

Technology Sophistication Across Establishments*

Xavier Cirera
World Bank

Diego Comin
Dartmouth College

Marcio Cruz
IFC

September 8, 2025

Abstract

We study technology sophistication using a novel approach that measures the sophistication of the most advanced (MAX) and the most widely used (MOST) technologies in each of the key business functions within establishments. Using data from over 21,000 establishments across 15 countries, we find that establishments generally underutilize the most sophisticated technologies available within a business function. These MAX-MOST gaps are persistent and strongly associated with productivity both across establishments and countries. At the establishment level, there is substantial variation in both MAX and MOST, with MOST showing a more skewed distribution. MAX and MOST follow different lifecycle patterns in low-income countries and among small establishments, and they exhibit different associations with several establishment characteristics and performance indicators. This evidence underscores the different nature of the technology upgrading processes that drive MAX and MOST.

*We thank David Baqaee, Jan DeLoecker, Mary Hallward-Driemeier, Apoorv Gupta, John Haltiwanger, Maurice Kugler, Kevin Lang, Bill Maloney, Martha Martinez Licetti, Paolo Mauro, Denis Medvedev, Dani Rodrik, Giacomo Ponzetto, Tommaso Porzio, Doug Staiger, Edouard Schaal, Chris Snyder, Eric Verhoogen, and seminar participants for insightful comments. Kyung Min Lee actively participated in earlier drafts of this paper and we are indebted to him. Harneet Singh provided outstanding research assistance. We thank Aman Mahajan, Magda Malec, Mariana Pereira, Santiago Reyes, Antonio Martins Neto, Soumya Agrawal, and Adrian Forest for inputs to sections of the paper. We also thank the industry experts and partner institutions for their support in designing and implementing the survey. An extended list of acknowledgements is provided in the appendix. Financial support from the infoDev Multi-Donor Trust Fund, the Korea World Bank Group Partnership Facility (KWPF), and the Competitive Industries and Innovation Program (CIIP) is gratefully acknowledged. This manuscript supersedes [Cirera et al. \(2020\)](#), [Cirera et al. \(2021\)](#) and [Cirera, Comin and Cruz \(2024a\)](#). The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the World Bank Group or its Board.

1 Introduction

Technology is central to some of the most fundamental economic questions. Yet, our understanding of these issues often depends on indirect and limited measures. A long tradition, dating back to [Ryan and Gross \(1943\)](#) and [Griliches \(1957\)](#), has characterized technology in establishments by the presence of a few (typically one) advanced technologies. This approach faces several limitations. First, the number of technologies is very small compared to those used in an establishment. Additionally, the tasks in which establishments employ these technologies are neither comprehensive nor representative of the business functions that comprise the establishment. Second, measures based on the presence of advanced technologies do not provide information on how establishments without them produce. In particular, we do not know how sophisticated the technologies used are relative to the frontier. This concern is particularly relevant in developing countries where advanced technologies are less widely diffused. Third, traditional measures do not capture how intensively a technology is used, which is crucial to explain income divergence across countries ([Comin and Mestieri, 2018](#)). This omission limits our understanding of whether establishments predominantly use the most sophisticated technologies they have adopted and the importance for productivity of these technologies relative to the most widely used technologies.¹

In this paper, we develop a new approach to directly and comprehensively measure the sophistication of technologies used in establishments. Our first step is to create a two-dimensional grid structure, which we refer to as ‘the grid’. Following the task-based production function approach ([Zeira, 1998](#); [Grossman and Rossi-Hansberg, 2008](#); [Acemoglu and Autor, 2011](#); [Acemoglu and Restrepo, 2018](#)), the horizontal dimension of the grid covers the key tasks an establishment conducts, grouped into broader categories that we call business functions (BF). To this, we add a vertical dimension that represents the range of technologies that can be used to perform the key tasks in each business function. The grid encompasses

¹Since the classic work of [Ryan and Gross \(1943\)](#) and [Griliches \(1957\)](#) on hybrid corn, many have applied this approach to measuring technology in establishments in other sectors. For example, [Davies \(1979\)](#) studies the diffusion of 26 different manufacturing technologies, each typically relevant in only a single narrow sector, [Trajtenberg \(1990\)](#) measures the presence of CAT-scanners in hospitals, [Brynjolfsson and Hitt \(2000\)](#); [Stiroh \(2002\)](#); [Bresnahan, Brynjolfsson and Hitt \(2002\)](#); [Akerman, Gaarder and Mogstad \(2015\)](#) measure the presence of some ICTs such as computers or access to the internet. Other efforts include the Survey of Manufacturing Technology by the Census Bureau discontinued after 1993, which covers 17 specific technologies, including numerically-controlled machines, computer-aided design or engineering technologies, programmable controllers and local area networks (see [Dunne \(1994\)](#)); or the Canadian Survey of Advanced Technologies with 41 and 50 technologies, depending on the round (see for example [Boothby, Dufour and Tang \(2010\)](#)). More recently, the Advanced Business survey, [Acemoglu et al. \(2022\)](#), also administered by the US Census Bureau and that focused on five generic, frontier technologies: AI, robotics, dedicated equipment, specialized software and cloud computing. Unlike the previous studies, the Advanced Business Survey asks for the intensity with which the firm uses these advanced technologies.

63 business functions: seven general business functions (GBF) relevant to all sectors, and 56 sector-specific business functions (SSBF) across 12 sectors (agriculture, livestock, food processing, apparel, leather goods, automotive, pharmaceutical, other manufacturing, wholesale and retail, financial services, land transport services, and health services). In total, the grid spans 305 technologies.

The grid has three properties. First, it is comprehensive both in terms of the business functions and of the technologies considered in each business function. Second, it is relevant for any establishment and country, regardless of its level of development. Third, the technologies in each business function are ranked according to their sophistication, from the simplest to the most complex which represents the world technology frontier.

We implement the grid in the Firm Adoption of Technology (FAT) survey, administered to a cross-section of over 21,000 establishments that constitute representative samples in 15 countries: South Korea, Poland, Croatia, Chile, the Brazilian state of Ceará, Georgia, Vietnam, the Indian states of Uttar Pradesh, Tamil Nadu, Gujarat and Maharashtra, Ghana, Bangladesh, Kenya, Cambodia, Senegal, Ethiopia, and Burkina Faso.² FAT collects three types of information. First, it gathers establishment-level data on sales, inputs, education of the workers and the managers, management practices, etc. Second, it records whether each sector-specific business function is conducted in-house. Third, and most relevant, FAT documents the technologies from the grid used by each establishment in each business function and, of these, which one is the most widely used technology.

Using the information from FAT, we develop two measures of technology sophistication at the business function-establishment level: ‘MOST’ for the most widely used technology, and ‘MAX’ for the most advanced technology available. These two measures capture distinct forces that jointly determine the use of technology at the BF-level. Changes in MAX reflect the adoption of new technologies in the BF. Changes in MOST can be driven by adoption but also by shifts in the intensity with which different available technologies are used in the BF.

We use the MAX and MOST measures to study two broad issues. First, we examine whether establishments underutilize the most sophisticated technology they have adopted within a business function. An ordinal comparison between the MAX and MOST indices reveals that in 63% of BFs where multiple technologies are in use, the most widely used technologies are not the most sophisticated ones available to establishments. These MAX-MOST gaps are not transitory, as their prevalence does not depend on the establishment age or on how long ago an advanced technology (e.g., MAX) was introduced in the function.

²The date of administration of the survey in the different countries ranged from 2019 to 2023. See [Table A.21](#).

Furthermore, MAX and MOST are differently related to the number of technologies an establishment uses in a BF. While MAX exhibits a near one-to-one relationship with the number of technologies adopted, an increase by one unit in the number of technologies is associated with only an increase by 0.25 units in MOST. These observations support the hypothesis that MAX and MOST are driven by distinct technology upgrading processes.

To better understand why establishments may not make extensive use of the sophisticated technologies they have adopted, we analyze the relationship between the frequency of MAX-MOST gaps and establishment-level characteristics that proxy for human and physical capital scarcity, and managerial mistakes. We find that the frequency of MAX-MOST gaps is positively correlated with indicators of these factors, and that their relevance varies by establishment size and between low- and high-income countries.

In addition to shedding light on how establishments use technology at the BF level, MAX and MOST can be used to measure the sophistication of technology at the establishment level. By aggregating these measures across all functions within an establishment, we generate establishment-level indicators of technology sophistication that retain the grid’s comprehensive scope. These indicators address the shortcomings of conventional technology metrics, which typically record only the presence or count of advanced technologies. In contrast, our measures capture the sophistication of the technology used in production even for establishments that have not adopted advanced technologies, while also reflecting the intensity with which sophisticated technologies are utilized.

Unlike MAX-MOST gaps, the establishment-level versions of MAX and MOST are cardinal measures. Therefore, they may be sensitive to the specific cardinalization of the ordinal sophistication rankings used (Bond and Lang, 2019). We conduct two exercises to validate the affine cardinalization used in defining MAX and MOST. First, we demonstrate that the relationship between technology sophistication and various measures of productivity across establishments is well approximated by a linear function.³ Second, we show that our technology sophistication measures exhibit strong correlations with two alternative cardinalizations: one based on the z-score method used by Bloom and Van Reenen (2007) to cardinalize the quality of management practices, and another based on independent estimates of the relative productivity of the technologies in conducting a representative task specific to the BF. These findings indicate that the patterns of technology sophistication across establishments identified with our baseline measures are robust to a wide range of reasonable cardinalizations.

The analysis of establishment-level sophistication reveals key patterns in the distribution of MAX and MOST and their relationship to establishment characteristics and performance.

³This aligns with the standard approach to cardinalize ordinal variables by projecting them onto relevant variables; see Cunha, Heckman, and Schennach (2010) and Bond and Lang (2018) for example.

First, both MAX and MOST vary substantially across establishments, within and across countries and within detailed sectors. MOST exhibits a more skewed distribution than MAX. Within-country dispersion in sophistication rises with per capita income, though more markedly for MOST.

Second, MAX and MOST are positively associated with several establishment-level characteristics, including size, worker human capital, management quality, exporter status, and affiliation with multinationals or multi-establishment firms. We identify distinct lifecycle patterns: MAX increases steadily with establishment age, while MOST flattens beyond ten years in low-income countries and among smaller establishments. These differences are partly driven by varying size effects across income levels.

Finally, we find robust associations between technology sophistication and establishment performance. Both productivity and profitability are positively related to MAX and MOST, with significantly stronger associations for MOST and for agricultural establishments. The associations between establishment performance and technology sophistication do not differ between high- and low-income countries.

Our analysis of technology sophistication contributes to several strands of research. The MAX–MOST gap relates to intra-firm technology diffusion, originally introduced by [Mansfield \(1963\)](#) to describe the gradual adoption of innovations (e.g., diesel locomotives) within firms. A small body of literature has examined intra-firm diffusion for a handful of technologies in a few countries and sectors,⁴ yielding mixed evidence on whether older technologies persist after newer ones are introduced, and whether similar factors drive intra- and inter-firm diffusion.

Our findings show that MAX–MOST gaps are widespread and persistent. Moreover, MAX and MOST reflect distinct technology upgrading processes, driven by different factors with MOST exhibiting a stronger association with productivity across establishments than MAX.

There are clear methodological and conceptual parallels between our effort to measure and study technology sophistication and the seminal work by [Bloom and Van Reenen \(2007\)](#) on management practices. As with technology, there is a long tradition of documenting specific management practices in a limited number of companies. The groundbreaking studies by [Bloom and Van Reenen \(2007\)](#) and [Bloom et al. \(2019\)](#) have greatly extended this scope by measuring the quality of management practices across 18 dimensions related to operations, planning, monitoring, and human resources, covering thousands of firms in many countries.

Similar to FAT, data on management practices is collected via firm surveys. Experts rank

⁴For example, numerically controlled machines in UK metalworking ([Battisti and Stoneman, 2003](#)), and e-commerce in the UK and Switzerland ([Battisti et al., 2007](#); [Hollenstein and Woerter, 2008](#)).

practices based on their quality, and an establishment-level score is constructed to study the drivers of management practices and their association with productivity. We perform a similar analysis for technology sophistication. Beyond the similarities in measurement methods, Bloom, Sadun and Van Reenen (2012) have hypothesized that technology sophistication and management practices are complementary. We explore the complementarity of technology and management finding supporting evidence.

The breadth of the FAT survey enables us to examine the relationship between technology sophistication and establishment performance across key subsamples—agricultural vs. non-agricultural establishments and low- vs. high-income countries. Our results contribute to the literature on the agricultural productivity gap (Caselli, 2005) and inform ongoing debates about the appropriateness of advanced technologies in low-income settings (Basu and Weil, 1998; Acemoglu and Zilibotti, 2001).

The remainder of the paper is organized as follows. Section 2 introduces the FAT survey and describes validation exercises for the sophistication rankings, along with the data collected. Section 3 presents the technology sophistication measures, discusses their properties, and illustrates their relevance using examples from specific establishments and sectors. Section 4 examines the use of technology at the business function level and documents and analyzes the MAX–MOST gaps. Section 5 explores technology sophistication across establishments. Section 6 concludes.

2 The Survey

The FAT survey (henceforth, “the survey”) collects detailed information for nationally representative samples of establishments in agriculture, manufacturing, and services about the technologies each establishment uses to perform key business functions necessary to operate in its respective sector. In the following sub-sections, we describe the survey design and implementation, relegating further details to section A in the Appendix.

2.1 Structure

The survey is composed of five modules. Module A collects information on the general characteristics of the establishment.⁵ Modules B and C cover the technologies used. Mod-

⁵The survey is designed, implemented, and weighted at the establishment level. For multi-establishment firms, the survey targets the establishment randomly selected in the sample. The survey can be downloaded at the following address (https://dcomin.host.dartmouth.edu/files/FAT_Survey_complete.pdf). The implementation manual which includes all instructions for interviewers, training materials and a full description of the technologies in the grid can be downloaded at (https://dcomin.host.dartmouth.edu/files/Implementation_Manual_TAS_29112023.pdf).

ule D focuses on barriers to, and drivers of, technology adoption, while Module E gathers information about the establishment's financial statements and employment.

The survey differentiates between general business functions (Module B), which comprise tasks that all establishments conduct, regardless of the sector where they operate, and sector-specific business functions (Module C), which are potentially relevant only for establishments in a given sector. All establishments in our sample respond to Module B, but only those belonging to the sectors for which we have developed a sector-specific module respond to C. To attain a wide coverage that allows a meaningful study of sector-specific technologies, we develop sector-specific modules for 12 significant sectors in the economy, including agriculture (crops and fruits), livestock, food processing, wearing apparel, leather goods and footwear, automotive, pharmaceutical, other manufacturing, wholesale and retail, financial services, land transport services, and health services.⁶ These sectors have been selected based on their share in aggregate value-added, employment and number of establishments and they cover all three industries (agriculture, manufacturing, and services).

2.2 The Grid

To design Modules B and C, we determined the business functions covered and the list of technologies, from most basic to most sophisticated, that can be used to implement the key tasks in each function. We call the resulting structure "the grid".

To construct the grid, we followed three steps. First, we conducted desk research reviewing the specialized literature. Second, we held meetings with World Bank Group experts on each of the sectors covered. Third, we reached out to external consultants with significant experience (at least 15 years) in a given sector. For example, the external experts in agriculture and livestock were agricultural engineers and researchers from Embrapa-Brazil. For food processing, apparel, automotive, pharmaceuticals, transportation, finance, and retail, as well as for the GBFs, we relied on senior external consultants selected by a large management consulting organization. For health, our team relied on consultants and physicians with practical experience in both developing countries and advanced economies. In total, more than 50 experts participated in the construction of the technology grid. The resulting grid is composed of 7 general and 56 sector-specific business functions and contains a total of 305 technologies (See Section A.1.1 of the appendix for details on the procedures followed to define the grid).

All technologies in the Grid are precisely described so that respondents and enumera-

⁶The granular information that can be obtained with the FAT survey allows us to explore central questions on technology policy in developing countries. One example, itself a product of this paper, is the World Bank policy report "Bridging the Technological Divide" (Cirera, Comin and Cruz, 2022).

tors can objectively establish their use. [Figure 1](#) presents the general business functions considered in the survey and the possible technologies that can be used to conduct each of them. The grid contains 7 GBFs: Business administration, production planning, sourcing and procurement, marketing, sales, payment methods, and quality control. Each function considers between 4 and 7 technologies. For example, an establishment can gather and analyze customer information for marketing purposes using face-to-face conversations, online chats via WhatsApp or the internet, structured customer surveys, customer relationship management (CRM) software to store contact information, interaction history, and communication preferences, or big data analytics and/or artificial intelligence to uncover trends and make informed marketing decisions. [Figure 2](#) presents the grid for one sector-specific module, agriculture. The grid considers six SSBFs for agriculture which include land preparation, irrigation, weeding and pest management, harvesting, storage and packaging. For example, to prepare the land for cultivation, a farm can use manual labor with simple tools such as hand-held hoes, or rakes, animal-aided instruments such as ploughs, equipment manually operated such as tractors, motor tillers, or rotators, or equipment supported by digital technologies such as GPS, software or precision agriculture tools. [Section A.1.1](#) of the appendix reports the grids for all other SSBFs and the implementation manual precisely defines each of the technologies in the grid.

2.3 Ranking of Technology Sophistication

In addition to identifying key business functions and relevant technologies, industry experts ranked the technologies in each function based on their sophistication. More sophisticated technologies can perform a wider variety of tasks, more complex tasks, or perform tasks with greater accuracy and speed. The experts' deliberations and resulting sophistication rankings, shown on the grid, were produced before the survey administration. This approach to ranking technologies resembles the World Management Survey ([Bloom and Van Reenen, 2007](#)), which relies on experts to rank management practices according to their quality.

Given the importance of the ranking for our analysis, we evaluated the coherence of the expert rankings using two distinct methodologies in two subsamples of BFs. The first is a three-stage validation process implemented in 14 of the 63 business functions on the grid including most of the GBFs and SSBFs in agriculture, food processing apparel, and retail. The three stages are as follows:

1. **Comparison of Key Features:** We compared the technologies in each business function along three dimensions invoked by the experts: functionality, integration, and automation. Functionality refers to the capabilities a technology offers to handle more complex

tasks, in a faster way, on a larger scale, with greater accuracy and reliability. Integration reflects a technology’s ability to connect and interact seamlessly with other systems by exchanging data and coordinating processes. Automation enables the technology to execute processes, make decisions, and generate outcomes independently, without human intervention.

2. **Novelty and Cost:** We documented the year of invention and the cost of each technology and studied their correlation with the experts’ rankings. Although novelty and cost do not define sophistication, more sophisticated technologies tend to be newer and more expensive.

3. **Large language models (LLMs):** We asked ChatGPT to rank the technologies based on their levels of sophistication. We replicate this exercise following specific definitions of sophistication based on functionality, integration, and automation.

To collect the information in the first two stages, we relied on multiple sources, including the official description of specific leading brands supplying these technologies. For GBFs we collected information from multiple companies websites, including Microsoft, Google, SAP, Oracle, QuickBooks, IBM, Sage, NetSuite, BambooHR, Trello, Salesforce, Workday, Meta, Qualtrics, Survey Monkey, Amazon, Shopify, LinkedIn, among others. These companies have more than 80% of the global market share for technologies used in business administration, such as standard software (e.g, spreadsheet) and enterprise resource planning (ERP) systems, and a large share of the market across GBFs.⁷ In addition, we consulted specialized websites (e.g, tech.co; erpresearch.com; getapp.com) that provide comparisons across these products, totaling more than 50 original sources of information. Similar exercises with multiple sources of information were replicated for SSBFs.

We illustrate the validation methodology using the example of business administration, a GBF that includes finance, accounting, and human resources processes. Table 1 summarizes the three-step validation procedure for each technology in business administration. The least sophisticated technology, handwritten processes, can only perform basic manual administration tasks such as transaction entry, bookkeeping, or employee records handling, without any integration or automation features. Standard software like Microsoft Excel or Google Sheets helps with basic functionality to perform mathematical and statistical operations, including charts, and handle financial account, and HR records. However, it requires manual inputs and knowledge to build specific applications, with limited integration and automation. Mobile apps, such as QuickBook online, are pre-designed to perform these tasks with some

⁷Estimates based on Enlyft dataset (Cirera, Comin and Cruz, 2022). These companies are recognized as key players by various specialized sources estimating market potential for ERP (e.g., Research and Market, Fortune Business Insight), even if there are variations in their market share estimations.

integration and automation features, but with limited scale and customization. Specialized software, such as Oracle Financial, have high functionality, integration, and automation capabilities, within specialized domains. Finally, enterprise resource planning (ERP systems), such as SAP and Oracle NetSuite provide comprehensive functionality with full integration within and across business functions, with a high level of automation. Comparing the features of business administration technologies results in a ranking that matches the grid’s sophistication ranking.

More sophisticated technologies in business administration are embodied in more expensive software. For example, standard software such as Microsoft 365 (which includes Microsoft Excel) and Google Sheets costs between \$12 and \$18 per user/month; apps such as Quickbooks Online and BambooHR cost between \$30 and \$200 per user/month; specialized software such as Oracle Financials, Intuit Quickbooks and Workday HCM cost between \$120 and \$600 per user/month; ERP systems such as SAP ERP or Oracle Netsuite cost over \$1700 per user/month. In business administration, there is no clear relationship between technology novelty and our sophistication ranking (e.g., the SAP ERP system was introduced in 1981, while Microsoft Excel was first available in 1985). As shown in section A.2 of the appendix, for most SSBFs, especially in agriculture and manufacturing, we observe a positive and strong association between technology novelty, cost, and our sophistication rankings.

ChatGPT’s ranking of the technologies in business administration based on functionality, automation, and integration coincides with the expert ranking. Furthermore, the ranking is robust to variations of the prompts provided to ChatGPT focusing on specific dimensions of technology sophistication (e.g, exclusively based on functionality, integration, or automation). The ChatGPT rankings are strongly and positively associated with the experts’ rankings, following a similar ranking order of sophistication in all iterations.

We conduct a second exercise to validate the grid technology rankings. This complements the previously described three-step procedure by offering a distinct approach and covering different BFs. Specifically, we use primary sources to independently compute the relative productivities of technologies within each BF, assessing 18 BFs across retail, agriculture, banking, healthcare, and general business functions. For each function, we identify a central task essential to its activities and define an appropriate productivity measure reflecting the time, effectiveness, or accuracy with which the task can be completed using each technology.

We then compute technology productivity in each BF using three types of sources: (i) research published in academic journals, (ii) field conversations with a distinct set of experts from those involved in the grid assembly, and (iii) industry studies and reports. Table A.14 summarizes the BFs, tasks, productivity measures, and source types used. The Appendix

provides detailed information on sources, calculations, and relative productivity estimates. In Section 3, we use these productivity values to construct an alternative cardinalization of the ordinal technology rankings, demonstrating the robustness of our findings to the chosen cardinalization. In all cases, productivity increases with technological sophistication.

Overall, the validation exercises across all the business functions explored supports the experts' ranking of sophistication.⁸

2.4 Information Collected in FAT

The survey collects information in three broad areas: the business functions conducted by an establishment, the use of technologies in each business function, and information on the establishment's financial statements, workers, and management.

Business functions. The business functions that comprise the horizontal dimension of the grid cover the key tasks involved in production. Explorations conducted at the piloting stage of the survey as well as the responses to the questions on the use of technologies in GBFs demonstrate that these functions are conducted in-house and that respondents are aware about the technologies their establishments use in the GBFs.⁹ We formally explore the relevance of each sector-specific business function in each establishment through a screener question that asks whether a sector-specific function is conducted in that establishment. This information helps us assess the relevance of establishment-level measures of technology sophistication based only on the technologies used in functions conducted in-house.

Technology questions. The survey has two types of questions about the technologies used to conduct each business function. First, it asks whether the establishment uses each of the technologies listed in the grid. After identifying the technologies that are used by the establishment in a business function, the survey asks which technology is the most widely used in that function. The answers to these questions permit us to differentiate between the range of technologies present in the business function vs. the intensity with which they are used.

FAT also asks whether the establishment uses "other technologies" in the business function

⁸In section A.2 of the appendix, we provide the results for the other 13 business functions where we implement the validation of the expert rankings.

⁹Due to space constraints in the survey and the information revealed during the pre-pilot, we decided to not directly ask about whether establishments conduct each GBF in FAT. Proxying the fraction of GBFs that are not conducted in house by the share of GBFs for which the establishment responds that either "does not use" or "does not know if it uses" to all the technologies in the grid for the BF, we find that only 3.9% of GBFs are not conducted in-house.

in addition to those contained in the grid. Only in 3.6% of the business functions establishments declare that “other” technologies are used in the business function, and only in 0.8% of the business functions "other" is the most widely used technology. The low frequency of “other” demonstrates the comprehensiveness of the technologies in the grid.

Other variables. The survey also includes other standard questions about financial statements’ information, employment, education of the employees, and education and experience of the manager. The survey collects information on four management practices from MOPS, including the presence of formal incentives, number of key performance indicators (KPIs), frequency of KPI review, and time frame of production targets. The answers to these questions are used to construct a management z-score following the methodology in [Bloom and Van Reenen \(2007\)](#). Despite covering only four of the 16 variables collected in MOPS, the FAT z-score based on this subset of questions accounts for 90.5% of the cross-establishment variance of the original MOPS z-score for Mexican establishments collected by ENAPROCE.

2.5 The Data

Our analysis is based on primary data collected from establishments in 15 countries: South Korea, Poland, Croatia, Chile, Brazil (Ceará), Georgia, Vietnam, India (Uttar Pradesh, Tamil Nadu, Gujarat and Maharashtra), Ghana, Bangladesh, Kenya, Cambodia, Senegal, Ethiopia, and Burkina Faso. Several factors were considered in deciding where to implement the FAT survey. We targeted countries on different continents (Asia, Africa, South America, and Europe), with different levels of income, for which there was access to a high-quality sampling frames. In these countries, we collected data from 21,055 randomly selected establishments from the sampling frames. [Table 2](#) shows the distributions of the sample by country, sector, and size groups and [Table C.1](#) provides descriptive statistics. The median establishment in our sample has 9 workers, with an average of 34 workers. 20% of workers have a college degree, 19% of firms were multi-establishments, 18% are part of a multinational firm, 17% are exporters, 18% are 5 years old or younger, and 76% have electricity, computers, and internet access.

2.5.1 Sampling

Our data is representative for a universe of about 2.1 million establishments. The samples are nationally representative for establishments with 5 or more workers. For each country, the sampling frame is based on the most comprehensive and up-to-date establishment-level census data available from the respective National Statistical Office (NSOs) or similar au-

thority. The survey is stratified on three dimensions - sector, firm size, and region - so that we can construct representative measures of technology for aggregates along these dimensions. Sampling weights are based on the inverse probability of selecting establishments within each stratum.¹⁰

2.5.2 Measures to minimize bias and measurement error

The literature on survey design has identified three types of potential bias and measurement errors based on whether they originate from non-responses, the enumerator, or the respondent (Collins, 2003). In what follows, we briefly describe the steps taken in designing and implementing the FAT survey to minimize these errors. Appendix A.6 provides a more detailed description of the measures implemented to minimize potential bias.

Non-response bias. To maximize response rates and minimize potential biases associated with non-response (Gary, 2007), we followed best practice procedures. First, we partnered with national statistical offices and industry associations to use the most comprehensive and updated sampling frame available. Second, we hired data collection companies or agencies which were supported by endorsement letters from local institutions and which had demonstrable experience in nationally representative firm-level surveys. Third, we followed a standard protocol in which each firm was contacted several times to schedule an interview. Fourth, we mostly used face-to-face or phone interviews, which usually have higher response rates than web-based interviews.¹¹

Enumerator bias and error counts. The survey, training, and data collection processes were designed to minimize enumerator biases and data collection errors. First, we used closed-ended questions to make coding the answers a mechanical task, thereby eliminating the need for the enumerator to interpret the answers or exercise subjective judgement when coding them. Second, the same standardized training was implemented in each country in the local language, with enumerators, supervisors, and managers leading the data implementation. Third, we conducted a pre-test pilot of the questionnaire in each country using

¹⁰Table A.22 provides information about the distribution of firms by country, sector, and size groups within the universe covered by the FAT survey. For the state of Ceará in Brazil and the Indian states of Tamil Nadu, Uttar Pradesh, Gujarat, and Maharashtra, it is representative at the state level. Section A of the Appendix provides more details on the sampling frame, survey implementation and data collection, and sampling weight.

¹¹These procedures are in line with suggestions of good practice for implementation by (Bloom et al., 2016). We use online surveys only for Georgia and Croatia. In Georgia, we partnered with the National Statistical Office, which resulted in exceptionally high response rate. Face-to-face interviews were not possible during the pandemic. See Table A.23 for the mode and date of data collection in each country.

establishments not included in the sample. Fourth, to attain greater quality control during the data collection process, enumerators recorded the answers via *Computer-Assisted Personal Interviews* (CAPI) or *Computer-Assisted Telephone Interviewing* (CATI) software, and we regularly monitored the data collection process using standard algorithms to analyze the consistency of the data.¹²

Respondent bias. We took several steps to minimize respondent bias. First, we ensured that the interview was arranged with the appropriate person or persons; main managers (and other managers, such as plant managers and HR managers, in larger firms). Second, we used a closed-ended design in the questionnaire such that the respondent was questioned about specific technologies one at a time and was not told beforehand all the technologies that were associated with each business function. This design reduced measurement error in respondent’s answers. Third, we pre-tested the questionnaire in each country to ensure that our questions were clearly worded within the specific geographical and cultural contexts of each country, reducing the need for subjective judgment in responses (Bertrand and Mullainathan, 2001). Fourth, to avoid *social desirability bias*, which may cause respondents to overstate the use of more sophisticated technologies, the survey avoided the words "technology" and "sophistication", employing more neutral terms such as "methods" and "processes" instead.

2.5.3 Ex-post checks and validation exercises

We conducted several ex-post checks to assess the quality of the collected data.

Non-response bias. The average (unit) response rate on the survey varies by country and ranges between 15% and 86%. For example, the response rate was 80% in Vietnam, 57% in Senegal, 39% in Ceará, Brazil, 24% in Korea, and 15% in Croatia. These response rates are high relative to typical response rates in establishment-level surveys, which are around 5 to 10% and are consistent with response rates observed for WMS which are around 40% (Bloom et al., 2016). To minimize potential non-response bias, we adjusted the sampling weights for unit non-response. The adjustment was calculated at the strata level, so that the weighted distribution of our respondent sample across strata (sector, size, region) exactly matches the distribution of establishments in the sampling frame.¹³ We conducted three

¹²Randomized survey experiments with household surveys have demonstrated that a large number of errors observed in *Pen-and-Paper Personal Interview* (PAPI) data can be avoided with CAPI or CATI (Caeyers, Chalmers and De Weerd, 2012). For Georgia and Croatia, we used Computer Assisted Web Interviewing (CAWI).

¹³Table A.24 in the Appendix A provides the response rate by country, defined as the ratio between establishments that responded to the survey and the total number of eligible establishments in the sample

tests to assess potential biases from unit non-response-rates.¹⁴ In each of these exercises, presented in Section A.7 of the Appendix, we find no statistical difference in the number of employees, technological sophistication, wages, and share of workers by skill and education between establishments in the group that proxies for the response sample and the group of establishments that proxies for the non-response sample.

Response bias. To assess the relevance of response bias, we conducted a parallel pilot in Kenya where we re-interviewed 100 randomly selected establishments with a short version of the questionnaire. For those establishments, we randomly selected three business functions and asked about the presence of the relevant technologies. We estimated a probit model to assess the likelihood of consistent answers between the original and the back-check interviews, controlling for establishment-level fixed-effects. Reporting the use of a technology in the back-check interview is associated with 80.6% of the likelihood of reporting the use of the same technology in the original interview. Conversely, reporting that a technology is not used in the back-check interview, is associated with a 70.7% likelihood of not being reported in the original survey. These estimates do not differ between establishments of different sizes.¹⁵

Validation using external sources. We evaluate the quality and reliability of the data collected by comparing it to external sources in Korea (KED) and Brazil (RAIS). We focus on variables related to establishment size, productivity and technology. Table A.31 shows that the weighted sample averages of the labor variables in the FAT data (number of workers, average wages, share of college workers, share of low- and high-skill workers) are not statistically different from the averages in the universe of firms from the RAIS dataset. In the Brazil matched establishments, we find a strong correlation between FAT measures of log value-added per worker and the log of average wages from RAIS (See Table A.30). In the Korean

for which we attempted to conduct an interview. The response rates were higher when national statistical agencies implemented the survey. Section A.5 of the appendix provides more details on sampling weights.

¹⁴First, using the information from the sampling frame, we check if there are differences in the average number of workers per establishment between respondents and non-respondents within stratum. Second, using information on the number of contact attempts, we compare the establishment-level technology sophistication in GBFs, described in the next section, between establishments with above and below the average number of attempts. Third, in a similar vein, we compare establishments in the first list of contacts provided to interviewers, versus those provided subsequently. See Table A.25 to A.31 in Appendix A.

¹⁵The re-interviews produced 1,661 answers, 106 interviews times 3 business functions times an average of 5.2 technologies per function. Both the original and back-end interviews in the pilot are conducted by phone by different interviewers. The correlation between the binary responses in survey and pilot is 73% ranging from 65% in business administration to 77% in sales across business functions, and from 85% among the most basic technologies to around 61% in intermediate, and 77% at the most advanced technologies across functions.

matched establishments, we find very high cross-establishment correlations (above 0.93) in the log levels and growth rates of sales and employment, as well as in log labor productivity (0.73).¹⁶ Additionally, the average adoption rate of ERP systems in Korean manufacturing establishments in FAT is similar to [Chung and Kim \(2021\)](#), who used a similar sampling frame (32% vs. 40% in [Chung and Kim, 2021](#)), and there is a strong cross-establishment association between the book value of machinery and equipment in KED and the establishment technology sophistication measures (MOST and MAX) from FAT, which will be explained in the next section.

Correlations between top-tier technologies and firm characteristics. We conduct an additional validation exercise of the technology measures, by studying whether establishments with larger sales, employment and sales per worker are more likely to use top-tier technologies, which are the more sophisticated technologies in each BF and are marked in bold in [Appendix A.1](#). Specifically, we estimate a linear probability model for each business function, where the dependent variable is binary and equal to 1 if the establishment uses one of the technologies classified as top-tier for the business function and 0 otherwise. The model includes a full set of country- and, for the GBFs, 2-digit sector fixed effects. The independent variables are either (log) sales, (log) employment or (log) sales per worker. We find that the coefficients for these variables are positive and significant in a large majority of business functions.¹⁷

These ex-post checks further reassure us about the soundness of the survey design, the data collection process, and the accuracy of responses.

3 Measures of Technology Sophistication

We next describe how we use information collected in FAT to construct technology measures at the business-function and establishment levels. First, we introduce the measures. We then assess the robustness of the establishment-level measures to the cardinalizations we impose. Finally, we present examples that illustrate how the measures characterize establishments’

¹⁶In Korea we merge FAT with the Korea Enterprise Data (KED), a leading supplier of business credit reports on Korean businesses. In Brazil, we merge the data with the *Relação Anual de Informações Sociais* (RAIS), which is an administrative database maintained by the Ministry of Labor providing information on salaries for all formal workers in Brazil. The FAT survey asks about sales and the number of employees for two periods. The most recent year for which the information is available (i.e. the year before the implementation of the survey) and two years before that. For Korea, these reference years are 2019 and 2017.

¹⁷For sales we find a positive coefficient in 100% of BFs (85% significant at 5% level); for employment 98% are positive (93% significant); and for productivity 80% are positive (52% significant, and never negative and significant).

technology sophistication as well as the granularity of the grid.

3.1 Measuring Technology Sophistication

Only 16% of establishments in FAT use an ERP system, the most sophisticated technology for business administration. Among establishments without an ERP, there is considerable variation in the most sophisticated technologies used to conduct business administration tasks: 27.5% use a computer with specialized software, 8% use a mobile app or digital platform, 38% use a computer with Excel, and 10% rely primarily on handwritten processes.

Our goal is to construct measures of technology sophistication that describe the technologies used in all establishments and that specifically reflect the diversity of technologies used when advanced technologies are not available. The number of technologies used in a function is not an accurate measure of sophistication because technologies differ widely in sophistication. Instead, the measures we propose reflect the sophistication of key technologies employed by establishments, relative to the world technology frontier in the business function.

BF-level measures We start by defining $ANUM_{f,j}$ as the number of different technologies from the grid used in business function f by establishment j . When more than one technology is used in a business function, we explore whether the technologies used are contiguous in the sophistication ranking of the grid or, instead, there are sophistication gaps in the vector of technologies used. Formally, we define the sophistication gap of establishment j in business function f ($SG_{f,j}$) as a binary variable that takes the value of 1 if the establishment uses technologies with sophistication rank τ and $\tau + k$ for $k \geq 2$ in function f but does not use the technology with sophistication rank $\tau + p$, for $1 \leq p < k$. $SG_{f,j}$ is 0 when there are no gaps and at least two technologies are used in the function.¹⁸

We study the sophistication of the technologies used in a business function with the help of two variables. $MAX_{f,j}$ measures the sophistication of the most sophisticated technology used in the given business function, while $MOST_{f,j}$ reflects the sophistication of the most widely used technology in the business function. The starting point to construct these measures is the experts’ rankings of the technologies, from least to most advanced, $r_f \in 1, 2, \dots, R_f$.¹⁹

¹⁸ $SG_{f,j}$ is not defined when less than two technologies are used in the function (i.e. $ANUM_{f,j} < 2$).

¹⁹Because several technologies may be assigned the same sophistication, the highest rank in a function R_f may be smaller than the number of possible technologies N_f . In a small number of business functions, the technologies covered are used in various subgroups of tasks. For example, in the body-pressing and welding functions of the automotive sector, the survey differentiates between technologies used for pressing skin panels, pressing structural components and welding the main body. In cases like this, we construct ranks of technologies for each subgroup of tasks within the business function, and then aggregate the resulting indices by taking simple averages across the tasks groups. See [Appendix B.1.1](#) for more details.

We define the relative rank of a technology with absolute ranking r_f as $\hat{r}_f = \frac{r_f - 1}{R_f - 1}$. Note that $\hat{r}_f \in [0, 1]$. Our baseline measure of technology sophistication is an affine transformation of the relative rank, \hat{r}_f . Specifically, we define $MAX_{f,j}$ and $MOST_{f,j}$ as

$$MAX_{f,j} = 1 + 4 * \hat{r}_{f,j}^{MAX}, \quad MOST_{f,j} = 1 + 4 * \hat{r}_{f,j}^{MOST} \quad (1)$$

where $\hat{r}_{f,j}^{MOST}$ and $\hat{r}_{f,j}^{MAX}$ are the relative sophistication rankings of the two technologies. We also use a similar transformation to define a scaled measure of the number of technologies used in a business function ($NUM_{f,j}$).²⁰

By construction, $MAX_{f,j}$ and $MOST_{f,j}$ are in the $[1, 5]$ interval, and $MAX_{f,j} \geq MOST_{f,j}$. Since the most sophisticated technologies in the grid reflect the current (world) technology frontier, $MAX_{f,j}$ and $MOST_{f,j}$ represent the closeness of an establishment to the technological frontier in a business function. Nevertheless, $MAX_{f,j}$ and $MOST_{f,j}$ are conceptually different. $MAX_{f,j}$ increases when a firm implements a new technology that is more sophisticated than those currently used in a given business function. This technology may not be new to the establishment, but it is new to the business function of the establishment. Therefore, increases in $MAX_{f,j}$ reflect the adoption of new technologies.

$MOST_{f,j}$ increases when an establishment changes the most widely used technology in a business function to one that is more sophisticated. The new most widely used technology may be entirely new to the business function in which case the change in MOST reflects the adoption of a new technology. However, changes in MOST may also reflect the extension in the use of a technology that the establishment was already using in the business function. In this case, $MOST_{f,j}$ is more related to Mansfield (1963)'s concept of technology diffusion within the firm rather than to traditional notions of adoption or innovation. The potentially different technology-upgrading processes that drive $MAX_{f,j}$ and $MOST_{f,j}$ motivate two key empirical questions in this paper. First, are the $MAX_{f,j}$ and $MOST_{f,j}$ technologies different in the typical business function? Second, are the same factors associated with these two dimensions of technology sophistication in a business function?

Establishment-level Measures. Relevant outcomes and observable characteristics are often reported at the establishment level. To study their relationship with technology sophistication requires constructing establishment-level measures by aggregating up the function measures. This raises two separate issues. The first is that FAT only provides information on the technology sophistication of the business functions an establishment conducts in-house. An establishment-level measure that focuses on these functions, omits the sophistication of

²⁰Formally, we define $NUM_{f,j}$ as $NUM_{f,j} = 1 + 4 * \frac{ANUM_{f,j} - 1}{N_f - 1}$, where N_f is the number of different technologies in the grid for the business function f .

technologies used in functions that the establishment outsources to other establishments. This omission is not important if establishments outsource a small number of functions. As discussed in [section 2](#), the pre-pilot, together with the answers to the technology questions in FAT, strongly suggest that GBFs are conducted in-house in an overwhelming majority of establishments. Similarly, 87% of the relevant sector-specific business functions are conducted in-house. Therefore, the technology sophistication of in-house functions is a good proxy for the sophistication of the technologies establishments have access to both directly and indirectly via the sourcing of functions. Reassuringly, all the establishment-level findings we present are robust to controlling for the fraction of functions an establishment conducts in-house ([Cirera, Comin and Cruz, 2024a](#)).

The second issue concerns the aggregation of function-level measures to construct measures of technology sophistication at the establishment level. We follow the standard practice [Bloom and Van Reenen \(2007\)](#) of computing establishment-level measures by the simple mean of the business function measures.²¹ Formally, we construct establishment-level counterparts of $NUM_{f,j}$, $MAX_{f,j}$ and $MOST_{f,j}$ as:

$$S_j = \sum_{f=1}^{N_j} \frac{S_{f,j}}{N_j} \quad (2)$$

where $S = \{NUM, MAX, MOST\}$, and N_j is the number of business functions covered for establishment j .

3.2 On the cardinalization of technology sophistication measures.

We conduct two types of exercises to validate the cardinalization imposed by our baseline sophistication measures. A common practice to cardinalize ordinal variables is to project them into a relevant cardinal variable ([Cunha, Heckman and Schennach, 2010](#); [Bond and Lang, 2019](#)). Our first exercise to validate our cardinalization follows this approach and studies whether the relationship between technology-sophistication and two measures of productivity across establishments is well approximated by a linear function. Specifically, we estimate

$$Y_j = \alpha_c + \alpha_s + \beta * X_j + \gamma(I(\bar{S}_j)) * \bar{S}_j + u_j \quad (3)$$

Y_j is establishment productivity measured either by the the (log) of real sales per worker or by the Solow residual,²² α_c and α_s are country and 2-digit sector \bar{S}_j is a simple average of

²¹This is the first-order approximation to a CES aggregator when the weights of the different functions are similar.

²²Real variables are constructed by deflating nominal variables by the country/year World Bank PPP price index. Solow residuals are constructed as the residuals from regressing the log of real sales on the log

MAX_j and $MOST_j$, and $\gamma(I(\bar{S}_j))$ indicates that we allow the coefficient of \bar{S}_j to differ across subsamples defined by intervals of \bar{S}_j . We consider both a two-interval split —distinguishing establishments above and below the median value of \bar{S}_j —and a four-interval split.²³

When using sales per worker as the dependent variable, we estimate specifications both with and without controls, X_j , for human and physical capital.²⁴ The estimates reported in [Table C.4](#) of the Appendix show a strong association between technology sophistication and establishment productivity, and that this relationship is well approximated by a linear function.

Beyond this evidence supporting linear cardinalizations of technology sophistication rankings, we explore the sensitivity of the analysis to our cardinalization by considering two alternative cardinalizations of the sophistication rankings. The first is the z-score cardinalization used by [Bloom and Van Reenen \(2007\)](#). We construct z-score versions of $MAX_{f,j}$ and $MOST_{f,j}$ by subtracting the mean sophistication rankings in the BF and dividing by the standard deviation:

$$MAX_{f,j}^Z = \frac{\hat{r}_{f,j}^{MAX} - \bar{r}_f^{MAX}}{std(\hat{r}_{f,j}^{MAX})}, \quad MOST_{f,j}^Z = \frac{\hat{r}_{f,j}^{MOST} - \bar{r}_f^{MOST}}{std(\hat{r}_{f,j}^{MOST})} \quad (4)$$

where \bar{r}_f^{MAX} and \bar{r}_f^{MOST} are the means of the sophistication rankings and $std(.)$ denotes the cross-establishment standard deviation. Unlike the baseline cardinalization, z-scores do not capture the absolute distance to the frontier sophistication in the function, but the distance relative to the cross-establishment dispersion in sophistication in the function.

The second cardinalization is non-linear and based on independent information collected from scientific papers, experts, industry reports, and company publications on the relative productivity of technologies for representative tasks within each business function. We have collected this information for 18 BFs, reported in [Table A.14](#) and discussed in sections [2.3](#) and [A.3](#). Formally,

$$MAX_{f,j}^Q = 1 + Q_{f,j}^{MAX}, \quad MOST_{f,j}^Q = 1 + Q_{f,j}^{MOST} \quad (5)$$

where $Q_{f,j}^{MAX}$ and $Q_{f,j}^{MOST}$ are, respectively, the productivity of the most sophisticated and most widely used technologies by establishment j in function f relative to the productivity of the least sophisticated technology in the function.

of the number of employees, the log of the real book value of capital, the log of the real cost of materials, and the log of the real cost of energy.

²³The four intervals are [1-1.5), [1.5-2.5), [2.5,3.5), and [3.5,5]. These intervals are constructed to ensure that each contains a substantial share of establishments in the sample.

²⁴ H_j is the fraction of college-graduated workers, and K_j is the log of the book value of capital per worker. Real values are computed as nominal values deflated by the World Bank PPP for the relevant year.

We assess the robustness of the establishment-level measures of sophistication to the cardinalization by conducting four exercises. First, we compute the correlation across establishments of the baseline and the z-measures of sophistication. For both MAX and MOST, the correlation is over 0.98. Second, for each BF where we compute the relative productivities $Q_{f,\tau}$, we calculate the correlation between the cardinalizations in the baseline and using $Q_{f,\tau}$. The median correlation across functions is 0.95, with a p25-p75 range of 0.9-0.98. This shows that the linearity imposed by the baseline cardinalizations is not a bad approximation to the actual productivity gaps between technologies in the BF reflected by the Q-cardinalizations.

In addition to linearity, the baseline cardinalization imposes a common range across BFs. To assess the impact of this restriction on establishment-level sophistication measures, we compute establishment-level MAX and MOST measures based on Q and baseline cardinalizations using the 18 functions for which we have Q-cardinalizations. The correlation between baseline and Q-based sophistication measures across establishment is 0.9 for MAX and 0.91 for MOST.²⁵ We also compare the correlation between baseline and Q-based measures of establishment-level sophistication and productivity — measured by the log of real sales per worker — and human capital — measured by the share of employees with college education.²⁶ The correlations between establishment characteristics and sophistication measures are similar for baseline and Q sophistication measures for both MAX and MOST (see Table 3).

These exercises confirm that establishment-level technology sophistication measures and their analysis are robust to reasonable cardinalizations of the ordinal sophistication rankings.

3.3 An Illustration

Before studying the general patterns of technology use in establishments, it is useful to become familiar with the grid and the measures of technology sophistication by exploring some examples from FAT.

To appreciate the level of detail in the grid, we examine two medium-sized establishments in apparel retail: one in India (establishment 1) and the other in Vietnam (establishment 2). Figure 3 plots the $MAX_{f,j}$ index in each business function for both establishments. For instance, establishment 1 uses a dynamic pricing system that automatically adjusts prices

²⁵Before computing the correlations we have residualized all variables by regressing them on 2-digit sector dummies. This corrects for the variation across sectors in the number of functions for which we have Q-cardinalizations. Alternatively, we have conducted the exercise using only information for the three general business functions with Q-cardinalizations. In this case, the correlation between baseline and Q-based sophistication across establishments is 0.97 for MAX and 0.98 for MOST.

²⁶In all cases, we residualize the variables by regressing them on a full set of 2-digit sector dummies.

based on demand conditions, while establishment 2 uses an automated markup technology that collects information on costs and applies a uniform markup. The pricing business function includes five technologies (Figure A.7). From least to most sophisticated, these are: manual pricing (prices set without a formal account of the costs), automated markup, automated promotional pricing (prices adjusted based on seasonal factors), dynamic pricing, and personalized pricing (prices adjusted at the individual customer level using data analytics such as data mining and machine learning). The value of the MAX index for establishment 1 (4) indicates that it is one notch below the technology frontier in pricing, while the value for establishment 2 (2) shows that it is three notches below the frontier.

Establishment 1 also has a higher MAX level in other business functions such as merchandising and inventory. In these functions, it uses digital merchandising systems (DMS) and automated inventory controls, respectively. In contrast, establishment 2 selects products to display on shelves manually and uses a warehouse management system with specialized software.²⁷

Nevertheless, establishment 2 has a higher MAX index than establishment 1 in other functions such as customer service, quality control, and sales. Specifically, establishment 2 attends to customer requests made online, checks product quality using statistical process control with software monitoring, and sells its products online using an external digital platform. In contrast, establishment 1 attends to customer requests over the phone, checks product quality manually with the support of digital technologies, and sells products directly at the establishment.

The variation in the relative technology sophistication rankings of establishments 1 and 2 across different business functions underscores the importance of a comprehensive coverage when characterizing the technological sophistication of an establishment. Focusing on one or a few functions or technologies provides an imprecise, and possibly biased, characterization of the sophistication of the technologies used in an establishment.

Next, we move from the establishment to the sector level and explore the cross-establishment distribution of technology sophistication in the food processing sector. We focus on the fabrication business function, which is relevant for all manufacturing establishments in FAT. Some of the technologies it covers, such as numerically controlled machines and robots, have been widely studied in automation research. The grid considers six classes of technologies. In increasing order of sophistication, these are (1) manual processes, (2) machines controlled by operators, (3) machines controlled by computers, (4) robots, (5) additive manufacturing in-

²⁷DMS is used to execute core merchandising activities, including product management, inventory replenishment, purchasing, vendor management, and financial tracking. One example of automated inventory controls is Computer Assisted Ordering (CAO), an inventory replenishment system that can use either sales or inventory algorithms to prepare a suggested reorder.

cluding rapid prototyping and 3D printing, and (6) other advanced manufacturing processes such as laser, plasma sputtering, high-speed machine, E-beam and micro-machining).

The top right panel of Figure 4 plots the distribution of $MAX_{f,j}$ in fabrication across food processing establishments in South Korea. The histogram reveals a significant dispersion across establishments in the most sophisticated technologies available for production in fabrication. Establishments that process food using the world frontier’s fabrication technologies coexist with others that just use manual processes. It is also worth noting that, in contrast to the popular perception, most establishments do not use robots or other more sophisticated technologies (i.e., $MAX_{f,j} \geq 3$) even in such an advanced economy as South Korea.²⁸

To explore the cross-country differences in technology sophistication, the top panel of Figure 4 also plots the histogram of $MAX_{f,j}$ in Senegal (left) and India (middle). There are stark cross-country differences in the distribution. The mean and variance increase uniformly with the level of development in the country (i.e. Senegal, India, South Korea). Additionally, the cross-establishment distribution of $MAX_{f,j}$ is most skewed to the right in Senegal, and least in Korea.

The bottom panel in Figure 4 shifts the focus to the most widely used technology. In particular, it shows the histogram of $MOST_{f,j}$ in fabrication across food processing establishments for each country. By construction, the distribution of $MAX_{f,j}$ stochastically dominates the distribution of $MOST_{f,j}$, as $MAX_{f,j} \geq MOST_{f,j}$. However, the distributions of MAX and MOST differ significantly. For example, in 65% of Indian food processing establishments the most sophisticated technology used in fabrication is ‘machines controlled by operators.’ Yet, this technology is the most widely in only 35% of establishments.

The gap between MAX and MOST in this example motivates a deeper exploration of whether MAX and MOST are statistically distinct across a broad range of business functions and countries and, if that is the case, their relative importance in shaping the relationship between technology sophistication and productivity across establishments.

4 Technology Sophistication at the BF level

We use the FAT dataset to examine the use of technologies at the business function level. First, we explore the relationship between MAX and the vector of technologies used in a BF. Second, we compare MAX and MOST and study the frequency of gaps between

²⁸This is also true in other sectors with greater penetration of robots such as the automotive sector. Even if we weight establishments by size (e.g., employment or sales) only 56% of establishments in automotive fabrication in South Korea use robots or a more sophisticated technology in fabrication.

MAX and MOST. Finally, we assess whether MAX-MOST gaps are temporary and examine their association with establishment characteristics that may shed light on potential drivers. Notably, the analysis in this section relies exclusively on the ordinal sophistication rankings imposed by the grid.

4.1 MAX and the Vector of Technologies Used in the BF

We begin our analysis of the vector of technologies used in a business function by counting the number of technologies from the grid that an establishment uses, denoted as $ANUM_{f,j}$. Table C.2 reports the average $ANUM_{f,j}$ for each function across all establishments that conduct the function in-house. On average, establishments use two different technologies per function, a pattern that holds consistently across both general and sector-specific business functions. The distribution of $ANUM_{f,j}$ reveals that 62.6% of functions use more than one technology, and 28.3% use at least three.

The vector of technologies used in a business function is relevant to two economic literatures. Schumpeterian models predict that when an establishment adopts a more sophisticated technology, it will abandon the less sophisticated ones previously used. Meanwhile, technological leapfrogging models suggest that late adopters may skip intermediate technologies altogether, choosing instead to adopt only the most advanced options available.

Using FAT, we examine whether establishments abandon or skip less sophisticated technologies. We document both instances where establishments clearly abandon or skip older technologies and cases where they retain them. We first analyze business functions where only one technology is used, which account for 37.4% of observations. Among these, in 52.8% of the cases establishments use the least sophisticated technology in the grid, meaning that only 18% ($37.4\% * (1-52.8\%)$) of functions have completely skipped or abandoned less sophisticated technologies in favor of more advanced ones.

A related but distinct phenomenon is the presence of sophistication gaps, where establishments use multiple technologies that are not contiguous in the sophistication space. These gaps are relatively infrequent, occurring in 25% of business functions – 27% in GBFs and 17% in SSBFs. The GBFs where gaps are more common include payments (48%), business administration (34%), and sales (28%).

Our analysis highlights the rarity of cases where less sophisticated technologies are fully skipped or abandoned. Instead, establishments frequently continue to use less sophisticated technologies even after adopting more advanced ones. In 70.4% of functions where multiple technologies are used, one of the technologies employed is the least sophisticated technology available in the grid for that function. This shows that establishments typically retain

dominated technologies rather than skipping or abandoning them.

The infrequency of sophistication gaps, combined with the regular use of the least sophisticated technology, implies that we can approximate the entire vector of technologies used in a business function just by the most sophisticated technology it uses. In other words, $MAX_{f,j}$ is a good proxy for the entire adoption history of the establishment in a given business function.

To further explore the relationship between $MAX_{f,j}$ and an establishment’s technology adoption history, we estimate the following regression:

$$MAX_{f,j} = \alpha_j + \alpha_f + \beta * NUM_{f,j} + u_{f,j} \tag{6}$$

where α_j and α_f are establishment and business function fixed effects. The point estimate of β , presented in column 2 of [Table 4](#), is 0.84.

The close to one-to-one movement of MAX and NUM suggests that the technologies that are introduced by an establishment in a business function are typically more sophisticated than existing ones. This supports the key premise of adoption models: establishments incur in the costs of implementing new technologies to gain access to more sophisticated technologies.

4.2 MOST

Research on technology has primarily focused on the presence of new or advanced technologies in establishments, neglecting the intensity of use of existing technologies. This conventional focus rests on the presumption that the technologies establishments use most intensively are the most sophisticated technologies they have adopted. Under this assumption, measuring the most widely used technology would be redundant since $MAX_{f,j}$ would be a sufficient statistic for $MOST_{f,j}$.

Departing from this tradition, we study the sophistication of the most widely used technology in business functions, $MOST_{f,j}$. Our primary objective is to determine whether $MOST_{f,j}$ is redundant or whether it provides distinct insights into the sophistication of technologies used in a business function.

We begin by examining the relationship between $MOST_{f,j}$ and $MAX_{f,j}$. Specifically, we study the fraction of within-establishment variance in $MOST_{f,j}$ that is accounted for by $MAX_{f,j}$ through the following regression:

$$MOST_{f,j} = \alpha_j + \alpha_f + \beta * MAX_{f,j} + u_{f,j}, \tag{7}$$

where α_j and α_f represent establishment and function effects. The estimates, reported in column 1 of [Table 4](#), reveal that while $MAX_{f,j}$ and $MOST_{f,j}$ are positively correlated, $MAX_{f,j}$ explains only 34% of the variance in $MOST_{f,j}$, showing that $MAX_{f,j}$ is not a sufficient statistic for $MOST_{f,j}$.

To explore why this is the case, we compare the estimates of regressing $MAX_{f,j}$ and $MOST_{f,j}$ on $NUM_{f,j}$. Specifically, we estimate

$$MOST_{f,j} = \alpha_j + \alpha_f + \beta * NUM_{f,j} + u_{f,j}. \quad (8)$$

Estimates, reported in column 3 of [Table 4](#), imply that a 1 unit increase in $NUM_{f,j}$ is associated with an increase in $MOST_{f,j}$ of just 0.25 (s.e.=0.01), suggesting a minimal impact of the adoption of new technologies on the most widely used technology. This estimate contrasts with the 0.84 (s.e.=0.01) estimate for $MAX_{f,j}$ (equation 6). The sizeable difference — statistically and economically — implies that extending the use of already-adopted sophisticated technologies is a process distinct from adopting new technologies.

Taken together, the low fraction of the variance in $MOST_{f,j}$ explained by $MAX_{f,j}$ and the divergent responses to $NUM_{f,j}$, suggest that the forces driving cross-firm technology adoption differ from those driving within-function extension of sophisticated technology use. Using a different methodology, these findings speak to a central question in the intra-firm diffusion literature ([Mansfield, 1963](#); [Battisti and Stoneman, 2003](#)): cross-firm diffusion models do not necessarily capture within-firm diffusion dynamics.

4.3 The MAX-MOST Gap

Now that we have established that MAX and MOST reflect distinct technology upgrading processes, we turn our attention to studying the frequency and nature of MAX-MOST gaps. We start by formally defining the MAX-MOST gap. In any business function where an establishment uses more than one technology, we define the MAX-MOST gap, $GAP_{f,j}$, as a binary variable that takes the value of 1 if $MAX_{f,j} > MOST_{f,j}$, and it is equal to 0 if $MAX_{f,j} = MOST_{f,j}$.

We first examine the prevalence of MAX-MOST gaps across establishments and functions. There are MAX-MOST gaps in 62% of the functions where establishments use multiple technologies. The frequency is similar for GBFs (0.62) and SSBFs (0.61). Except in health services and pharmaceuticals, the frequency of MAX-MOST gaps is above 50% in all SSBFs (see [Figure C.3](#)).²⁹

²⁹Observations are weighted by establishment sampling weights. The averages are similar if instead the weights are the sampling weights times the inverse of the number of BFs in the establishment where more

Distribution of MAX-MOST gaps across establishments. Next, we turn our focus to how MAX-MOST gaps differ across establishments. We define the establishment-level MAX-MOST gap, GAP_j , as the average of $GAP_{f,j}$ across all the functions where establishment j uses more than one technology. Figure 5 plots the distribution of GAP_j across the establishments in our sample. The median gap is 0.67 with a $p20 - p80$ range of 0.33-1 and a standard deviation of 0.3. The distribution of MAX-MOST gaps is skewed to the left. 7.2% of establishments have no MAX-MOST gaps, while 20.9% of establishments have MAX-MOST gaps in all function where they use multiple technologies.³⁰

To better understand the sources of variation of MAX-MOST gaps across establishments, we investigate their variation across and within countries. MAX-MOST gaps are the norm at all levels of development, as in each country in our sample the average MAX-MOST gap is above 50%. However, there are significant differences in the frequency of MAX-MOST gaps across countries. The country-level frequency of MAX-MOST gaps ranges from 51% in South Korea to 83% in Burkina Faso. Figure 6 plots the country-level frequency of MAX-MOST gaps against its per capita income. The plot shows that the frequency of MAX-MOST gaps decreases with income, exhibiting a correlation of -0.55.

Figure C.4 plots the within-country dispersion in MAX-MOST gaps, measured by the standard deviation of GAP_j across the establishments, against per capita income. In contrast to the average MAX-MOST gap, the within-country dispersion exhibits no statistically significant correlation with per-capita income (correlation of 0.04). However, it is remarkable that the dispersion in the MAX-MOST gaps within countries is, on average, three times greater than the between-country dispersion (0.28 vs. 0.09).

Is the MAX-MOST gap transitory? One possible explanation for MAX-MOST gaps is that they reflect a transitory state in which establishments gradually expand the use of newly adopted technologies. Under this interpretation, the gaps would exhibit dynamics akin to vintage capital models (e.g., Benhabib and Rustichini, 1991), wherein firms replace obsolete technologies embedded in depreciating capital at a gradual pace.

Alternatively, MAX-MOST gaps may represent a persistent condition, shaped by frictions or constraints that lead establishments to systematically under-utilize the most advanced technologies available within a given function. To empirically differentiate between these hypotheses, we undertake three exercises

than one technology is used. C.2 reports the frequency of MAX-MOST gaps by business function, computed as the average of $GAP_{f,j}$ across establishments using sampling weights. The frequency of the MAX-MOST gaps by type of function (e.g., GBF, agricultural SSBF, etc) result from averaging these frequencies across the relevant BFs.

³⁰The dispersion, median and skewness of the distribution are robust to conditioning on the minimum number (or share) of functions where $ANUM > 1$.

First, we examine how MAX-MOST gaps vary with establishment age. We divide establishments into age deciles and compute the average MAX-MOST gap within each decile. [Table 6](#) (first column) shows a weak negative association between MAX-MOST gaps and establishment age across deciles (correlation of -0.29) and individual establishments (-0.01). We continue to explore lifecycle patterns dividing the sample by income level and establishment size.³¹ The first two columns of [Table 5](#) and the first three columns of [Table C.5](#) reveal no clear lifecycle pattern in MAX-MOST gaps across income groups or size classes. Only among large establishments do we observe a significant decline in the MAX-MOST gap for establishments aged 16 years or older, relative to younger counterparts.

Second, we study how the gaps vary with the time since establishments adopted advanced technologies in the function. We focus on the subsample of functions where establishments have adopted top-tier technologies which are those marked in bold in the grids. (See [subsection A.1.](#)) For these technologies, FAT collects information on the year of adoption in the business function. We then divide this sample into two groups: those where $MAX_{f,j} = MOST_{f,j}$ and those where $MAX_{f,j} > MOST_{f,j}$. For each group, we examine the distribution of the years since adopting the top-tier technology in the function. [Figure 7](#) shows that the distributions are similar, indicating that time is not a significant factor in closing the MAX-MOST gap.

Third, we split the sample of establishments that have adopted a top-tier technology between early adopters and recent adopters. Early adopters are those that adopted a top-tier technology before the median year of adoption of top-tier technologies in the specific business function and recent adopters are those that have adopted after the median year of adoption. [Figure 8](#) plots the histogram of $MOST_{f,j}$, for each of these sub-samples. Conditional on adopting a top-tier technology, the distributions of $MOST_{f,j}$ for early and late adopters are quite similar.

These exercises confirm that MAX-MOST gaps persist long after the adoption of new technologies; therefore, they are not transitory.

Correlates of MAX-MOST Gaps. Having shown that establishments consistently underutilize the most advanced available technologies in business functions, and that this pattern persists over time, we next examine the association between MAX-MOST gaps (GAP_j) and establishment characteristics other than its age.

As with age, we start by calculating the average value of GAP_j within each decile of the distribution of each characteristic across establishments (See [Table 6](#)). The characteristics we

³¹High-income countries include Korea, Croatia, and Poland; the remaining countries are classified as low-income for brevity. Establishments are grouped into three size categories: small (5–19 employees), medium (20–99), and large (100 or more workers).

consider include sales per worker, employment, human capital (fraction of college-educated workers), quality of management practices, and manager’s bias in technology assessment (difference between manager’s assessment and actual technology sophistication in the establishment).³² Additionally, we also report average MAX-MOST gaps for establishments in each category of four binary variables: exporter status, multinational status, family ownership, and whether a loan application was rejected.

There are significant differences in the average frequency of MAX-MOST gaps between establishments in the top and bottom deciles of the distribution of sales per worker (-0.16), employment (-0.08), human capital (-0.14), management practices (-0.11), and bias in the perceived technological sophistication (0.05). Across deciles, MAX-MOST gaps are strongly negatively associated with sales per worker, establishment size (log employment), and human capital; they show a weaker negative relationship with establishment age; no systematic link with the z-score of management practices; and a positive association with the bias in the managerial assessment of technological sophistication. MAX-MOST gaps are also more frequent among family-owned businesses, establishments with rejected loan applications, multinationals, and non-exporters.

We further explore the association between MAX-MOST gaps and establishment characteristics by estimating the following specification:

$$GAP_j = \alpha_s + \alpha_c + \beta_x * X_j + \beta_0 * Multiple_j + \beta_a * D_j^a + u_j, \quad (9)$$

where α_s are two-digit sector fixed effects, α_c are country fixed effects, X_j is a vector that includes most of the establishment characteristics in [Table 6](#). Additionally, we control for the number of BFs where the establishment uses multiple technologies, $Multiple_j$, and a set of dummy variables that capture the establishment age, D_j^a .

The signs of the estimated coefficients reported in the first column of [Table 7](#) are consistent with the unconditional correlations. Specifically, the fraction of functions with a MAX-MOST gap is positively associated with a loan application rejected, a lower fraction of college educated workers, a positive bias in the perceived sophistication of technology, being part of a family business or a multinational and not being an exporter. Additionally, an establishment’s age of between 11 and 15 years is associated with a lower frequency of the gap.

While partial correlations should be interpreted with caution, the estimates are consistent with a role for two types of factors in MAX-MOST gaps. The first factor reflects input constraints that limit the establishment’s ability to fully utilize advanced technologies. Es-

³²FAT asks managers to rate their establishment’s technology relative to other establishments worldwide; we convert this to a 1–5 scale and compare it to \bar{S}_j , which is the simple average of MAX_j and $MOST_j$.

establishments with limited access to skilled labor or facing financing constraints may rely more heavily on less sophisticated technologies, that are operable by less-educated workers or requiring less capital, despite having adopted more sophisticated technologies. The negative coefficient for human capital and the positive coefficient for the loan denied dummy are consistent with the role of input constraints in the MAX-MOST gap.

The second factor is managerial mistakes. In contrast to input constraints, poor management or biased perceptions of an establishment’s technological sophistication can lead to a suboptimal underutilization of already-adopted advanced technologies. Although the association between management practices scores and GAP_j is weak, the positive correlations with both bias in perceived technology and family ownership suggest a role for managerial misjudgement. The negative association with exporter status may also reflect that competitive pressure induces managers to use technologies more efficiently, narrowing the MAX-MOST gap.

In the Appendix, we examine the relevance of these two types of factors across environments, by re-estimating specification (9) on subsamples defined by establishment size and country income level (see Table 7). Scarcity of human and physical are more prominent in low-income countries while managerial mistakes proxies are more salient among large establishments, likely due to the increased complexity of managing scale.

These findings suggest that the relative importance of input constraints and managerial mistakes varies systematically across establishments and environments.

5 Technology Sophistication Across Establishments

After documenting the presence of MAX-MOST gaps in the use of technology at the BF level, we turn to analyzing technology sophistication across establishments. We proceed in three steps. First, we examine the variation in MAX and MOST across establishments. Second, we analyze how technology sophistication correlates with establishment characteristics. Third, we explore its relationship with measures of establishment performance such as productivity and profit rates. Throughout, we connect the patterns of technology sophistication we identify to relevant strands of the literature.

5.1 Cross-establishment Patterns of Technology Sophistication.

As discussed in section 3.1, establishment-level measures of sophistication are computed as the simple average of BF-level sophistication across all the BFs conducted in-house. Figure 9 plots the distributions of MAX and MOST across the establishments and Table 8 reports

key statistics.

The average values for MAX and MOST are 2.6 and 2.0, with medians of 2.5 and 1.9, respectively. What do these mean in practice? Both indicators reflect how far establishments are from the technological frontier across their business functions. Each establishment has technology scores for 7 to 13 functions, with each function rated from 1 (manual/basic) to 5 (frontier technology). The typical business function has 5 technologies. On average, establishments adopt their most advanced technologies (MAX) at a level between 2 and 3 on this scale, while their most commonly used technologies (MOST) are around level 2. This suggests that the average establishment is one notch above purely manual processes, using basic machinery, equipment, or software. For example, in business administration or production planning, a MOST score near 2 indicates reliance on standard tools like computers with spreadsheets rather than advanced systems such as ERP. In fabrication, it means using operator-controlled machines rather than computer-controlled machines or robotic systems. These patterns are consistent across sectors (agriculture, manufacturing, services) and business functions, both GBFs and SSBFs, as shown in [Table C.2](#).

Both MAX and MOST exhibit substantial dispersion across establishments, as indicated by their large standard deviation and the p80-p20 gap relative to the maximum possible range. However, the dispersion is more pronounced for MAX than for MOST. Furthermore, the distributions of MAX and MOST differ in shape—MOST displays greater skewness than MAX.

To further examine the variation in technology sophistication, we analyze the variation by countries and sectors.

Countries Country-level measures of technology sophistication are computed as the average sophistication across establishments using the sampling weights. Technology sophistication differs considerably across countries. The difference between the average sophistication in the countries with the highest and lowest levels is 1.53 for MAX and 1.01 for MOST. [Figure 10](#) illustrates the relationship between technology sophistication and per capita income, showing strong positive correlations—0.78 for MAX and 0.94 for MOST. The particularly high correlation for MOST suggests that establishments in richer countries tend to use intensively more sophisticated technologies.

Technology sophistication varies within countries even more than across countries. Measuring within-country dispersion by the p80-p20 gap in the country, we find that on average within-country dispersion is 2.17 for MAX and 1.56 for MOST. Interestingly, this gap varies considerably across countries, as shown in [Figure 11](#). The p80-p20 gap in a country is positively associated with per capita income, though the strength of the association differs: 0.33

for MAX and a much stronger 0.95 for MOST.

Sectors Technology sophistication also varies widely within sectors. Table 8 reports the dispersion of technology sophistication measures within each 1-digit sector, revealing that within-sector variation is comparable to the variation across establishments in the overall sample. Notably, both MAX and MOST exhibit greater dispersion among agricultural establishments than within manufacturing or services. This finding is potentially relevant for the agricultural productivity gap literature, as the larger dispersion in sophistication may help explain the larger dispersion in productivity observed across agricultural establishments than across non-agricultural establishments (Gollin, Lagakos and Waugh, 2014).

The large variation in productivity across establishments has led Syverson (2011) to raise an important question: Does this dispersion primarily reflect differences in the tasks performed and goods produced by each establishment, or does it stem from variation in the productivity with which tasks are conducted and goods produced? An analogous question arises in the context of technology sophistication: Does its variation reflect differences in the types of tasks and outputs across establishments, or disparities in the sophistication of the technologies used to carry out similar activities?

To examine the sources of productivity variation, Syverson (2011) showed that dispersion remains large even within narrowly defined sectors where establishments perform similar tasks and produce similar products. We extend this approach to investigate the sources of variation in technology sophistication across establishments. Table 9 reports the fraction of the cross-establishment variance in MAX_j and $MOST_j$ explained by sector dummies at different levels of aggregation (2-digit, 3-digit and 4-digit ISIC classifications). For comparison, we present the same statistic for productivity, measured by the log of sales per worker.³³

Consistent with Syverson (2011), sector-level dummies account for a small share of the variance in productivity across the establishments in the FAT dataset—14% at the 2-digit level, 20% at the 3-digit level, and 23% at the 4-digit level. Notably, these dummies explain a comparable share of the variation in technology sophistication across establishments.

This finding suggests that cross-establishment variation in technology sophistication is not primarily driven by differences in the tasks performed or the goods and services produced but rather by differences in the technologies employed to execute these activities.

³³The sample is restricted to establishments for which sales per worker data and 4-digit ISIC codes are available. Additionally, we require at least 5 establishments per 4-digit sector. The results are robust to omitting this requirement or to focusing only in sectors stratified in FAT. The resulting data includes 238 4-digit sectors, 144 3-digit sectors and 66 2-digit sectors.

5.2 Technology Sophistication and Establishment Characteristics

Having documented substantial variation in technology sophistication across establishments — between and within countries and within detailed sectors — we examine factors associated with this variation. We begin by exploring lifecycle patterns in technology sophistication and then analyze its association with other establishment characteristics.

Technology sophistication over the lifecycle We examine how MAX_j and $MOST_j$ vary with establishment age across country income levels (Table 5) and size groups (Table C.5). For MAX_j , we find gradually upward-sloping age profiles in both high- and low-income countries and across all three size categories. The age profile of $MOST_j$ is similarly upward-sloping in high-income countries and among medium and large establishments. However, in low-income countries and among small establishments, there is a sharp increase in $MOST_j$ between establishments aged 6–10 years and those aged 5 years or younger, but it remains flat thereafter.

The different lifecycle patterns of $MOST_j$ in high- and low-income are reminiscent of the slower establishment growth documented by Hsieh and Klenow (2014) in low-income countries. However, an examination of FAT establishments reveals that older firms (16 years or more) exhibit comparable average employment levels in both contexts (49.7 employees in low-income vs. 47.9 in high-income countries). The flatter age-size profile observed in low-income countries is primarily driven by the relatively larger size of younger establishments (less than 6 years old), which employ on average 19.3 workers compared to 12.4 in high-income countries. This pattern suggests that differences in establishment size are unlikely to fully account for the the flattening of MOST with age in low-income economies.

Correlates of technology sophistication We examine how technology sophistication relates to establishment characteristics beyond age and size by estimating the following specification:

$$S_j = \alpha_c + \alpha_s + \beta * X_j + u_j \quad (10)$$

where $S_j = \{MAX_j, MOST_j\}$, α_c and α_s denote country and 2-digit sector fixed effects, respectively, and X_j reflects the characteristics of the establishment including fraction of employees with college degree, quality of management practices, manager’s bias on the establishment technology sophistication, family business, exporter and multinational status, size, and age. We also estimate a version of (10) with size effects that can potentially vary between high- and low-income countries. Columns 2 through 5 of Table 7 report the estimates.

Both measures of technological sophistication are positively associated with employees' human capital, management-practice quality, exporter and multinational status, and establishment size, and negatively associated with family business status and manager's bias in perceived technology sophistication. For MAX_j , size effects are similar across high- and low-income countries. In contrast, $MOST_j$ exhibits significantly stronger size effects in high-income countries.

These findings offer new insights into several strands of the literature. In particular, the weaker size effects for $MOST_j$ in low-income countries help explain the flattening of its lifecycle profile documented earlier. Specifically, the limited expansion of sophisticated technologies beyond age 10 is partly driven by the lower sensitivity of $MOST_j$ to establishment size in low-income economies.

The positive association between exporter status and technological sophistication mirrors the well-established link between exporting and productivity (Bernard and Jensen, 1999). Whether this reflects selection into exporting or productivity gains from exporting, our findings suggest that technological sophistication may be a key channel through which firms are selected or experience the gains from exporting.³⁴

Consistent with prior evidence that better management practices enhance productivity (Bloom and Van Reenen, 2007), we find a strong association between management quality and both MAX_j and $MOST_j$. This may reflect a complementarity between technological sophistication and managerial capability (Bloom, Sadun and Van Reenen, 2012).³⁵ Separately identifying associations between technology sophistication and different proxies for managerial quality (z scores, a family business dummy, and managers' bias in perceived technology sophistication) suggests these proxies capture different dimensions of managerial quality. For example, managers with upward biases in perceived technology may be slow to adopt or extend the use of more sophisticated technologies because they (mistakenly) believe their establishments already use highly sophisticated technologies. This channel contrasts with reductions in the marginal product of sophisticated technologies from poor management practices—such as the absence of KPIs, poor inventory management, or family business governance.

Finally, although both MAX_j and $MOST_j$ are strongly associated with workforce human capital, managerial quality proxies, and international status, we often find significant differences in the magnitudes of their coefficients. For example, human capital plays a more

³⁴See Bernard and Jensen (1999) and Melitz (2003) for evidence and theory on selection into exporting, and De Loecker (2007), Lileeva and Trefler (2010), and Brandt et al. (2017) for evidence on productivity gains from exporting.

³⁵In the working paper version, we document a positive interaction between management practices and technology sophistication in productivity regressions, consistent with this hypothesis.

prominent role in adoption (MAX_j) than in subsequent extension of use ($MOST_j$). This heterogeneity suggests the relevance of these factors differs between adoption and the extension in the use of sophisticated technologies.

5.3 Technology sophistication and performance

We conclude our analysis by exploring the association between technology sophistication and establishment performance measures such as productivity and profitability. While our data do not permit causal inference, these correlations can inform about a range of key research questions, including the sources of productivity and profitability differences across establishments, the agricultural productivity gap, and the suitability of advanced technologies in low-income countries.

To explore these correlations, we estimate variations of the following regression:

$$Y_j = \alpha_c + \alpha_s + \gamma * S_j + \beta * X_j + u_j \quad (11)$$

Y_j is establishment performance. We use three measures, real sales per worker, the Solow residual, and the profit rate (i.e. profits over sales). S_j are measures of technology sophistication such as MAX_j , $MOST_j$ or \bar{S}_j defined as the simple average of MAX_j and $MOST_j$, and X_j is a vector of controls that we include when the dependent variable is sales per worker and that comprises the log of the real book value of capital per worker (K_j) and the fraction of workers with college education (H_j). α_c and α_s are country and 2-digit sector effects, respectively.³⁶

Across establishments There is a strong positive association between technology sophistication and both productivity and profit rates across establishments (See columns 1-4 in [Table 10](#) and [Table C.7](#)). The association between productivity and technology sophistication is stronger when country fixed effects are excluded, reflecting the strong association between these variables across countries.

Decomposing \bar{S}_j into MAX_j and $MOST_j$, we find a much stronger association between the measures of establishment performance –labor productivity, Solow residuals and profit rates– and $MOST_j$ than with MAX_j . This finding adds a new dimension to the differences between adoption and extension of the use of sophisticated technologies. Beyond differences in their nature, distributions across establishments, lifecycle patterns and associations with

³⁶The working paper version, (Cirera, Comin and Cruz, 2024b) shows the robustness to a broader set of controls that include the quality of management practices, establishment age and markups. We report the estimates with Solow residuals and profit rates as dependent variables and those with real sales per worker without controls in [Appendix C.2](#).

establishment characteristics, the two technology upgrading processes have different relations to productivity dynamics.

Across sectors Next, we study whether the association between technology sophistication and productivity differs across sectors. Specifically, we re-estimate the productivity regression (11) separately for each one-digit sector. The estimates, reported in columns 5 through 7 of Table 10 and in Table C.8, reveal a strong positive association between technology sophistication and productivity in all three sectors. However, the magnitude of the coefficients varies substantially, being highest in agriculture and lowest in services.

The larger coefficient in agriculture, combined with the greater cross-establishment dispersion in technology sophistication within that sector (Table 8), implies that technology sophistication accounts for a larger share of productivity variation across establishments in agriculture than in non-agricultural sectors.

This finding offers a new perspective on the so-called agricultural productivity gap (Caselli, 2005; Gollin, Lagakos and Waugh, 2014)—the observation that productivity differences between rich and poor countries are markedly larger in agriculture than in other sectors.³⁷

Across countries We conclude the analysis by exploring whether the association between technology sophistication and establishment performance – labor productivity, Solow residual and profit rate – differs across countries. Specifically, we divide the sample into high- and low-income countries, and re-estimate specification (11) allowing the coefficient of technology sophistication to vary across samples.

Column 8 of Table 10 and Table C.9 present the results. The main finding is that the coefficient of technology sophistication in both the productivity and profitability regressions is not smaller for low-income countries than for high-income ones. This result is informative in the context of the “appropriate technology” hypothesis (e.g., Basu and Weil, 1998; Acemoglu and Zilibotti, 2001), which posits that establishments in poor countries underutilize sophisticated technologies because the scarcity of human and physical capital limits the potential productivity gains these technologies embody.

In contrast, our findings indicate that using more sophisticated technologies is associated with comparable gains in productivity and profitability in both low- and high-income countries. This evidence suggests that advanced technologies are appropriate across countries regardless of income level.

³⁷In our sample, productivity ratio between establishments at the 90th and 10th percentiles of the productivity distribution is 6.75 times larger in agriculture than in services.

6 Conclusions

We conclude by summarizing our main contributions and outlining directions for future work. The paper has proposed a new approach to characterizing technology use at both the business function and establishment levels, based on three core elements. First, the design of a grid that links each key business function in an establishment to the technologies it employs to conduct its main tasks. Second, the collection of data on both the technologies used and the intensity of their use within each function. Third, leveraging a ranking of technologies by sophistication, we construct measures for the most widely used (MOST) and the most sophisticated (MAX) technologies, along with an indicator that reflects when the most sophisticated technology is not the most widely used.

Using this methodology, we have constructed the FAT dataset that covers over 21,000 establishments that constitute representative samples in 15 countries at all stages of development. Upon exploring the data set, three main findings emerge. First, the most widely used technology in a business function (MOST) is typically not the most sophisticated technology available (MAX). The gap between MAX and MOST is persistent, and its frequency varies widely across establishments and countries. Specifically, the frequency of MAX-MOST gaps is positively associated to proxies of input scarcity, and managerial mistakes and negatively associated with productivity.

Second, there are large differences across establishments and countries in MAX and MOST measures of technology sophistication. Variation in sophistication is larger in high- than in low-income countries and across agricultural than non-agricultural establishments. As with productivity, variation in sophistication is driven by differences in the technologies used to produce goods and conduct tasks in establishments rather than on the type of goods and tasks produced across establishments.

Third, technological sophistication is positively associated with establishment characteristics such as human capital, management practices, size, exporting, multi-national and multi-establishment status, and performance metrics such as productivity and profitability. The correlation with establishment performance is stronger for MOST than for MAX, for agricultural than non-agricultural establishments and is not weaker in low- than in high-income countries.

A key theme emerging from these findings is that MAX and MOST reflect fundamentally distinct technology upgrading processes. Their differences arise not only from their inherent nature but also from being shaped by different drivers, exhibiting distinct distributions across establishments, divergent lifecycle patterns, and varying associations with key performance metrics such as productivity and profitability.

Building on these methodological and empirical contributions, we envision several avenues for future research. First, we are expanding the FAT dataset to include additional countries and sectors, constructing grids that offer richer technological profiles across a broader range of establishments, and introducing a longitudinal dimension by revisiting establishments to document the evolution of their technology use.

Second, the distinct nature of MAX and MOST documented in this paper deserves scrutiny. On the theoretical side, we plan to develop frameworks that connect productivity not only to the range of technologies used but also the intensity of their use, and that help explain the observed gap between MAX and MOST. Empirically, it is essential to investigate the drivers of MAX-MOST gaps.

Third, we intend to study the variation in technology sophistication within establishments (i.e. across the business function of an establishment). FAT is an ideal dataset to study the magnitude and drivers of the within-establishment variation in technology sophistication because its unit of observation is the business function. One by-product of this analysis is the construction of establishment-level measures of technology sophistication consistent with the aggregators that best explain the cross-function variation in technology sophistication.

Fourth, we have intentionally avoided studying the range of business functions conducted in establishments. This topic, along with developing models that rationalize the relationship between technology sophistication and the limits of the establishment and its specialization from a task perspective, warrants separate attention.

References

- Acemoglu, Daron, and David Autor.** 2011. “Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings.” In . Vol. 4 of *Handbook of Labor Economics*, , ed. David Card and Orley Ashenfelter, 1043 – 1171. Elsevier.
- Acemoglu, Daron, and Fabrizio Zilibotti.** 2001. “Productivity Differences.” *The Quarterly Journal of Economics*, 116(2): 563–606.
- Acemoglu, Daron, and Pascual Restrepo.** 2018. “The race between man and machine: Implications of technology for growth, factor shares, and employment.” *American economic review*, 108(6): 1488–1542.
- Acemoglu, Daron, Gary W Anderson, David N Beede, Cathy Buffington, Eric E Childress, Emin Dinlersoz, Lucia S Foster, Nathan Goldschlag, John C Haltiwanger, Zachary Kroff, et al.** 2022. “Automation and the workforce: A firm-level view from the 2019 Annual Business Survey.” *NBER Working Paper No. 30659*.
- Akerman, Anders, Ingvil Gaarder, and Magne Mogstad.** 2015. “The Skill Complementarity of Broadband Internet.” *The Quarterly Journal of Economics*, 130(4): 1781–182.
- Autor, David H, and David Dorn.** 2013. “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market.” *American Economic Review*, 103(5): 1553–97.
- Basu, Susanto, and David N. Weil.** 1998. “Appropriate Technology and Growth.” *The Quarterly Journal of Economics*, 113(4): 1025–1054.
- Battisti, Giuliana, and Paul Stoneman.** 2003. “Inter-and intra-firm effects in the diffusion of new process technology.” *Research Policy*, 32(9): 1641–1655.
- Battisti, Giuliana, Heinz Hollenstein, Paul Stoneman, and Martin Woerter.** 2007. “Inter And Intra Firm Diffusion Of Ict In The United Kingdom (Uk) And Switzerland (Ch) An Internationally Comparative Study Based On Firm-Level Data.” *Economics of Innovation and New Technology*, 16(8): 669–687.
- Behaghel, Luc, Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon.** 2015. “Please Call Again: Correcting Nonresponse Bias in Treatment Effect Models.” *The Review of Economics and Statistics*, 97(5): 1070–1080.
- Benhabib, Jess, and Aldo Rustichini.** 1991. “Vintage capital, investment, and growth.” *Journal of Economic Theory*, 55: 323–339.
- Bernard, Andrew B, and J Bradford Jensen.** 1999. “Exceptional exporter performance: cause, effect, or both?” *Journal of international economics*, 47(1): 1–25.
- Bertrand, Marianne, and Sendhil Mullainathan.** 2001. “Do People Mean What They Say? Implications for Subjective Survey Data.” *American Economic Review*, 91(2): 67–72.
- Bloom, Nicholas, and John Van Reenen.** 2007. “Measuring and Explaining Man-

- agement Practices Across Firms and Countries.” *The Quarterly Journal of Economics*, 122(4): 1351–1408.
- Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen.** 2019. “What Drives Differences in Management Practices?” *American Economic Review*, 109(5): 1648–1683.
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen.** 2012. “Americans do IT better: US multinationals and the productivity miracle.” *American Economic Review*, 102(1): 167–201.
- Bloom, Nicholas, Renata Lemos, Raffella Sadun, Daniela Scur, and John Van Reenen.** 2016. “International Data on Measuring Management Practices.” *American Economic Review*, 106(5): 152–56.
- Bond, Timothy N, and Kevin Lang.** 2019. “The sad truth about happiness scales.” *Journal of Political Economy*, 127(4): 1629–1640.
- Boothby, Daniel, Anik Dufour, and Jianmin Tang.** 2010. “Technology adoption, training and productivity performance.” *Research Policy*, 39(5): 650–661. Special Section on Government as Entrepreneur.
- Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang.** 2017. “WTO accession and performance of Chinese manufacturing firms.” *American Economic Review*, 107(9): 2784–2820.
- Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt.** 2002. “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence.” *The Quarterly Journal of Economics*, 117(1): 339–376.
- Brynjolfsson, Erik, and Lorin M. Hitt.** 2000. “Beyond Computation: Information Technology, Organizational Transformation and Business Performance.” *Journal of Economic Perspectives*, 14(4): 23–48.
- Caeyers, Bet, Neil Chalmers, and Joachim De Weerd.** 2012. “Improving Consumption Measurement and Other Survey Data through CAPI: Evidence from A Randomized Experiment.” *Journal of Development Economics*, 98(1): 19–33.
- Caselli, Francesco.** 2005. “Accounting for Cross-country Income Differences.” *Handbook of Economic Growth*, 1: 679–741.
- Chung, Sunghoon, and Minho Kim.** 2021. “How smart is a ‘smart factory’?: Drivers and effects of factory smartization.”
- Cirera, Xavier, Diego A Comin, and Marcio Cruz.** 2022. *Bridging the technological divide: Technology adoption by firms in developing countries*. World Bank Publications.
- Cirera, Xavier, Diego A Comin, and Marcio Cruz.** 2024a. “Anatomy of Technology and Tasks in the Establishment.” National Bureau of Economic Research Working Paper

32281.

- Cirera, Xavier, Diego A Comin, Marcio Cruz, and Kyung Min Lee.** 2020. “Technology within and across Firms.” *Policy Research Working Paper, n. 9476. The World Bank.*
- Cirera, Xavier, Diego A Comin, Marcio Cruz, and Kyung Min Lee.** 2021. “Anatomy of Technology in the Firm.” *NBER Working Paper Series, No. w28080. National Bureau of Economic Research.*
- Cirera, Xavier, Diego Comin, and Marcio Cruz.** 2024b. “TECHnological Factor Productivity.” *Mimeo.*
- Coase, R. H.** 1937. “The Nature of the Firm.” *Economica*, 4(16): 386–405.
- Collins, Debbie.** 2003. “Pretesting Survey Instruments: An Overview of Cognitive Methods.” *Quality of Life Research*, 12(3): 229–238.
- Comin, Diego, and Martí Mestieri.** 2018. “If Technology Has Arrived Everywhere, Why Has Income Diverged?” *American Economic Journal: Macroeconomics*, 10(3): 137–78.
- Cunha, Flavio, James J Heckman, and Susanne M Schennach.** 2010. “Estimating the technology of cognitive and noncognitive skill formation.” *Econometrica*, 78(3): 883–931.
- Davies, Stephen.** 1979. *The Diffusion of Process Innovations.* Cambridge University Press.
- De Loecker, Jan.** 2007. “Do exports generate higher productivity? Evidence from Slovenia.” *Journal of international economics*, 73(1): 69–98.
- Dunne, Timothy.** 1994. “Plant Age and Technology use in U.S. Manufacturing Industries.” *The RAND Journal of Economics*, 25(3): 488–499.
- Gary, Pike R.** 2007. “Adjusting for Nonresponse in Surveys.” In *Higher Education: Handbook of Theory and Research.* 411–449. Springer.
- Gollin, Douglas, David Lagakos, and Michael E Waugh.** 2014. “The agricultural productivity gap.” *The Quarterly Journal of Economics*, 129(2): 939–993.
- Griliches, Zvi.** 1957. “Hybrid Corn: An Exploration in the Economics of Technological Change.” *Econometrica*, 25(4): 501–522.
- Grossman, Gene M, and Esteban Rossi-Hansberg.** 2008. “Trading tasks: A simple theory of offshoring.” *American Economic Review*, 98(5): 1978–1997.
- Hollenstein, Heinz, and Martin Woerter.** 2008. “Inter- and intra-firm diffusion of technology: The example of E-commerce: An analysis based on Swiss firm-level data.” *Research Policy*, 37(3): 545–564.
- Hsieh, Chang-Tai, and Peter J. Klenow.** 2014. “The Life Cycle of Plants in India and Mexico *.” *The Quarterly Journal of Economics*, 129(3): 1035–1084.
- Lileeva, Alla, and Daniel Trefler.** 2010. “Improved access to foreign markets raises plant-level productivity... for some plants.” *The Quarterly journal of economics*, 125(3): 1051–

1099.

- Mansfield, Edwin.** 1963. "Intrafirm Rates of Diffusion of an Innovation." *The Review of Economics and Statistics*, 348–359.
- Melitz, Marc J.** 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *econometrica*, 71(6): 1695–1725.
- Ryan, B., and N. Gross.** 1943. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology*, 8(1): 15–24.
- Särndal, Carl-Erik, Bengt Swensson, and Jan Wretman.** 1992. *Model Assisted Survey Sampling*. Springer Science & Business Media.
- Stiroh, Kevin J.** 2002. "Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?" *American Economic Review*, 92(5): 1559–1576.
- Syverson, Chad.** 2011. "What Determines Productivity?" *Journal of Economic Literature*, 2(49): 326–365.
- Trajtenberg, Manuel.** 1990. *Economic Analysis of Product Innovation: The Case of CT Scanners*. Vol. 160, Harvard University Press.
- Ulyssea, Gabriel.** 2018. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review*, 108(8): 2015–47.
- Williamson, Oliver E.** 1979. "Transaction-Cost Economics: The Governance of Contractual Relations." *The Journal of Law and Economics*, 22(2): 233–261.
- Williamson, Oliver E.** 1988. "Technology and transaction cost economics: A reply." *Journal of Economic Behavior Organization*, 10(3): 355–363.
- Wood, Lexi.** 2024. "How Much Does ERP Cost?" <https://softwareconnect.com/learn/erp-pricing/>.
- World Bank.** 2009. "Enterprise Survey and Indicator Surveys: Sampling Methodology." *The World Bank, Manuscript*.
- Zeira, Joseph.** 1998. "Workers, machines, and economic growth." *The Quarterly Journal of Economics*, 113(4): 1091–1117.

Figures and Tables

Figure 1: General Business Functions and Their Technologies

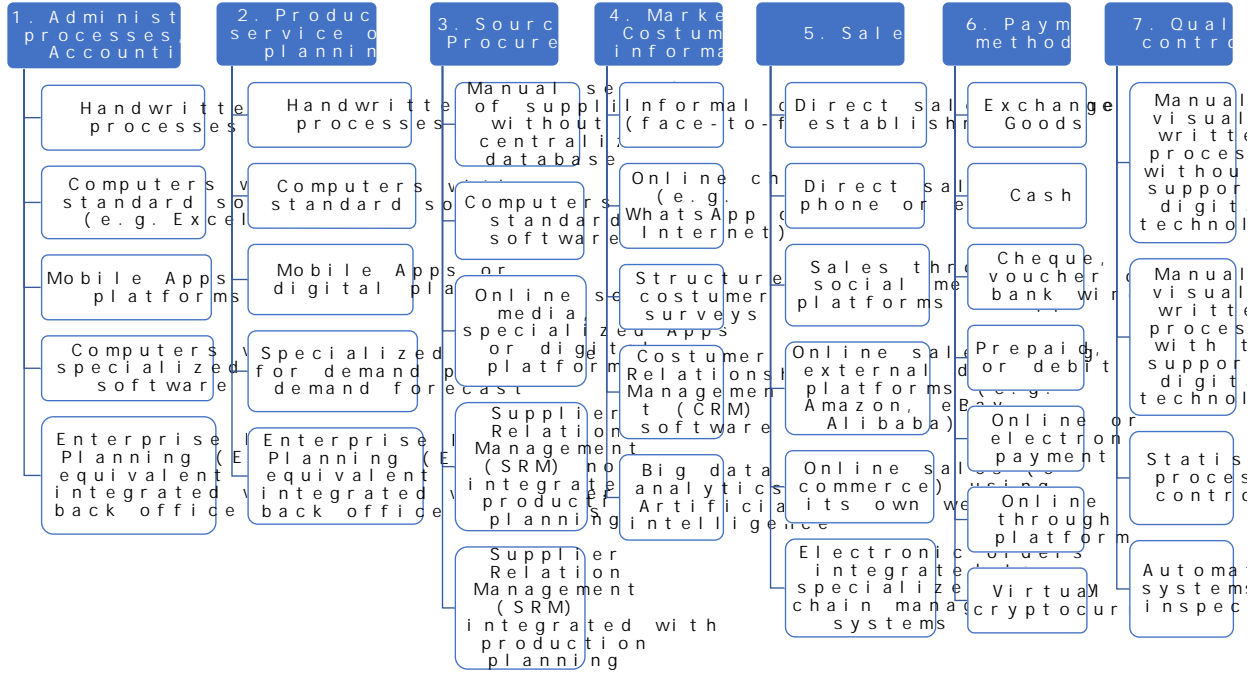


Figure 2: Sector Specific Business Functions and Technologies in Agriculture

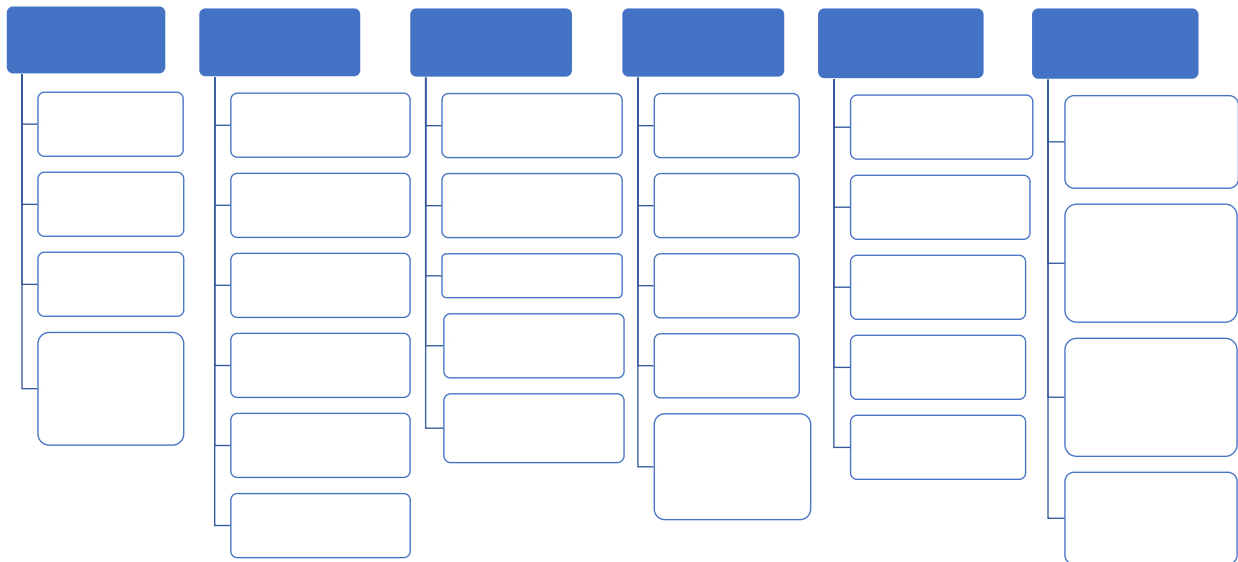
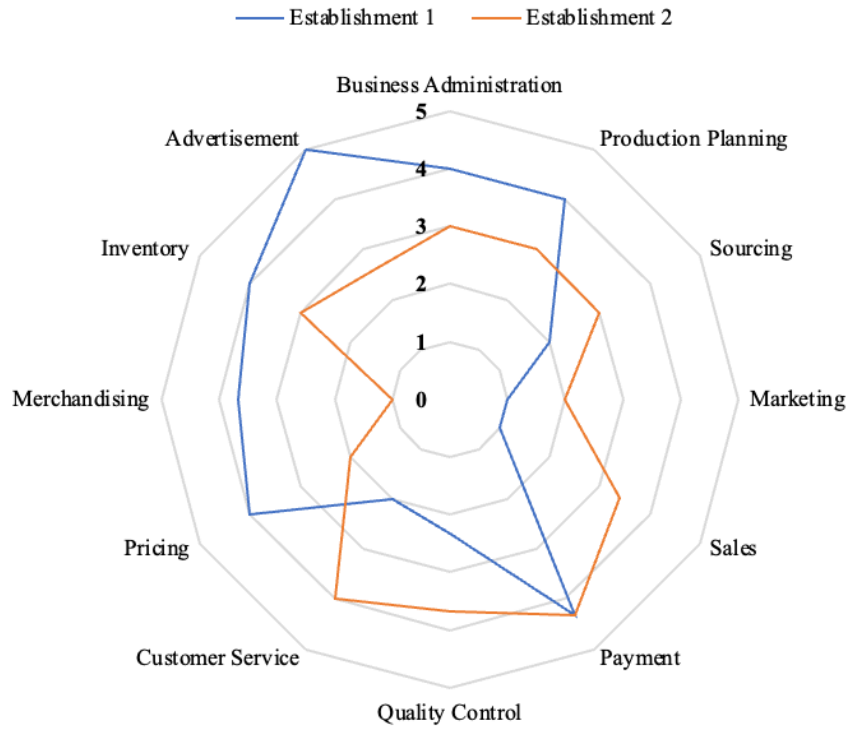
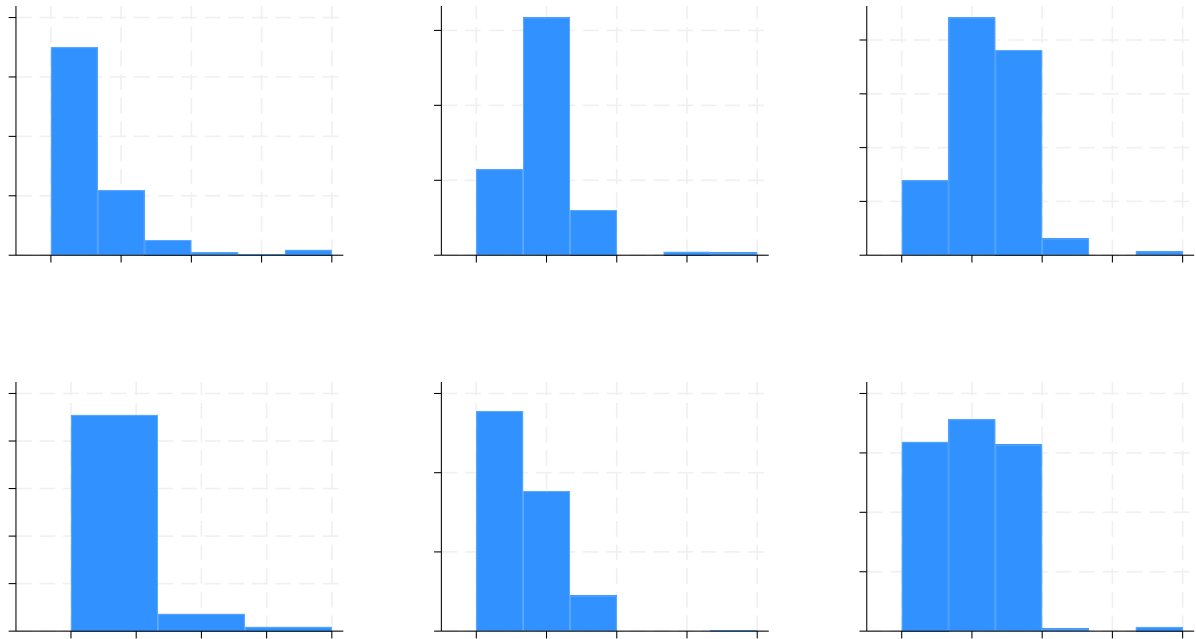


Figure 3: $MAX_{f,j}$ in two establishments in retail services



Note: Figure displays the technology index $MAX_{f,j}$ across all business functions for two individual establishments in retail services.

Figure 4: Distribution of technology sophistication in Food Processing (Fabrication)



Note: Figure displays the distribution of the technology measures $MAX_{f,j}$ and $MOST_{f,j}$ for the fabrication function across establishments in the food processing manufacturing sector in Senegal, India, and Korea. Each column in the histograms corresponds to one technology. From least to most sophisticated these are: (i) manual processes, (ii) machines controlled by operators, (iii) machines controlled by computers, (iv) robots, (v) additive manufacturing including rapid prototyping and 3D printers, and (vi) other advanced manufacturing processes such as plasma sputtering, high speed machine, E-beam, and micromachining.

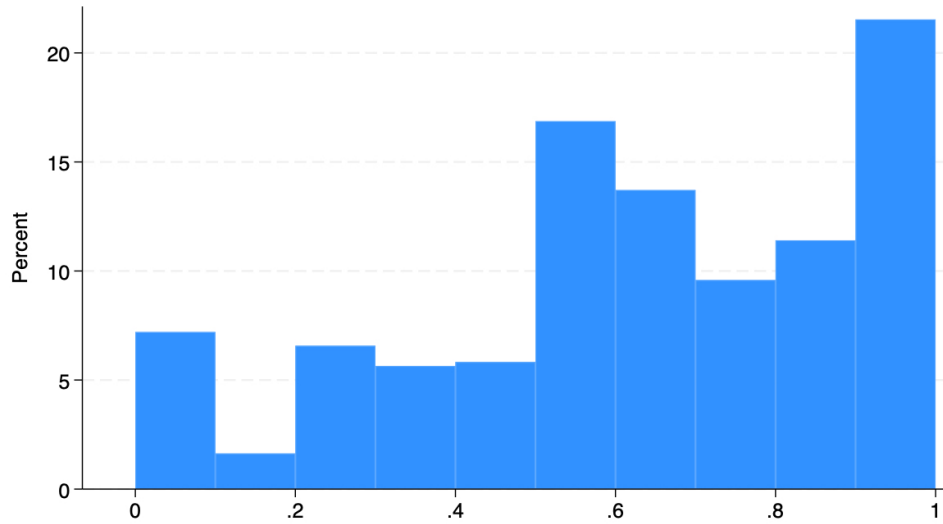
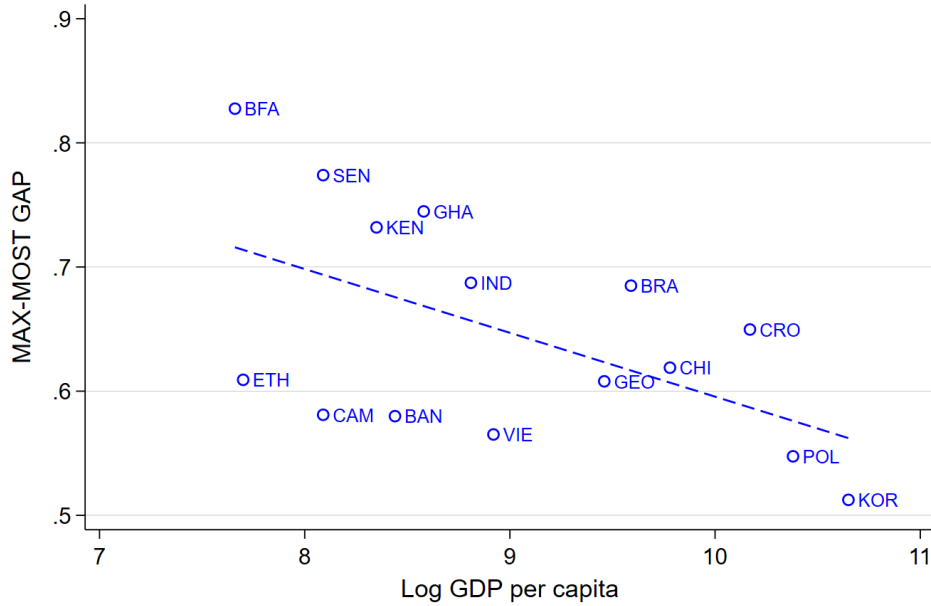


Figure 5: Distribution of MAX-MOST gap across establishments

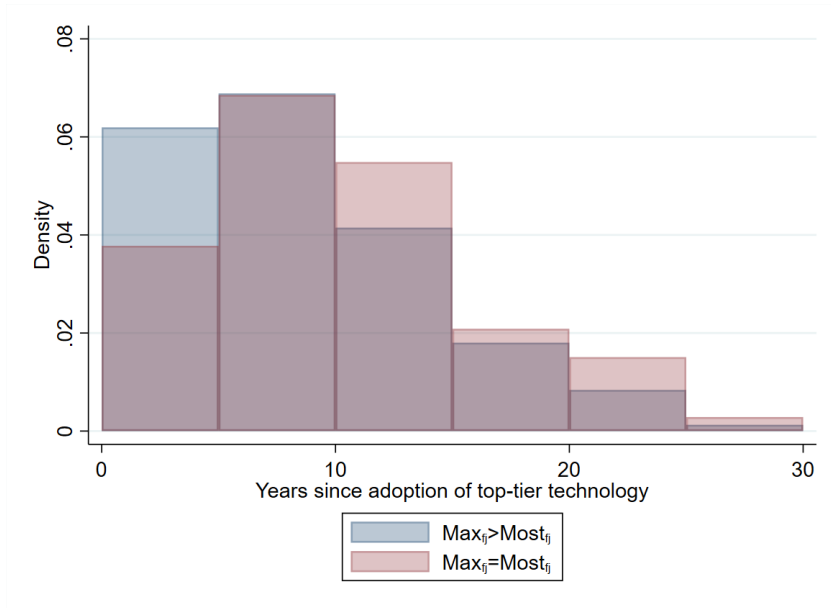
Note: MAX-MOST gap is the number of BFs in the establishment with $MAX_{f,j} > MOST_{f,j}$ relative to the number of functions where $NUM_{f,j} > 1$. Establishments are weighted by sampling weights.

Figure 6: MAX-MOST GAP across countries



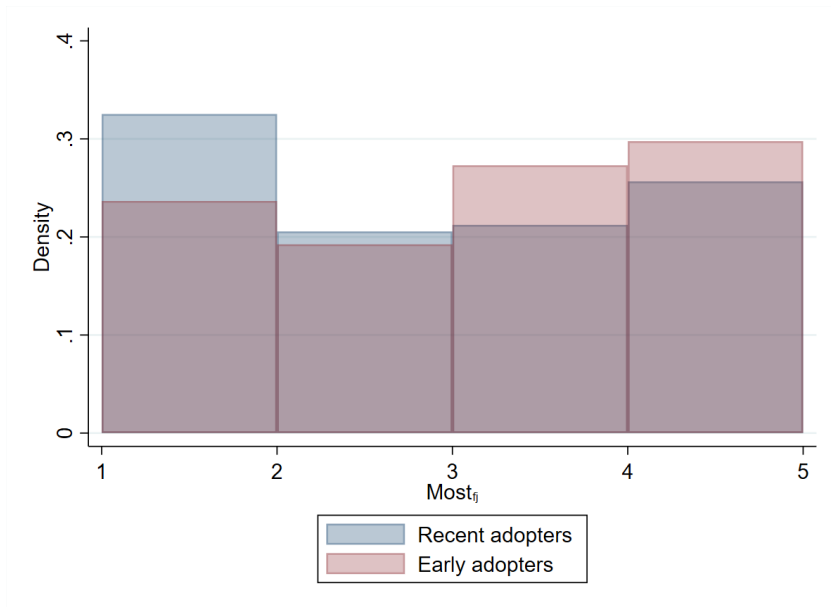
Note: The country-level MAX-MOST Gap is the average MAX-MOST Gap across the establishments in the country.

Figure 7: Distribution of Years since Adoption of Top-tier Technologies Conditional on MAX-MOST Gap



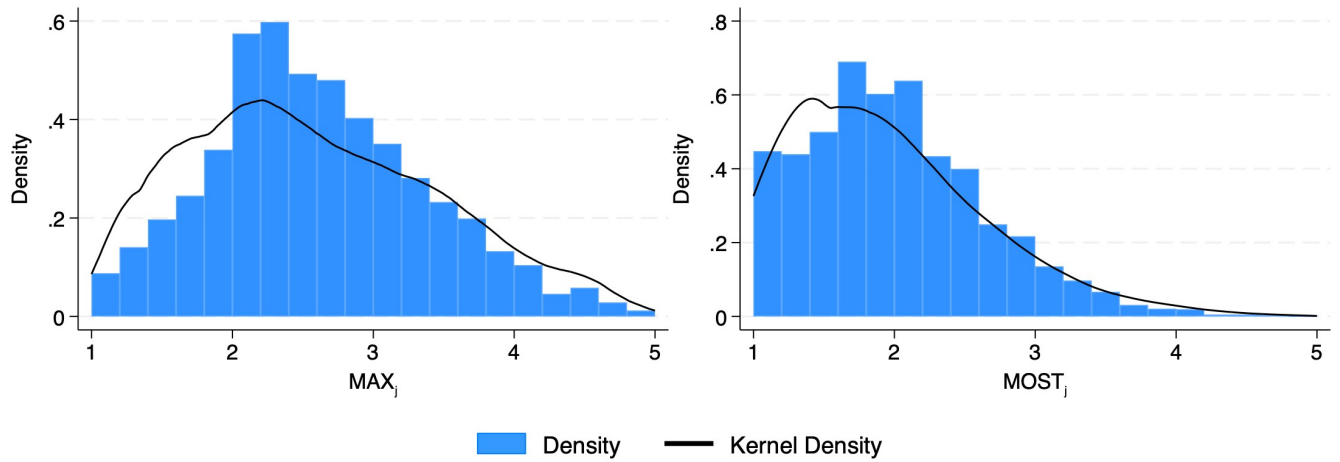
Note: Top-tier technologies are listed in [Appendix A](#). Only functions where the establishment uses multiple technologies are considered. MAX-MOST gap is present in the business function if $MAX_{fj} > MOST_{fj}$. It is absent otherwise.

Figure 8: Distribution of $MOST_{f,j}$ conditional on adopting Top-tier technology



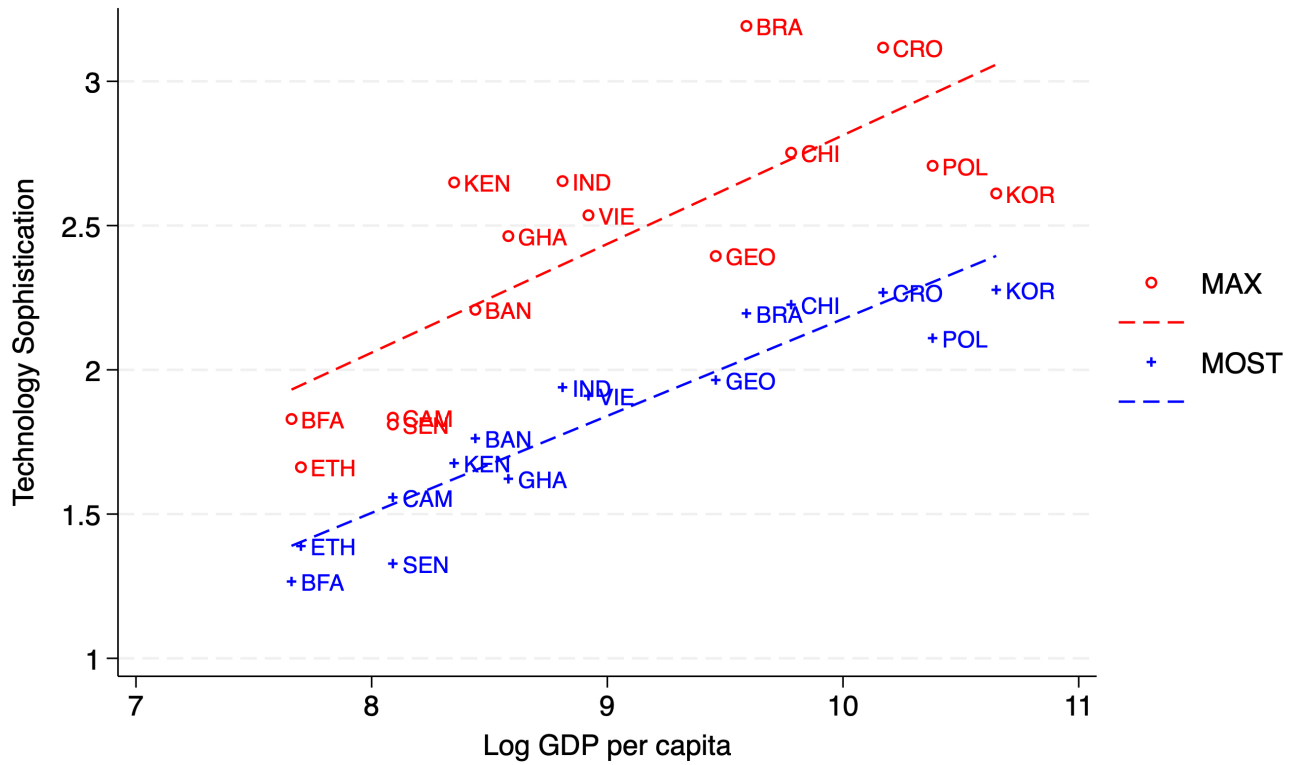
Note: Top-tier technologies are listed in [Appendix A](#). Only functions where the establishment uses multiple technologies are considered. Recent adopters are establishments that have adopted top-tier technologies after the median adoption year for top-tier technologies in the specific business function.

Figure 9: Distribution of MAX_j and $MOST_j$ across establishments



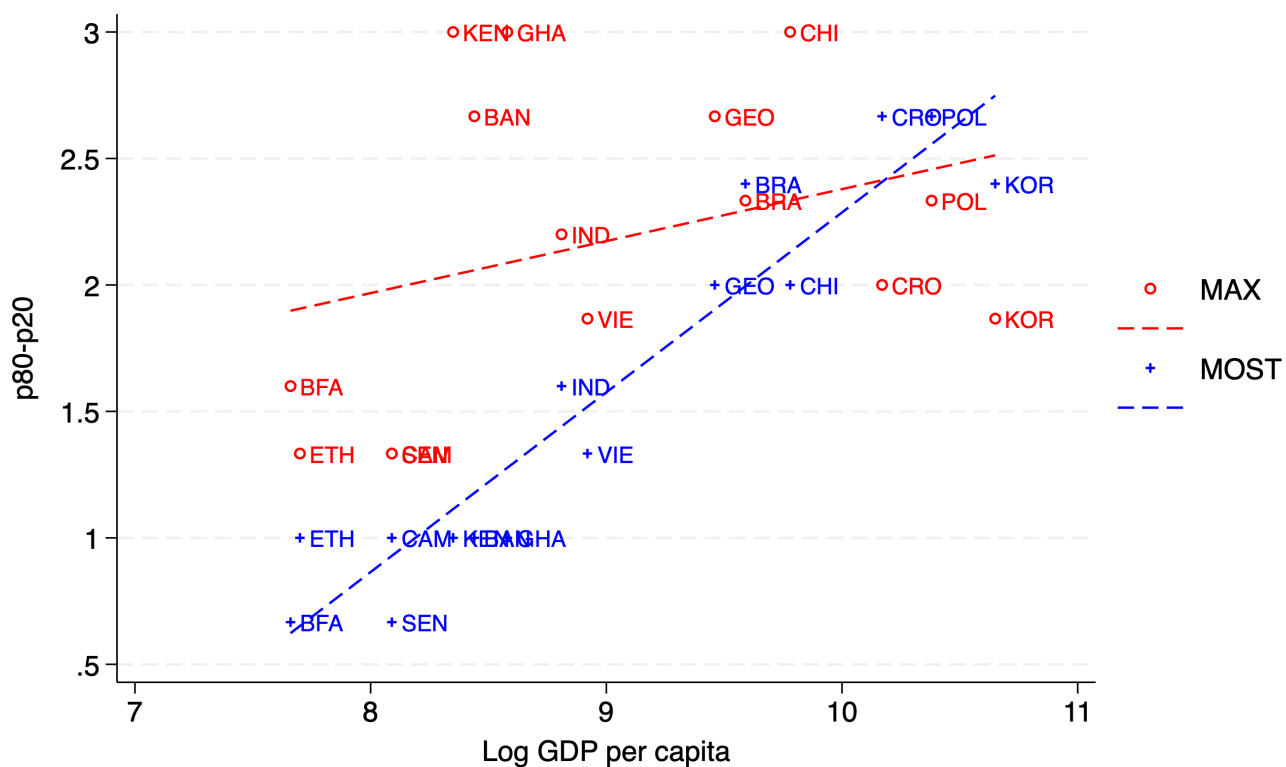
Notes : This figure represents the distribution of sophistication measures at the establishment level - MAX_j and $MOST_j$. The histogram and kernel density curves are calculated using establishment-level sampling weights.

Figure 10: MAX and MOST across countries



Note: Country-level MAX and MOST are, respectively, the weighted averages of establishment-level MAX_j and $MOST_j$, where the weights are sampling weights.

Figure 11: Within-country dispersion in technology sophistication and per capita income



Note: P80-p20 is the difference between the technology sophistication (MAX or MOST) of establishments in the 80th and 20th percentile of technology sophistication in the country. Percentiles are computed using establishment weights.

Table 1: Comparison of Technology Categories: Business Administration

	Handwritten process	Standard Software	Mobile apps	Specialized Software	Enterprise Resource Planning (ERP)
Functionality	Basic manual tasks (e.g., simple book-keeping, and employee records).	Handles financial, accounting, and HR record, with manual inputs or built in functions.	Pre-designed to handle financial, accounting, and HR record. Limited scale and customization.	Extensive specialized tools for complex financial, accounting, and HR management.	Comprehensive management of all finance, accounting, and HR processes.
Integration	No.	Limited. It requires manual processing or additional templates.	Good integration capabilities with limited customization.	Integration with systems and customizable reporting tools.	Full integration with a wide range of functions and customization.
Automation	No.	Limited. It requires manual scripting.	Good automation for specific processes with limited scale.	High-level of automation within their specialized domains.	High-level of automation across all functions.
Experts ranking	1	2	3	4	5
Reason for Ranking	Manual processes with minimal functionality, no automation, and no integration capabilities.	Basic functionality and some level of automation. Limited integration capabilities.	Good functionality, integration, and automation, but limited scale and customization.	High functionality, integration, and automation capabilities within specialized domains.	Comprehensive functionality, full integration, and advanced automation.
Technology Example	Paper Ledger	Microsoft Excel, Google Sheets	QuickBooks Online, BambooHR	Oracle Financials, Intuit QuickBooks (Desktop), Workday HCM	SAP ERP, Oracle NetSuite
Cost for acquiring the technology	Negligible	Microsoft Excel: \$159.99; Google Sheets (Free-\$18/user/month)	QuickBooks Online: \$30-\$200/month; BambooHR: \$108/month for 20 employees	Oracle Financials: \$600+/month; Intuit QuickBooks: \$1,481+/year/user;	SAP ERP; Oracle NetSuite price varies. Average ERP \$1,740-\$9,330/month.
Launch year	Pre- 1900	Microsoft Excel: 1985; Google Sheets: 2006	QuickBooks Online: 2001; BambooHR: 2008	Oracle Financials: 1989; Intuit QuickBooks: 1998; Workday HCM: 2006	SAP ERP: 1981; Oracle NetSuite: 1998
ChatGPT ranking	1	2	3	4	5

Sources: Product description on the websites of various companies, including Microsoft, Google, QuickBooks, Bamboo HR, Oracle, SAP, and Workday. Wood (2024) provides estimates of average costs for ERP software. The prompts for ChatGPT ranking are available in the appendix.

Table 2: Number of establishments in FAT by country, sector and size

	Total	Sector			Size		
		Agri.	Manu.	Serv.	Small	Medium	Large
Bangladesh	903	-	744	159	361	232	310
Brazil*	1531	96	726	709	690	563	278
BurkinaFaso	600	80	142	378	335	187	78
Cambodia	794	-	333	461	583	142	68
Chile	1095	44	321	730	545	390	160
Croatia	710	46	272	392	472	183	55
Ethiopia	1476	149	747	580	999	330	147
Georgia	1800	196	768	836	741	632	427
Ghana	1262	85	350	827	774	382	106
India**	3242	101	1841	1300	1822	912	508
Kenya	1305	155	438	712	499	421	385
Korea	1551	128	658	765	656	569	326
Poland	1500	90	624	786	779	394	327
Senegal	1786	204	679	903	1219	395	172
Vietnam	1499	110	806	583	774	426	299
Total	21055	1485	9449	10121	11249	6158	3646

Note : * Brazil refers to state of Ceará; ** States of Tamil Nadu, Uttar Pradesh, Gujarat, and Maharashtra in India. The survey does not cover agriculture or services in Bangladesh, nor agriculture in Cambodia. In India, only the states of Gujarat and Maharashtra have agriculture included in the survey.

Table 3: Comparison of baseline and Q-measures of technology sophistication

	MAX_j	$MOST_j$	H_j	$\log(\text{Sales per Worker})$
MAX_j^Q	0.90		0.08	0.22
MAX_j			0.11	0.25
$MOST_j^Q$		0.91	0.09	0.26
$MOST_j$			0.13	0.29

Note: MAX_j^Q and $MOST_j^Q$ are sophistication measures constructed with independent estimates of relative productivity across technologies in conducting a representative task in the business function. See Section A.3. All variables are residualized at 2-digit sector. All correlations use sampling weights.

Table 4: Relationship between technology measures

	$MOST_{f,j}$	$MAX_{f,j}$	$MOST_{f,j}$
$MAX_{f,j}$	0.55*** (0.01)		
$NUM_{f,j}$		0.84*** (0.01)	0.25*** (0.01)
N	187497	187497	187497
R-squared	0.66	0.75	0.50
BF FE	Y	Y	Y
Firm FE	Y	Y	Y
Variation Explained	0.34	0.47	0.05

Notes : This table reports the regression estimates of specifications 7, 6, and 8. To compute the last row, we first residualize the dependent and independent variables by regressing them on the fixed effects, and then we regress the residuals of the dependent on those of the independent. The reported number is the corresponding R^2 . Regressions are estimated using establishment-level sampling weights. Standard errors are clustered at the establishment level. *, ** and *** denote 10%, 5% and 1% significance respectively. *, ** and *** denote 10%, 5% and 1% significance respectively.

Table 5: Average GAP, MAX, and MOST, by age and country income groups

Age Group	GAP		MAX		MOST	
	Poor	Rich	Poor	Rich	Poor	Rich
<=5	0.67	0.54	2.31	2.50	1.71	2.11
6-10	0.66	0.56	2.54	2.61	1.93	2.19
11-15	0.64	0.53	2.60	2.60	1.97	2.20
>=16	0.68	0.52	2.78	2.67	1.96	2.25

Notes : Averages of GAP_j , MAX_j and $MOST_j$ across establishments in specific age and country groups. Averages computed using establishment-level sampling weights.

Table 6: Average MAX-MOST gap across establishment distribution across observable characteristics

Decile	Age	Productivity	Size	H_j	Manag.	Bias
1	0.62	0.74	0.66	0.69	0.70	0.63
2	0.65	0.70	0.65	0.68	0.60	0.56
3	0.62	0.70	0.66	0.63	0.52	0.61
4	0.64	0.62	0.65	0.62	0.62	0.64
5	0.61	0.66	0.64	0.63	0.66	0.58
6	0.59	0.61	0.61	0.61	0.58	0.63
7	0.65	0.62	0.61	0.59	0.71	0.62
8	0.64	0.61	0.59	0.61	0.60	0.63
9	0.59	0.56	0.58	0.57	0.65	0.62
10	0.62	0.58	0.58	0.56	0.59	0.68
P10-P1	0.00	- 0.16	- 0.08	- 0.14	- 0.11	0.05
Overall Corr.	- 0.01	-0.11	- 0.10	- 0.12	- 0.01	0.06
Between Corr.	- 0.27	- 0.92	- 0.80	- 0.85	- 0.03	0.55
	Exporter	MNC	Family	Loan Reject.		
1	0.53	0.69	0.63	0.67		
0	0.63	0.60	0.56	0.60		
D(1)-D(0)	- 0.10	0.09	0.07	0.07		
Overall Corr.	- 0.11	0.13	0.07	0.11		

Notes: This table displays the average MAX-MOST gap across the establishments in a decile of the distribution of each variable in the columns. Productivity is log sales per worker. Employment is the total number of workers. Human capital is the share of workers with a college degree. Manag. is the z-score for management practices. Age is the age of the establishment. Exporter, MNC, Family and Loan Reject. are binary variables that take the value of 1 if establishment is exporter, part of a multinational corporation, a family business and if a loan application has been rejected. P10-P1 is the difference between the average in the tenth and first deciles. Overall Corr is the correlation between GAP_j and the variable in the column across establishments. Between Corr is the correlation across deciles of GAP_j and the variable in the column. All statistics computed using establishment sampling weights.

Table 7: Technological sophistication and establishment characteristics

	GAP_j	MAX_j	$MOST_j$	MAX_j	$MOST_j$
	(1)	(2)	(3)	(4)	(5)
H_j	-0.07** (0.03)	0.42*** (0.04)	0.26*** (0.04)	0.42*** (0.04)	0.27*** (0.04)
Management (Z-Score)	-0.00 (0.01)	0.14*** (0.01)	0.11*** (0.01)	0.14*** (0.01)	0.11*** (0.01)
Family Business	0.04** (0.02)	-0.10*** (0.04)	-0.23*** (0.04)	-0.10*** (0.04)	-0.23*** (0.04)
Foreign owned	0.05*** (0.01)	0.27*** (0.03)	0.23*** (0.03)	0.28*** (0.03)	0.22*** (0.03)
Exporter	-0.05*** (0.02)	0.20*** (0.04)	0.15*** (0.03)	0.20*** (0.04)	0.15*** (0.03)
Bias	0.01** (0.01)	-0.12*** (0.01)	-0.09*** (0.01)	-0.12*** (0.01)	-0.09*** (0.01)
Age: 6 - 10 Years	-0.00 (0.02)	0.08** (0.03)	0.11*** (0.03)	0.08** (0.03)	0.11*** (0.03)
Age : 11 - 15 Years	-0.05** (0.02)	0.03 (0.04)	0.13*** (0.03)	0.03 (0.04)	0.13*** (0.03)
Age : 16+ years	-0.03 (0.02)	-0.00 (0.04)	0.02 (0.03)	-0.00 (0.04)	0.02 (0.03)
D(Rejected Loan)	0.04*** (0.01)				
Size : Medium		0.32*** (0.03)	0.21*** (0.02)	0.34*** (0.04)	0.19*** (0.03)
Size : Large		0.69*** (0.04)	0.47*** (0.05)	0.69*** (0.05)	0.43*** (0.05)
Size : Medium \times Rich Country				-0.09 (0.06)	0.07 (0.06)
Size : Large \times Rich Country				0.01 (0.11)	0.19* (0.11)
N	15685	16518	16518	16518	16518
R-squared	0.12	0.48	0.40	0.48	0.40
2-Dig. Sector FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

Note: *, ** and *** denote 10%, 5% and 1% significance respectively. GAP_j is the share of functions f in establishment j with multiple technologies and $MAX_{f,j} > MOST_{f,j}$. The first column controls for $Multiple_j$, which is the number of functions where multiple technologies are used in the establishment.

Table 8: Cross-establishment distribution of MAX and MOST

Sector	MAX _j				
	Mean	SD	P50	P80-P20	Skewness
Overall	2.61	0.76	2.52	1.26	0.41
Agriculture	2.63	0.89	2.51	1.92	0.39
Manufacturing	2.61	0.69	2.53	1.16	0.51
Services	2.61	0.79	2.51	1.33	0.38
Sector	MOST _j				
	Mean	Stdev	P50	P80-P20	Skewness
Overall	2.01	0.63	1.93	1.06	0.77
Agriculture	2.07	0.68	2.03	1.31	0.26
Manufacturing	2.01	0.62	1.92	1.05	0.76
Services	2.02	0.64	1.93	1.05	0.78

Notes: Mean is the average of establishment-level MAX_j or $MOST_j$. SD is the standard deviation. P50 is the median. P80-P20 is the difference between the technology sophistication of establishments in the 80th and 20th percentiles of technology sophistication. All statistics computed using establishment sampling weights. Overall refers to the entire sample, Agriculture, Manufacturing and Services restrict the sample to establishments in the 1-digit sector.

Table 9: Percentage of cross-establishment variance accounted for by sector dummies

	Productivity	MAX _j	MOST _j
2-digit	0.14	0.16	0.13
3-digit	0.20	0.20	0.19
4-digit	0.23	0.24	0.24

Notes: This table displays the share of the cross-establishment variance of log sales per worker, MAX_j and $MOST_j$ accounted for a full set of sector dummies. Sectors are defined using the International Standard Industrial Classification of All Economic Activities (ISIC), revision 4. Rows refer to the level of aggregation of sectoral fixed effects: division (2-digit), groups (3-digit), or classes (4-digit). At the 4-digit level, we kept sectors for which there were at least 5 establishments in the sample.

Table 10: Productivity and Sophistication

	ln(Sales per worker)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
K_j	0.24*** (0.03)	0.27*** (0.03)	0.24*** (0.03)	0.27*** (0.03)	0.43*** (0.04)	0.27*** (0.06)	0.27*** (0.04)	0.24*** (0.03)
H_j	0.20* (0.11)	0.67*** (0.12)	0.25** (0.11)	0.71*** (0.12)	1.13*** (0.43)	0.72*** (0.20)	0.71*** (0.14)	0.21** (0.11)
\bar{S}_j	0.50*** (0.05)	0.62*** (0.05)			0.99*** (0.13)	0.66*** (0.07)	0.56*** (0.07)	0.52*** (0.05)
$MOST_j$			0.46*** (0.06)	0.76*** (0.06)				
MAX_j			0.07 (0.05)	-0.06 (0.06)				
\bar{S}_j * D(Rich Country)								-0.07 (0.11)
Constant	6.02*** (0.31)	5.81*** (0.37)	6.01*** (0.31)	5.85*** (0.36)	3.19*** (0.31)	6.44*** (0.64)	7.09*** (0.39)	5.99*** (0.31)
N	12677	12677	12677	12677	792	6292	5593	12677
R-squared	0.42	0.24	0.42	0.26	0.58	0.33	0.21	0.42
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	No	Yes	No	No	No	No	No
Sample	All	All	All	All	Agri.	Manu.	Serv.	All

Note: \bar{S}_j is the simple average of MAX_j and $MOST_j$. $D(RichCountry)$ is a dummy variable that takes the value of 1 if country is high-income (South Korea, Poland and Croatia), and 0 otherwise. Agri. refers to agriculture sector, Manu. refers to manufacturing sector, Serv. refers to the services sector. *, ** and *** denote 10%, 5% and 1% significance respectively.