every time period and tend to have smaller (nonconvex) adjustment costs. Otherwise, the LP method is very similar to the OP method. However, because of their variability, the timing assumptions are less realistic for intermediate inputs than for investment. In the OP method, one uses current investment, which does not affect current output, as discussed. However, current use of material or energy directly affects the current output. These timing considerations partly inspired the ACF method.

- **ACF.** The ACF method provides a correction to either the OP or the LP approach. In technical terms, the labor demand and control function are partially collinear (related) and should not be estimated one after the other. The ACF method estimates both necessary equations in one step, instead of two. Further, it adjusts the theoretical assumptions on the timing of optimal input choices, which is especially useful when using intermediate inputs as a proxy (as in the LP method). Typically, labor is considered somewhat less flexible than before but still more flexible than (and chosen after) the capital stock. The additional ACF assumptions are not considered strong compared to those of the OP and LP methods. Further, compared to the OP or LP method, the ACF method does not require additional data, is often statistically more efficient owing to the one-step estimation (yielding more precise estimates), and can solve their technical collinearity issues. As a result, the ACF method has become very popular.

- **WRDG.** Wooldridge (2009) also proposed a method to solve the technical issues in the OP and LP approaches, as addressed by the ACF method. It also estimates both equations in one step, addressing collinearity and reducing the standard errors of the productivity estimates. Further, the WRDG approach allows the inclusion of so-called dynamic panel instruments. As a result, this method has also become quite popular. However, the advantage of allowing panel instruments is often not very interesting in practice, because such instruments may be of limited benefit and are not always used.

### 7.3 TFP Estimation in Stata

In Stata, to estimate TFPQ (quantity-based TFP) or TFPR (revenue-based TFP) with the OP, LP, ACF, or WRDG method, users can use the "prodest" command. Alternatively, for the ACF method, the command "acfest" can also be used. These
commands estimate the production function parameters. The residual representing TFP can then be obtained, for example, with the "predict" command. Both methods allow for a wide variety of options, such as using either a Cobb-Douglas or Translog production function or using value added instead of revenue or quantities. Further, prodest allows for an additional correction to attrition from firm exits (survivorship bias), by choosing the "attrition" option. More details can be found in Rovigatti and Mollisi (2018).

There are no commands to readily estimate the TFP index in Stata. To do so, the following steps should be applied. First, calculate the wage bill and employment if not already available. If possible, this step should be done separately for production and nonproduction workers, owing to differences in skills and wages. Then, calculate the industry cost shares (as a share of revenue) for labor and (if applicable) for other inputs. For labor, this is done through dividing wages by revenue. Then, calculate the share of capital costs as 1 minus the shares for labor and other inputs, which forces the constant returns-to-scale assumption. If possible, use industry price deflators to calculate “real” physical materials and output amounts, by dividing costs and revenues by the deflators. Finally, calculate the TFP index by dividing (real) output by the sum of each (real) input, multiplied by its expenditure share. When variables are in logs, researchers should take a difference instead of dividing.

24. Although rarely possible in practice, even more detailed breakdowns could be advantageous in theory.
References


References


References


