TECHnological Factor Productivity*

Xavier Cirera
World Bank

Diego Comin
Dartmouth College

Marcio Cruz
IFC

July 12, 2023

Abstract

We construct a two-dimensional grid that covers the key business functions and technologies that are used to conduct the main tasks associated with each function. This grid examines 303 technologies, spanning 7 general business functions and 56 sector-specific business functions for 12 broad sectors. We populate the grid with data from over 12,000 establishments which constitute representative samples from 10 countries. We use this data to construct measures of the sophistication of the technologies used in each business function of each establishment. We document that the variance of technological sophistication is larger within establishments than across establishments and develop a model to study the sources of variation in technological sophistication across functions. The model predicts a cross-establishment relationship between technological sophistication in a business function and an establishment-level index that aggregates the technology choices of the establishment across functions. This relationship is the technology curve; the index is Technological Factor Productivity (TechFP). We develop and implement an estimation strategy that takes advantage of variation in technology curves across business functions to produce measures of TechFP for each establishment. Technology curves account for 43% of the enormous within-establishment variance in technology. TechFP is strongly correlated with productivity across establishments, accounting for 15% of cross-establishment dispersion in productivity.

*We thank David Baqaee, Paulo Bastos, Najy Benhassine, Jiyoung Choi, Paulo Correa, Juanjo Dolado, Mark Dutz, Ana Margarida Fernandes, Cecilia Fieler, Caroline Freund, Mary Hallward-Driemeier, Leonardo Iacovone, Talip Kilic, Maurice Kugler, Bill Maloney, Denis Medvedev, Jorge Meza, Silvia Muzi, Mariana Pereira, Giacomo Ponzetto, Edouard Schaal, Chris Snyder, Stephen Yeo, Diego Zardetto, seminar participants and especially to Doug Staiger for comments and feedback that have helped us in developing and presenting the analysis. Kyun Min Lee actively participated in earlier drafts of this paper and we are indebted to him. We also thank João Belivaqua Basto, Tanay Balantrapu, Carmen Contreras, Pedro Martinez, Antonio Martins Neto, and Caroline Nogueira for their support through the implementation of the survey and preparation of the questionnaire, as well as the inputs from several external sector and production experts. We also thank the National Statistics Agencies and local partner institutions from Bangladesh, Brazil, Burkina Faso, Ghana, India, Kenya, Korea, Poland, Senegal, and Vietnam for their support in implementing the survey. An extended list of acknowledgement 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. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the World Bank or its Board.

1 Introduction

Suppose that you could open up an establishment and zoom into its functions to observe what technologies it uses to conduct the tasks associated with each function, and determine which of these technologies it uses most intensively. Access to this information can shed light on four first order questions. 1. What are the technologies used in a given business function? 2. How different (in some generalizable metric) are the technologies used in the various business functions of an establishment? 3. How can we aggregate all the technologies used by an establishment across its functions in an index that describes the relationship between technology and productivity in that establishment? 4. How much of the cross-establishment dispersion in productivity can be accounted for by the variation in this technological index? This paper explores these four questions, taking advantage of a new dataset that we have assembled, a novel model of technology choice inside the establishment, and several innovative empirical strategies that bring the model to the data.

FAT (Firm-level Adoption of Technology) is a representative[1] establishment-level dataset that currently covers 10 countries.[2] In the background of FAT, there is a grid structure (henceforth, the grid) that we used to organize business functions and technologies. This grid is a matrix that identifies the key business functions an establishment in a given sector conducts; we assembled it with the help of over 50 experts. The grid examines 7 business functions (BFs) that are common to establishments across all sectors, which we call general business functions (GBF)[3], and 56 additional BFs that are specific to one of 12 large sectors[4] and that we call sector-specific business functions (SSBF). For each business function, the experts identified the technologies, from the simplest to the most complex (i.e., the world technology frontier), that can be used to conduct the key tasks of that function. Overall, the grid covers 303 technologies. The grid has two key properties. First, it is relevant for any establishment and country, regardless of its sector and level of development. Second, for a given business function, the technologies in the grid can be ranked according to their sophistication from the simplest to the most complex.

To populate the grid for each establishment, we created the FAT survey and implemented it in representative samples (extracted from official sampling frames) in each of the 10 countries covered by FAT. The survey asks which of the technologies in the grid are used by and establishment and, among those used, which one is the most widely used. With this infor-

[1]FAT is representative at national, regional, sector, and firm-size levels.

[2]These include South Korea, Poland, the Brazilian state of Ceará, Vietnam, the Indian states of Uttar Pradesh and Tamil Nadu, Senegal, Kenya, Ghana, Bangladesh and Burkina Faso.

[3]See Figure 1 for the grid for GBF.

[4]Agriculture (crops), agriculture (fruits), livestock, food processing, apparel, automotive, pharmaceutical, other manufacturing, retail and wholesale, banking services, land transport services, and health services.

mation, we constructed business function-level measures of the sophistication of the most widely used technology (MOST) and of the most sophisticated technology used (MAX).

We first use these measures to document some facts about the use of technology in establishments and countries. We show that, on average, the technologies used by establishments are far from the technology frontier; that the most widely used technologies are significantly less sophisticated than the most sophisticated technologies available in a establishment/business function; that establishment size, and exporter, multinational, and multi-establishment status are associated with greater technological sophistication, and that there is large cross-country variation in technological sophistication which is very strongly correlated with per capita income.

The FAT dataset is designed so that the unit of observation is a business function in an establishment. This perspective allows us to pose new questions such as whether technological sophistication is relatively homogeneous across the different business functions of an establishment or whether it differs widely. We document that there is larger variance in technological sophistication within an establishment than across establishments. Interestingly, the within-establishment variance in technological sophistication is larger in more sophisticated establishments and is uncorrelated with establishment size and age.

To study the drivers of technology adoption across business functions we develop a model that allows for both heterogeneity in the marginal value and marginal costs of implementing more sophisticated technologies to affect technology choices. Additionally, the model introduces the concept of technological factor productivity, which is an index that aggregates all the technologies used by the establishment across functions and that is a sufficient statistic of the relationship between technology and productivity at the establishment level. A key prediction of the model is the existence of a parsimonious cross-establishment relation between the sophistication of technology in a business function and the technological factor productivity of the establishment. We name this relation the technology curve.

Using an empirical strategy inspired by Aguiar and Bils (2015),[5] we estimate the slopes of technology curves for each business function and the technological factor productivity (TechFP) of each establishment. We use these estimates to construct the predicted technology curves which account for 43% of the enormous within-establishment variance in technology sophistication.

Endowed with establishment-level estimates of TechFP, we explore various relevant questions. One that is central to our effort of building a comprehensive dataset of technology at the function/establishment level is how much information we would lose if we restricted

[5] Aguiar and Bils (2015) estimate household-level expenditure by exploiting the heterogeneity in the slopes of Engel curves for different consumption categories.

attention to the presence of specific technologies in the establishment. Relatedly, we study how TechFP relates to various moments of the distribution of different measures of technology sophistication. We also explore how TechFP is related to establishment characteristics, such as the education of the workers, the quality of management practices, size, age, and status as exporter, multi-establishment or multinational. Finally, we study the association between TechFP and establishment productivity finding that the two are strongly associated and that cross-establishment variation in TechFP accounts for 15% of the gap in labor productivity between the establishments in the 90^{th} and 10^{th} percentiles of the distribution. We also find considerable sectoral heterogeneity in the fraction of cross-establishment dispersion accounted for by TechFP ranging from 30% in agriculture to 13% in services.

Our analysis is related to several literatures. First, the FAT dataset is inspired by a long tradition in economics, sociology and marketing in the measurement of technology initiated with Ryan and Gross (1943) and Griliches (1957). Since then, it has become a common practice to characterize the technology of a firm by the presence of a few (often just one) advanced technologies.[6] Recent efforts such as the Canadian SAT have added detail by increasing the number (between 41 and 50, depending on the round) of advanced technologies that firms are asked about. FAT extends these measurement efforts while introducing some significant differences with prior studies. First, the number of technologies covered is much greater covering both general and sector-specific technologies. Second, FAT covers a full range of technologies from the most basic to the state-of-the-art. Third, the unit of observation in FAT is the business function in a establishment. Importantly, this new unit of analysis permits to know what technologies are used for and to study variation in technology across the business functions of an establishment. Finally, FAT measures not just the presence of a technology but also what is the most widely used technology in the function. The intensity of use is an important dimension for the technological productivity of the establishment.

There are interesting parallels between our contribution to the measurement of technology and pathbreaking efforts by Bloom and Van Reenen (2007) and Bloom et al. (2019) to measure the quality of management practices also using establishment surveys.[7] Beyond the obvious conceptual differences between technologies and managerial practices, the main

[6]For example, Davies (1979) studies the diffusion of 26 manufacturing technologies but, typically, each technology is relevant only in one 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 internet. This practice has been adopted in surveys of ICT conducted by the statistical offices in a number of advanced economies, including the US Census Bureau (ICTS and ABS), the Eurostat (Community Survey of ICT Usage), and Statistics Canada (Survey of Advanced Technology (SAT).

[7]These surveys include the World Management Survey (WMS) and the Management and Organizational Practices Survey (MOPS). The WMS is a telephone based survey that uses double blinded methodologies. MOPS is an online and paper based survey.

difference is that we have a genuine interest in studying differences in the sophistication of technology across the business functions of a given establishment, and that we exploit this dimension together with the structure of the model to estimate the technological factor productivity of the establishment.

The estimation of TechFP is related to the macro literature that has intended to construct aggregate measures of TFP that are purged from non-technological factors such as cyclical capacity utilization and profits (e.g., Basu, Fernald and Kimball, 2006, and Comin, et al., 2023). This literature starts with measures of the Solow residual and adjusts them using proxies for unmeasured variation in factor utilization and markups. In contrast, our measures of TechFP are constructed bottom up, relying on a comprehensive number of direct measures of the technologies used by the establishment across business function and the structure of a model of technology choice across functions.

Finally, there is a long tradition studying the relationship between technology and productivity across establishments including the so-called Solow Paradox by which the anticipated effect of the adoption of computers in productivity measures was elusive in empirical analyses.[8] These studies typically consider a limited number of technologies.[9] Our finding that measures of establishment technology that rely on one or a few technologies provide a poor and imprecise approximation of TechFP provides a rationale for the limits of this approach to study the relationship between technology and productivity.

The rest of the paper is organized as follows. Section 2 introduces the FAT survey. Section 3 presents the measures of technology we construct with the information collected in the survey, and studies technology choice at the business function level, and across the business functions of a establishment. Section 4 develops our model of technology choice across business functions. Section 5 presents the strategy to estimate technological factor productivity at the establishment level and analyzes the estimates. Section 6 concludes.

[8] A related literature has explored this question using more aggregate units of analysis such as countries (e.g., Comin and Hobijn (2010) and Comin and Mestieri (2018)) or sectors (e.g., Comin (2000), Jorgenson et al. (2005), Jorgenson, Ho and Stiroh (2008), Oliner, Sichel and Stiroh (2007), Van Ark, O'Mahoney and Timmer (2008)).

[9] For example, Hubbard (2003) studies the effects of adopting on-board computers in trucks, Bartel, Ichniowski and Shaw (2007) study the effects of the adoption of computer numerically controlled (CNC) machines and computer-aided design (CAD) software in the productivity of valve manufacturing. Hjort and Poulsen (2019) analyzes how the access to fast Internet connection increases firm entry, productivity, and exports in African countries. Gupta, Ponticelli and Tesei (2020) study how the adoption of cellphones by Indian farmers increased their productivity by reducing their informational barriers.

2 The Survey

The FAT survey ("the survey" henceforth) collects detailed information for nationally representative samples of establishments in agriculture, manufacturing, and services about the technologies that each establishment uses to perform key business functions necessary to operate in its respective sector. In what follows we describe in detail the survey design and implementation.[10]

2.1 Structure

The survey is composed of five modules. Module A collects information on general characteristics about the establishment.[11] Modules B and C cover the technologies used. Module D focuses on barriers and drivers of technology adoption, while module E gathers information about the establishment's balance sheet and employment.[12]

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 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 ten significant sectors in the economy.[13] These sectors have been selected to cover all three industries (agriculture, manufacturing, and services) and based on their share in aggregate value-added, employment and number of establishments.[14]

2.2 Technology grid

To design modules B and C, we determine the business functions covered and the list of technologies that can be used to implement the key tasks in each function. We call the resulting structure: the grid.

---

[10]See Appendix A for more details.

[11]The survey is designed, implemented, and weighted at the establishment level. For multi-establishment firms, the survey targets the establishment randomly selected in the sample.

[12]The survey can be downloaded in the following address (https://dcomin.host.dartmouth.edu/files/FAT_Survey_complete.pdf).

[13]The ten sectors for which we have developed sector-specific modules are: agriculture (crops and fruits), livestock, food processing, wearing apparel, automotive, pharmaceutical, retail and wholesale, banking services, land transport services, and health services.

[14]The granular information that can be obtained with the FAT survey allows to explore central questions on technology policy in developing countries. One example, and product of this paper, is the World Bank policy report "Bridging the Technological Divide"(https://openknowledge.worldbank.org/server/api/core/bitstreams/5a5e55f7-edf8-530e-8e11-aa2e15421a9d/content).

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 in each of the sectors covered. Third, we reached out to external consultants with significant experience in the field (at least 15 years). For example, the external experts in agriculture and livestock were agricultural engineers and researchers from Embrapa-Brazil. For food processing, wearing apparel, automotive, pharmaceutical, transport, finance, and retail, as well as for the GBFs, we relied on senior external consultants selected by a large management consulting organization. For health, the team relied on consultants and physicians with practical experience in both developing countries and advanced economies. In total, over 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 a total of 303 technology/business functions pairs (See subsubsection A.1.1 of the appendix for details on the process followed to define the grid). Figure 1 presents the general business functions considered in the survey and the possible technologies that can be used to conduct each of them. Figure 2 presents the grid for food-processing, one of our SSBFs.[15]

The grid in FAT has three characteristics. First, it is comprehensive. It includes the main business functions and the full array of technologies in each function, from the most basic to the most advanced technologies available. Second, the business functions and technologies in the grid are potentially relevant to all firms within any given sector. For example, the business functions and technologies in the grid for crop agriculture should allow us to characterize the technologies used both by large producers of soybeans in Brazil, and small producers of rice in Vietnam. Third, the technologies are precisely defined so that their use in a firm can be objectively established by respondents and enumerators.

In addition to identifying the key business functions and relevant technologies, experts also identified the ranking of the technologies in each business function based on their sophistication. The sophistication of a technology reflects the complexity of the technology and is often associated with the novelty of the technology (e.g., crypto payments are more complex than cash). More sophisticated technologies may be able to perform more tasks, tasks of greater complexity or may perform them with greater accuracy or speed, but do not necessarily imply grater productivity. Thus, the sophistication rankings are not based on ex-ante or ex-post information about the relative productivity associated to these technologies.

[15] The grids for the GBFs and the eleven SSBFs are available in section A.1.1 of the appendix.

2.3 Information collected in FAT

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

2.3.1 Business functions

The business functions that comprise the vertical dimension of the grid are comprehensive (i.e. cover the key tasks involved in production) and relevant (i.e. the tasks they group are conducted in most establishments in the sector). Explorations conducted at the piloting stage of the survey confirm the relevance of general business functions as they confirm that the functions are conducted in house, and respondents are aware about the technologies used for those functions.[16] For sector-specific business functions we formally explore the relevance of each business function in each establishment through a screener questions that ask whether the specific business function is conducted in the establishment and, in case it is not, whether it is in-sourced, outsourced, or irrelevant. We find that in 80% of the business function/establishment pairs, the function is relevant and conducted in the establishment.

2.3.2 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. FAT also asks whether the establishment uses "other technologies" in the business function in addition to those contained in the grid. The frequency that respondents declare that "other" technologies are used in the business function allows us to assess the comprehensiveness of the technologies in the grid. Establishments use "other" technologies in 3.65% of the business functions, which confirms that the technologies in the grid are comprehensive.[17]

After identifying the technologies that are used by the establishment in a business function, the survey asks which one is the most widely used in the business function. In addition to collecting information on technology used based on the grid, the FAT survey also asks establishments about the presence of three standard general purpose technologies: computers, internet, and electricity. Unlike the FAT technologies, the presence of these general purpose technology is not connected to any specific business function.

[16]This premise is confirmed ex-post by the very low frequency of missing GBFs, defined as the share of functions for which firms replied "no" or "do not know" for all options of technologies available in a given business function. Specifically, the frequency of "missing" GBFs, by taking into consideration all GBFs across all firms, is 4%.

[17]"other" is the most widely used technology in 0.8% of the business functions.

The survey also includes other standard questions in establishment surveys, such as balance sheet information, employment and occupations, management characteristics and a few variables from MOPS to measure management quality (see Appendix A for a more detailed description).

2.4 The Data

Our analysis is based on primary data collected across 10 countries: South Korea, Poland, Brazil (State of Ceará), Vietnam, India (States of Uttar Pradesh and Tamil Nadu), Senegal, Kenya, Ghana, Bangladesh, and Burkina Faso. Several factors were taken into consideration to decide where to implement the FAT survey. We targeted countries in different continents (Asia, Africa, America, and Europe) with different levels of income, and where there was access to a good quality sampling frame. In these countries, we collected data from 12,636 establishments, including 903 from Bangladesh, 711 from the State of Ceará, in Brazil, 600 from Burkina Faso, 1,262 from Ghana, 1,519 from India, 1,305 from Kenya, 1,551 from South Korea, 1,500 from Poland, 1,786 establishments in Senegal, and 1,499 from Vietnam. Table 1 shows the distributions of the sample by country, sector, and size groups. These establishments were randomly selected based on the sampling frame for each country.

Table 2 provides descriptive statistics for the sample we used in this study. The median firm in our sample has 20 workers, with an average of 98 workers. About 23% of workers have a college degree, 21% of firms are multi-establishments, 11% has foreign capital, 21% are exporters, 14% has 5 or less years of age, and 76% has electricity, computer, and internet available.

2.4.1 Sampling

Our data is representative for a universe of about 1.2 million establishments.[18] The samples are nationally representative for establishments with 5 or more workers.[19] For each country, the sampling frame is based on the most comprehensive and updated establishment-level census data available from the respective National Statistical Office (NSOs) or similar administrative information.[20] The survey is stratified along three dimensions - sector, firm size, and region - so we can construct representative measures of technology for aggregates

[18]Table A.2 provides information of the distribution of firms by country, sector, and size groups, in the universe covered by the FAT survey.

[19]For the state of Ceará in Brazil, and the Indian states of Tamil Nadu and Uttar Pradesh it is representative at the state-level.

[20]Section A of the appendix provides more details on sampling frame, survey implementation and data collection, and sampling weight.

along these dimensions. Sampling weights are based on the inverse probability of selecting establishments within each stratum.

2.4.2 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 the non-response, the enumerator, and the respondent (Collins, 2003). In what follows we briefly describe the steps taken in the design and implementation of the FAT survey to minimize these errors. Appendix A.4 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 follow best practice procedures. First, we partner with national statistical offices and industry associations to use the most comprehensive and updated sampling frame available. Second, we hire data collection companies or agencies with demonstrable experience in nationally representative firm level surveys; which are supported by endorsement letters from local institutions.[21] Third, we follow a standard protocol in which each firm is contacted several times to schedule an interview. Fourth, we use face-to-face or phone interviews which lead to higher response rates than web-based interviews.[22]

Enumerator bias and error counts. The survey, training, and data collection processes are largely designed to minimize enumerator biases and data collection errors. First, we use closed-ended questions to make coding the answers a mechanical task, eliminating the reliance on the enumerator's interpretation of the answer and subjective judgement to code them. Second, the same standardized training is implemented in each country in local language with enumerators, supervisors, and managers leading the data implementation. Third, in each country, we conduct a pre-test pilot of the questionnaire with establishments out of the sample. Fourth, to attain greater quality control during the data collection process, enumerators record the answers via *Computer-Assisted Personal Interviews* (CAPI) or *Computer-Assisted Telephone Interviewing* (CATI) software and we regularly monitor the data collection process using standard algorithms to analyze the consistency of the data and assure quality control.[23]

[21] These procedures are in line with suggestions of good practice for implementation by Bloom et al. (2016).

[22] Note that face-to-face interviews were not possible during the pandemic and in some countries as shown in Table A.3 the survey was implemented by telephone.

[23] Randomized survey experiments with household survey has demonstrated that a large number of errors observed in *Pen-and-Paper Personal Interview* (PAPI) data can be avoided in CAPI (Caeyers, Chalmers and De Weerdt, 2012).

Respondent bias. We take several steps to minimize respondent bias. First, we ensure that the interview is arranged with the appropriate person or persons; main managers (and in larger firms other managers such as plant managers and HR managers). Second, we use a closed-ended design in the questionnaire that reduces measurement error in the answers as the respondent is questioned about specific technologies (one at a time), not knowing ex ante all technologies to be asked in each business function. Third, we pre-test the questionnaire in each country to ensure that questions are clear in their wording in the specific geographical and cultural contexts, so that the response does not require any subjective judgement (Bertrand and Mullainathan, 2001). Fourth, to avoid *social desirability bias*, by which respondents may overstate the use of more sophisticated technologies, the survey avoids the words "technology" and "sophistication" and employs more neutral terms such as "methods" and "processes".

2.4.3 Ex-post checks and validation exercises

We conduct 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 24 percent and 80 percent. The response rates were higher when the survey was implemented by national statistical agencies. For example, response rate was 80 percent in Vietnam, 57 percent in Senegal, 39 percent in Ceará, Brazil, and 24 percent in Korea.[24] 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.[25]

We conduct three tests to assess potential biases from unit non-response-rates.[26] In each of these exercises, presented in section A.5 of the appendix, we find no statistical difference

[24] Table A.4 in the appendix A provides the response rate by country, defined as ratio between firms that responded the survey and the total number of eligible firms in the sample, for which we attempted to conduct the interview. The high response rate for Vietnam is associated with the fact that the survey was implemented by the national statistical office. These response rates are high relative to typical response rates in firm-level surveys, which for the U.S. are around 5 to 10 percent, and are consistent with response rates observed for WMS and MOPS (Bloom et al., 2016).The average response rate for the WMS is around 40 percent.The response rate for MOPS, implemented by the United States Census Bureau, was around 80 percent.

[25] See section A.3 of the appendix for more details on sampling weights.

[26] First, using the information from the sampling frame, we check if there are differences in the average number of workers per firm between respondents and non-respondents within stratum. Second, using information on the number of attempts, we compare the firm-level technology sophistication in GBFs, described in the next section, between firms with above and below the average number of attempts. Third, in a similar vein, we compare firms in the first list of contacts provided to interviewers, versus those provided subsequently.

in the number of employees, technology sophistication, wages, and share of workers by skill and education between firms in the group that proxies for the response sample and the group of firms that proxies for the non-response sample.[27]

Response bias. To assess the relevance of response bias, we conduct a parallel pilot in Kenya where we re-interview 100 randomly selected firms with a short version of the questionnaire. For those firms, we randomly select three business functions and ask about the presence of the relevant technologies.[28] We estimate a probit model to assess the likelihood of consistent answers between the original and the back-check interviews, controlling for firm-level fixed-effect. Reporting the use of a technology in the back-check interview is associated with 80.6% of 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.[29] These estimates do not differ across establishments of different size.

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).[30] We focus on variables related to establishment size, productivity and technology. Table A.11 shows that the sampling weighted average 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 Korean matched establishments we find very high correlations (above 0.93) in the log levels and growth rates of sales, employment and sales.[31]

With regards to sales per worker, in Korea, we find a cross-establishment correlation between FAT and the KED of 0.73. While, in Ceará, Brazil, Table A.10 shows that average log of wages from RAIS are strongly correlated with the FAT measures of log value-added per worker. With regards to technology, we find similar average adoption rate of ERP systems in Korean manufacturing establishments in FAT as Chung and Kim (2021) who use a similar

[27]See Table A.5 to A.11 in Appendix A.

[28]The re-interviews produce 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.

[29]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.

[30]In Korea we merged FAT with the Korea Enterprise Data (KED), a leading supplier of business credit reports on Korean businesses. In Brazil, we merged the data with the *Relação Anual de Informações Sociais* (RAIS) which is an administrative database maintained by the Ministry of Labor providing information of salary for all formal workers in Brazil.

[31]The FAT survey asks about sales and number of employments for two periods. The most recent year for which the balance sheet is available, which is the year previous the implementation of the survey and two years before. For Korea, these reference years are 2019 and 2017.

sampling frame (32% vs. 40% in Chung and Kim (2021)). Additionally, we find a strong positive cross-establishment correlation between the book value of machinery and equipment in KED and the establishment sophistication measures (MOST and MAX) from FAT, to be explained in the next section. These checks support the soundness of the survey design, data collection process, and accuracy of responses.

3 Facts about Technology

The horizontal dimension of the grid provides detailed information about the technologies used in each business function of each establishment and about which ones are most intensively used. To reduce the dimensionality of this information, we construct measures at the business function/establishment level that reflect the sophistication of two specific technologies: the most widely used technology, and the technology with maximum sophistication in the business function used in the establishment. After discussing the constructions of these measures, we first characterize their distribution across business functions and establishments and explore the establishment characteristics that are associated with technology sophistication. Then, we take advantage of the granularity of the FAT dataset and pose a new question to the literature: how much heterogeneity in technology there is across the business functions of an establishment? To answer this question we compute the within-establishment variance in technology sophistication and compare it to the cross-establishment variance. We then study possible sources of within establishment variation in technology sophistication and present the technology curve.

3.1 Measures of technology sophistication

We measure the sophistication of each technology in each business function by proceeding in two steps. First, we order the technologies in each function based on their sophistication (i.e., complexity). Second, we map the ordinal sophistication rankings into a cardinal measure.

The starting point is the experts' rankings of the technologies in each business function, from least to most advanced, $r_f \in 1, 2, ..., R_f$.[32] We define the relative rank of a technology as $\hat{r}_f = \frac{r_f - 1}{R_f - 1}$. Note that $\hat{r}_f \in [0, 1]$.

[32]Because 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.

Our analysis focuses on the sophistication of two types of technologies: the most widely used and the technology with maximum sophistication among those used in the business function by the establishment. We construct cardinal measures of sophistication of a technology by applying an affine transformation to the relative rank of a technology, $\hat{r}_f$. Specifically, the sophistication of the most widely used technology in function f of establishment j, $MOST_{f,j}$, and of the technology with highest sophistication, $MAX_{f,j}$, are computed by:

$$MOST_{f,j} = 1 + 4 * \hat{r}_{f,j}^{MOST} \tag{1}$$

$$MAX_{f,j} = 1 + 4 * \hat{r}_{f,j}^{MAX}, \tag{2}$$

where $\hat{r}_{f,j}^{MOST}$ and $\hat{r}_{f,j}^{MAX}$ are the relative sophistication rankings of the two technologies. By construction, $MOST_{f,j}, MAX_{f,j} \in [1, 5]$, and $MAX_{f,j} \geq MOST_{f,j}$.

Since the technology with highest rank in the grid corresponds to the (world) technological frontier in that business function, we can interpret cardinalizations of the relative ranking of a technology (e.g., $MOST_{f,j}$ and $MAX_{f,j}$) as the closeness of a technology to the frontier. The affine nature of the transformations in (1) and (2) implies that these measures display constant increments as we move up in the sophistication ranking. This is a natural assumption commonly used to construct cardinal variables in social sciences. We have carefully studied the robustness of the key findings in our analysis to the linearity of (1) and (2) (see Cirera et al. (2020)) and have found that the findings hold generally and do not hinge on the linearity of the sophistication measures. For example, we have checked that a sufficient condition for the robustness of the sign of the correlations between sophistication measures and firm observables to arbitrary monotonic cardinalizations of the sophistication rankings holds.[33] Similarly, we have checked the robustness of the technology curves and their slopes to using cardinalizations of relative rankings that involve a wide range of concave and convex functions.

Out of all the possible technologies we could focus on, those that define $MOST_{f,j}$ and $MAX_{f,j}$ seem probably the most relevant as they define the technological frontier in the business function and establishment. Additionally, $MOST_{f,j}$ and $MAX_{f,j}$ are of independent importance as they reflect different aspects of the technology upgrading processes in the business function. $MAX_{f,j}$ increases when a firm implements a new technology, that is more

[33] This is the so-called LMA condition Schroeder and Yitzhaki (2017) which required that the LMA curve (line of independence minus absolute concentration curve) does not cross the horizontal axis. The LMA curve is the vertical difference between two curves. The first curve is the absolute concentration curve of a variable X (e.g., firm size) given the technology sophistication s_j under the assumption that the two variables are statistically independent. The second curve is the absolute concentration curve of X as a function of the cumulative distribution $F(.)$ of the sophistication measure across establishments.

sophisticated than those currently in use in the business function.[34] Therefore, increases in $MAX_{f,j}$ capture technology improvements as those in quality ladder (e.g., Aghion and Howitt, 1992) or horizontal variety (e.g., Romer, 1990) conceptualizations of technology in production.

Increases in $MOST_{f,j}$ occur because the establishment expands the use of a technology which becomes the new most widely used in the business function.[35] The new most widely used technology may be either new in the business function or it may be a technology that was used marginally and whose use has been expanded. Therefore, $MOST_{f,j}$ is more closely connected to Mansfield (1963)'s notion of technology diffusion in the firm (in our case in the BF) than to innovation.

Relevant outcomes and observable characteristics are often reported at the establishment level. We construct establishment-level average sophistication measures as the simple average of $MOST_{f,j}$ and $MAX_{f,j}$ across the business of an establishment:

$$S_j = \sum_{f=1}^{N_j} \frac{S_{f,j}}{N_j} \tag{3}$$

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

3.2 Describing technology sophistication

We start exploring technology sophistication by studying the distributions of MAX and MOST at the business function and establishment levels (see Figure 4). Table 3 presents key statistics of these distributions. Fact 1 summarizes our main observations.

Fact 1.

A. The average technology sophistication across establishments/functions is relatively low, and it is considerably higher for MAX (2.41) than for MOST (1.78).

B. There is a large dispersion of technology sophistication across function and establishments.

C. The distributions of MAX and MOST, both at the business function and establishment level, are skewed to the right.

[34] This technology may not be new to the establishment, but it is new to the BF in the establishment.

[35] Obviously, for MOST to increase, the newly most widely used technology must be more sophisticated than the previous most widely used technology.

Part A of Fact 1 shows that the average establishment is quite far from the world technology frontier. The low average sophistication together with the right skewness of the distribution (Part C) signify that the technology used in many BFs and establishments are relatively unsophisticated. The right skewness of technology sophistication measures is reminiscent of the well-documented skewness in the distribution of log productivity across establishments (e.g. Syverson (2011)).

Part A also points to the fact that the average gap between MAX and MOST (0.63) represents 16% of the range in the support of the technology sophistication measure and roughly half of the cross function/establishment dispersion in sophistication. This gap implies that establishments have more sophisticated technologies than those they use most widely.

Part B of Fact 1 highlights the large dispersion of sophistication both at the business function and establishment levels. We assess the magnitude by comparing the standard deviations with the range of the support of technology sophistication (i.e., (5-1)=4). At the establishment level, this ratio is 20.5% for MAX and 15% for MOST. At the business function level the ratios are 32% for MAX and 25% for MOST.

We next explore the sources of variation in technology sophistication across establishments. To this end, we regress the establishment-level technology measures (MAX_j and $MOST_j$) on dummies that reflect establishment size (5-19, 20-99, 100+ employees), age (6-10, 11-15, and 16+ years), 1-digit sector (agriculture, manufacturing and services), export, multi-establishment and foreign ownership status, and the country. Table 4 reports the estimates from these regressions. Fact 2 summarizes the key findings.

Fact 2.

A. There is a positive association between technology sophistication and establishment size.

B. Being an exporter, part of multi-establishment firm or a multinational are also positively associated with the three technology measures.

C. There is no significant association between technology and the age of the establishment.

D. Size, age, exporting, multinational and multi-establishment status, together with the country and one-sector dummies account for 41% and 45% of the cross-establishment variance in technology sophistication measured by MAX and MOST, respectively.

The coefficients on the country dummies reflect the average technology sophistication levels of the establishments in each country. Table 5 reports those coefficients as well as the coefficients of similar regressions where we introduce sector-specific country dummies and that capture the average sophistication in each one-digit sector (i.e., agriculture, manufacturing and services) by country. Table 5 also reports the dispersion and correlation with per capita income of the country-level sophistication measures.

Fact 3.

A. There is a large cross-country variation in average sophistication levels for both MAX and MOST.

B. The correlation between per capita income and technology sophistication across countries is a strong (0.93 for $MOST$ and 0.76 for MAX).

C. The cross-country dispersion in average technology sophistication and the covariance between sophistication and per capita income are largest for agriculture.

Parts A and B of Fact 3 imply that technology sophistication is a relevant factor in a development accounting sense. In particular, MOST accounts for 36% to the cross-country variance in per capita income while MAX accounts for 30%.[36]

Part C of Fact 3 is relevant for the literature on the agricultural productivity gap (e.g., Caselli (2005), Lagakos and Waugh (2013)) that has documented the presence of larger cross-country differences in productivity in agriculture than in other sectors. The fact that cross-country differences in technology sophistication are also larger in agricultural establishments than in establishments that operate in other sectors suggests a new line of research that should help us understand better this important puzzle.

3.3 Technology inside the establishment

The granularity of the information collected in the FAT survey offers a unique opportunity to study the variation of technology across the business functions of an establishment. Specifically, we explore whether the technologies used by companies across the different functions of a establishment have a relatively uniform sophistication or whether there are large differences in technology sophistication across the business functions of an establishment. After establishing the magnitude of the variation in technology within establishments, we proceed in two directions. First, we study the variation across establishments in the within-establishment variance in technology sophistication. Specifically, we study the magnitude of variation across establishment in the within-establishment variance in technology and how the within-establishment variance in technology sophistication covaries with establishment

[36] The contribution is computed as the covariance divided by the variance of per capita income.

characteristics. Second, we make the business function/establishment the unit of observation to study the cross-establishment relation between technology sophistication in a specific business function and average sophistication in the establishment. This relationship, which we call the technology curve, can be interpreted as the expansion path of technology sophistication, as its slope informs us about how much a establishment loads on a business function as it improves the overall sophistication of technology. We are interested in studying the variation in the slopes of technology curves across business functions and how much of the within-establishment variance in technology is explained by variation in the slopes of technology curves.

Within-establishment variance in technology

We study the sources of variation in technology sophistication at the establishment/function level (i.e., $MAX_{f,j}$ and $MOST_{f,j}$, generically denoted by $S_{f,j}$), by decomposing them into a establishment component (α_j), a business function-country component ($\alpha_{f,c}$), and a residual ($u_{f,j}$):

$$S_{f,j} = \alpha_j + \alpha_{f,c} + u_{f,j}, \tag{4}$$

The business functions dummies α_f absorb variation in $S_{f,j}$ driven by the nature of business functions in a firm. By purging this component, the levels and variances of the establishment effects and the residuals are comparable across firms with potentially different business functions (as they may have different SSBFs). We then compute the variance of the establishment, business function/country and within-establishment components in (4). (See Table 6.)

Fact 4.

> The within-establishment variance in technology (WVAR) is larger than the cross-firm variance in technology. In particular, it is 38% larger for MAX, and more than twice as large for $MOST$.

The lack of datasets on technology at the business-function-level has made impossible for prior research to explore the within-establishment variation in technology. Models of the firm have reduced the technological landscape of a firm to a unique parameter, implying that the technologies used by a firm across the different business functions were uniformly sophisticated or unsophisticated. The enormous variation in technology inside the establishment documented in Fact 4 debunks this notion and motivates new questions about the sources of this variation that we formulate and address in the remaining of this paper.

Additionally, Fact 4 has implications for the measurement of technology. A common practice in studies of technology across firms is to reduce the technology of a firm/establishment to a single technology (e.g., the presence of a specific technology, or the intensity with which a specific technology is used). The large variation across the technology sophistication of the business functions of a establishment implies that relying on information on one (or a few technologies) will provide a very imprecise measure of the technology in the establishment/firm. We directly explore this issue in section 5.2.

WVar across establishments

We start unpacking the within-establishment variance in technology by exploring two questions. First, how much does the within-establishment variance in technology measures vary across firms. Second, what establishment characteristics co-vary with the within-establishment variance in technology.

To explore these questions, we define the within-establishment variance in technology sophistication for establishment j ($WVar_j$) as the variance across the business functions of establishment j of $u_{f,j}$ in (4). Figure 5 plots the kernel densities of $WVar_j$ across establishment for both MAX and $MOST$, and Table 7 reports some moments of the distribution that we summarize in Fact 5.

Fact 5.

> There is large variation across establishments in their within-establishment variance in technology measures. The ratios of the within-establishment variance in technology between the establishments at the 90^{th} and 10^{th} percentiles of the distribution are 7 for MAX and 10 for $MOST$.

We explore the sources of within-establishment variance in technology by regressing $WVar_j$ on establishment-level observable characteristics, X_j, and a second order polynomial of average establishment sophistication. Table 8 reports the estimates of equation (5).

$$WVar_j = \alpha_c + \beta_1 S_j + \beta_2 S_j^2 + \gamma X_j + u_j \tag{5}$$

Fact 6.

A. There is a strong positive association between $WVAR_j$ and the establishment-level technology sophistication, for both MAX and MOST. The relationship between these variables is concave but the peak is close to the maximum possible level of technology. (See Table 8.)

B. There are weak and/or insignificant associations between within-establishment variance in technology and establishment size, age, number of business functions in the establishment, export, multi-national, and multi-establishment status. (See Table 8.)

The insignificance of the number of business functions covered in the establishment by FAT in (5) provides reassurance about the signal to noise ratio of $S_{f,j}$ as if noise was an important driver of the cross-function variation in technology sophistication, we should find a strong negative coefficient.

The fact that, conditional on S_j, establishment size is also insignificantly associated with $WVar_j$ reveals a limited role of heterogeneity in the marginal costs of technological sophistication for within-establishment variance in technology. Specifically, as the marginal cost of technology sophistication in capital- and knowledge-intensive functions tends to be disproportionately higher for small establishments, if cross-function heterogeneity in the marginal costs of technological sophistication was an important driver of within-establishment variance we would expect a negative coefficient of size in (5). The findings in Part B of Fact 6 hinge on the inclusion of S_j in specification (5). However, there are various motivations to controlling for S_j. Suppose, for example the following model for $S_{f,j}$

$$S_{f,j} = \alpha_f + S_j + v_{f,j} \tag{6}$$

, where $v_{f,j}$ is the deviation in sophistication in function f from the average establishment sophistication. In this case, the definition of within-establishment variance in technology sophistication implies that $WVar_j = Var(v_{f,j})$. There is an upper bound for the variance of $v_{f,j}$ which is $MAXVar((v_{f,j}) = (S_j - 1) * (5 - S_j) = 6S_j - S_j^2 - 5$. This quadratic function of the average sophistication in the establishment S_j peaks at $S_j = 3$.

The estimates in Table 8 confirm the quadratic nature of the relation between $WVar_j$ and S_j but they imply a peak variance attained at significantly higher levels of establishment sophistication (3.6 for MAX and 4.1 for MOST). This discrepancy suggests that the cross-establishment relationship between $WVar_j$ and S_j may be due to other mechanisms different than the mechanical association induced by the functional form of $MAXVar$. We explore one such alternative next.

The technology curve

Suppose that establishments choose the technological sophistication of a business function according to the following rule:

$$S_{f,j} = \alpha_f + \varepsilon_f * S_j + v_{f,j}. \tag{7}$$

Expression (7) is a generalization of (6), where the parameter ε_f captures how much a establishment loads in business function f as it increases the average sophistication, S_j. Since ε_f is indexed by f, the expansion paths/rays defined by 7 can have different slopes across functions. Consequently, the rule in (7) generates a greater within-establishment variance in technology sophistication as an establishment implements a higher S_j.

We study graphically whether technology sophistication choices in our data are consistent with expression (7) by collapsing establishments into deciles of the distributions of $MOST_j$ and MAX_j, respectively. For each decile and business function, we calculate the average value of $MOST_{f,j}$ or $MAX_{f,j}$. Figure 6 plots, the average value of $MOST_{f,j}$ (vertical axis) against the average of $MOST_j$ (horizontal axis) for each decile of $MOST_j$. For example, in the top left panel of Figure 6 we observe that the average sophistication level in "payments" for establishments in the bottom decile of the distribution of average sophistication is 1.7, while the average sophistication for these establishments is 1.1. We name the expansion ray for a given business function **the technology curve**.[37]

The top panel plots the technology curves for the seven GBFs, while the other four panels plot the technology curves for the sector-specific business functions in each of the four sectors with more establishments in sample (crops-agriculture, food processing, apparel, and retail and wholesale).[38]

Figure 6 reveal interesting patterns. Not surprisingly, technology curves are upward sloping. That is, as we move to higher deciles in the distribution of average establishment sophistication, the sophistication in any given business function tends to increase. However, what is remarkable is that the slope of the technology curves varies a lot across business functions. For example, among the GBFs, the most technology-elastic (i.e., largest slope) functions are business administration and planning, while the least technology-elastic is sales. SSBFs also display heterogeneity in the slope of technology curves. The most technology-elastic functions in each sector are irrigation in agriculture, design and finishing in apparel, packaging in food processing, and advertising and inventory in retail and wholesale. Finally, these patterns are quite similar in technology curves based on $MOST_j$ and MAX_j although the dispersion in slopes across business functions is smaller for MAX.

Fact 7.

> There are large, statistical-significant differences in the slope of technology curves across business functions.

[37]See Appendix C for the plot of the technology curves using MAX instead of $MOST$ as the measure of technology sophistication.

[38]These are sectors for which the survey was stratified in all countries. In grey we plot 95% confidence bands.

Fact 7 confirms that rule (7) provides a good characterization of technology choices across functions. It also raises two questions. The first is whether differences in the slopes of technology curves are due to differences across functions in the nature of the technologies in the grid or to the nature of the function. We explore this question in the next section with the help of our model. Fact 7 also prompts the question of whether technology curves are stable across countries. To investigate this, we classify establishments in three groups based on the income of their country.[39] Then, for each function and decile of $MOST_j$, we compute the average of $MOST_{f,j}$ across the establishments in each country group. This yields three technology curves per business function. Figure 7 plots them for four specific business function (business administration, payments, irrigation and packing in agriculture). Visual inspection suggests that, in each of the four functions, the technology curves in the three country groups line up quite closely. To quantify the significance of the role of cross-country differences vs. cross-function differences in technology curves, we conduct a variance decomposition and find that 88% of the variance in the (country grouping-level) technology curves is due to the variance in technology curves across functions, with the remaining variance due to the cross-country grouping differences in technology curves.[40] This finding motivates Fact 8.

Fact 8.

Technology curves are remarkably stable across countries.

Next, we build a model of technology choice within establishments to micro found the technology curves.

4 A Model of technology use across business functions

We develop a model of technology sophistication across business functions that helps us account for the within-establishment variance in technology and provides a micro-foundation for the technology curves. The model also provides a framework to study how the vector of technology sophistication levels across the business functions of a establishment aggregate into an establishment-level index that represents the technological component of TFP: the TechFP index. In the next section, we present and implement a strategy to estimate the TechFP of each establishment.

[39]Korea, Poland and Brazil are high income, Vietnam, India, and Ghana middle and Bangladesh, Kenya, Senegal and Burkina Faso low-income.

[40]See Appendix y for details.

4.1 Optimal technology sophistication across business functions

Establishment j selects the sophistication of technology in each business function f, $s_{f,j}$, to maximize profits net of adoption costs.[41] The levels of technology sophistication across business functions are linked to profits through an index that reflects the overall technology of the establishment, a_j, and that we call the technological factor productivity (TechFP) of the establishment. One can think of technological productivity as the component of TFP determined by the technology choices of the firm, and it is implicitly defined by the following non-homothetic constant elasticity of substitution (nh-CES) aggregator:[42]

$$\sum_{f=1}^{N_f} \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}} \right) e^{\frac{\sigma-1}{\sigma} s_{f,j}} = 1 \tag{8}$$

where $\sigma \in (0,1)$ is the elasticity of substitution between sophistication across business functions, $\Omega_f > 0$ (with $\sum_{f=1}^{N_f} \Omega_f = 1$) and $\varepsilon_f (> 0)$ affect the importance of business function f for the technology index. To better understand their role, we rearrange (8) as follows

$$e^{a_j} = \left[\sum_{f=1}^{N_f} \underbrace{\left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{(\varepsilon_f - (1-\sigma)) a_j}{\sigma}} \right)}_{\widetilde{\omega}_f(a_j)} e^{\frac{\sigma-1}{\sigma} s_{f,j}} \right]^{\frac{\sigma}{\sigma-1}}. \tag{9}$$

The weight of a business function in TechFP, $\widetilde{\omega}_f(a_j)$, has one fixed component, Ω_f, and one that varies with the level of TechFP. If $\varepsilon_f > (1-\sigma)$ the weight of function f increases with TechFP, while if $\varepsilon_f < (1-\sigma)$ it decreases. If $\varepsilon_f = 1-\sigma, \forall f$, then the weights of the business functions in TechFP are constant and (9) becomes the standard homothetic CES aggregator.

Gross profits, $\Pi_j(a_j)$, are increasing and concave in a_j. The cost for establishment j with characteristics X of implementing technology sophistication $s_{f,j}$ in business function f is given by the convex function $C_f(s_{f,j}) = C_j C_{f,X} e^{s_{f,j}}$, where C_j is a establishment-specific component, and $C_{f,X}$ is a function-specific component that can potentially vary across groups of establishments based on their characteristics X. Given this set up, the technology choice

[41] In a clear abuse of notation, we denote the technology sophistication choice of establishment j in function f by $s_{f,j}$. So far, we have used the capital letter version $S_{f,j}$ to denote a generic measure of technology sophistication we construct from FAT which could be either $MAX_{f,j}$ or $MOST_{f,j}$. When bringing the model to the data in the next section, we will assume that the data counterpart of $s_{f,j}$ is $(MAXf, j + MOST_{f,j})/2$.

[42] In Appendix B.3 we introduce a version of this aggregator that incorporates business function-specific technical change (i.e. the productivity embodied in more sophisticated technologies differs across business functions). We derive the optimal adoption choices and discuss the implications this has for within-establishment variance in technology sophistication, the estimation of the slopes of technology curves and TechFP.

problem of establishment j is:

$$Max_{\{a_j,\{s_{f,j}\}\}}\Pi_j(a_j) - \sum_{f\in\aleph} C_f(s_{f,j}) \qquad s.t. \quad (8) \tag{10}$$

Proposition 1

A. The technology sophistication in business function f by establishment j is given by the following expression:

$$s_{f,j} = \underbrace{\sigma ln(\Pi_j'(a_j)) + \overbrace{\sigma ln\left(\frac{1-\sigma}{\sigma}\right) + ln\left(\Omega_f\right) + \varepsilon_f a_j - \sigma ln(\overline{\varepsilon}_j/\sigma)}^{\propto\partial a_j/\partial s_{f,j}}}_{\propto\partial\Pi_j/\partial s_{f,j}} - \underbrace{\sigma(ln(C_j) + ln(C_{f,X}))}_{\propto\partial C_f/\partial s_{f,j}} \tag{11}$$

where

$$\overline{\varepsilon}_j/\sigma \equiv \sum_f \frac{\varepsilon_f}{\sigma}\omega_{f,j} = \left[\left(\frac{\sigma}{1-\sigma}\right)\sum_f \frac{\varepsilon_f}{\sigma}\left(\Omega_f e^{\varepsilon_f a_j}\left(\frac{C_{f,X}C_j}{\Pi_j'(a_j)}\right)^{1-\sigma}\right)\right]^{1/\sigma} \tag{12}$$

with $\omega_{f,j} \equiv \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}s_{f,j}}$.

B. The technological factor productivity of establishment j is implicitly defined by the following expression:

$$\sum_{f=1}^{N_f} \Omega_f e^{\varepsilon_f a_j}\left(\frac{C_{f,X}C_j\overline{\varepsilon}_j}{\Pi_j'(a_j)\sigma}\right)^{1-\sigma}\left(\frac{1-\sigma}{\sigma}\right)^{\sigma} = 1. \tag{13}$$

Part A of Proposition 1 establishes the factors that influence the optimal level of technology sophistication implemented by an establishment in a business function. $s_{f,j}$ increases with factors that affect positively the marginal benefit of sophistication and decreases with factors that increase the marginal cost of technology sophistication. With the non-homothetic formulation (9), the marginal benefit of $s_{f,j}$ increases with the importance of the function for TechFP which is reflected in $\varepsilon_f * a_j$.

We can express the optimal technology choice (11) as

$$s_{f,j} = \kappa_j + \kappa_f + \varepsilon_f * a_j - \sigma ln(C_{f,X}). \tag{14}$$

Optimal technology sophistication depends on function-specific factors, κ_f, establishment-specific factors, κ_j, factors that are specific to the function and the class of establishment, $ln(C_{f,X})$, and the interaction between ε_f and a_j. Expression (14) is very similar to the heuristic specification for the technology curves in (**??**), with the key differences being the inclusion of the marginal cost term $-\sigma ln(C_{f,X})$, and that the technology curve is a function of a_j instead of S_j. This latter difference arises because, according to our model, the slope of the technology curve reflects how the technological importance of the function varies with the technological sophistication of a establishment measured by a_j. Variation in the slopes of technology curves results from differences across functions in how their importance changes with a_j.

Of course, a_j is an endogenous variable. Part B of Proposition 1 provides an expression for a_j in terms of the fundamentals of the model. Corollary 1 shows how these fundamentals affect the establishment's technological productivity using a first order approximation around the productivity of a establishment with average costs and marginal profits.[43]

Corollary 1 The TechFP index of establishment j, a_j, to a first order, can be expressed as follows

$$a_j \simeq a + \alpha_\Pi ln(\Pi'_j/\Pi') - \sum_f \alpha_{C_f} ln(C_{f,X}/C_f) - \alpha_C ln(C_j/\overline{C_j}) \tag{15}$$

where α_Π, $\left\{\alpha_{C_f}\right\}_{f=1}^{N_f}$, and α_C are positive constants.

Corollary 1 implies that the technological productivity of an establishment is increasing in the marginal profit, and decreasing with the establishment- and function-specific components of the marginal costs of adopting more sophisticated technologies.

4.2 Drivers of within-establishment variance

Expression (14) implies that the residual in sophistication after removing establishment and function effects is $u_{f,j} = \varepsilon_f * a_j - \sigma * ln(C_{f,X})$.

Proposition 2

The within-establishment variance in technology implied by our model is given by:

$$WVar_j = Var(u_{f,j}) = a_j^2 Var(\varepsilon_f) + \sigma^2 Var(ln(C_{f,X})) - \sigma a_j Cov(\varepsilon_f, ln(C_{f,X})) \tag{16}$$

Heterogeneity in ε_f causes the marginal benefit of sophistication to grow at different rates across functions as establishments try to reach higher levels of a_j. This induces dispersion in technology sophistication reflected in the first term in expression (16). Additionally, cross-function heterogeneity in the non-separable component of the marginal cost of adoption $(C_{f,X})$ also leads establishments to implement different sophistication levels across functions

[43] See Appendix B.2.

as captured by the second term in equation (16). Finally, the third term in (16) results from the potential covariance between these two forces.

Before proceeding to estimate the model, the Facts uncovered so far shed some light on the relative importance of these sources of within-establishment variance in technology sophistication. The variance of the marginal costs of adoption across functions is likely to be decreasing with establishment size as smaller establishments to suffer from limited technical capacity and access to finance and those limitations will amplify the cross-function variation in the marginal costs of sophistication between the functions that are capital/knowledge-intensive and those that are not intensive in capital/knowledge. The fact that we do not see an association in Table 8 between within-establishment variance in sophistication and size (Fact 6B) suggests that heterogeneity in adoption costs is not the main driver of within-establishment variance in technology. Conversely, the strong association between WVAR and average sophistication in the establishment (Fact 6A) as well as the heterogeneity in the slopes of the technology curves (Fact 7) is consistent with the presence of non-homotheticities in TechFP which imply that it is optimal for establishments to increase the difference in technology sophistication across business functions as they implement higher levels of TechFP.

5 Estimation of technology curves and technological productivity

We take advantage of the model' structure and the FAT dataset to obtain estimates of the slope of technology curves for each business function and the technological factor productivity of each establishment. In contrast with measures of TFP, our estimates of TechFP rely only on direct information of the technologies used by an establishment. By combining the estimates of the slopes of technology curves and TechFP we will be able to explore the sources of variation in technology sophistication across business functions.

5.1 Estimation strategy

The starting point of our estimation is the optimal technology sophistication levels selected by an establishment as stated in part A of Proposition 1. Equation (11) implies the following econometric specification:

$$s_{f,j} = \kappa_j + \kappa_f + \varepsilon_f * a_j + \beta_f * X_j + u_{f,j} \tag{17}$$

where κ_j and κ_f are establishment and function effects, X_j is a vector of relevant establishment characteristics that affect the differential marginal cost of sophistication in business function f, ε_f is the technology elasticity of function f, a_j is the technological productivity of establishment j and $u_{f,j}$ is classical measurement error.

Specification (17) is a linear mixed model, which is a class of models that have both fixed and random effects.[44] The estimation of (17) delivers both estimates of the slopes of the technology curves, ε_f, and the technological productivity of each establishment, a_j.

To estimate (17), we follow a two-step approach similar to how Aguiar and Bils (2015) simultaneously estimate the slopes of Engel curves and household expenditure from the consumer expenditure survey (CEX). First, we estimate ε_f replacing a_j by a proxy $\overline{a}_j$ which is an affine transformation of a_j plus classical measurement error:

$$s_{f,j} = \kappa_j + \kappa_f + \varepsilon_f * \overline{a}_j + \beta_f * X_j + v_{f,j} \tag{18}$$

Under this assumption, the OLS estimates $\widehat{\varepsilon}_f$ are unbiased estimates of ε_f (up to a constant scaling factor). In a second step, we use the values of $\widehat{\varepsilon}_f$ to estimate the establishment-level technological factor productivity, a_j, by estimating the mixed model

$$s_{f,j} = \kappa_j + \kappa_f + \widehat{\varepsilon}_f * a_j + \beta_f * X_j + w_{f,j} \tag{19}$$

where κ_j and a_j are (potentially correlated) random establishment effects, and $w_{f,j}$ is classical measurement error. The random slope ($\widehat{a_j}$) is an unbiased estimate of a_j, up to a scaling constant.[45] Furthermore, because the constant scaling factors of the estimates $\widehat{a_j}$ and $\widehat{\varepsilon}_f$ are inversely related, the contribution of the technology curve to within-establishment variance is equal to the variance of $\widehat{a_j} * \widehat{\varepsilon}_f$ relative to the variance of $s_{f,j} - \kappa_j - \kappa_f$.

Our proxy of a_j, $\overline{a}_j$, is based on the first-order approximation of establishment technological productivity (a_j), implicitly defined by (8), around the productivity of an average establishment that implements the average technology sophistication level in each business function. As shown in the appendix, a_j can be expressed as the following linear function of average establishment sophistication and within-establishment dispersion in sophistication:

$$a_j \simeq \alpha_0 + \frac{1-\sigma}{\overline{\varepsilon}} \left(s_j + \varphi \sqrt{WVAR_j} \right) \tag{20}$$

where φ is a constant, $\overline{\varepsilon}$ is the average of ε_f, and $WVAR_j$ is the within-establishment

[44]See for example Searle, Casella and McCulloch (1992), McCulloch, Searle and Neuhaus (2008).

[45]In addition, we obtain estimates of the random intercept ($\widehat{\kappa_j}$) and the variance covariance matrix between the random intercept and slopes. The post-estimation commands use this variance covariance matrix to compute the best linear unbiased predictor (BLUP) of $\widehat{a_j}$ which shrinks the estimate to adjust for measurement error (Robinson, 1991).

variance of establishment j.[46] Accordingly, we set $\overline{a}_j = s_j + \varphi * \sqrt{WVAR_j}$. That is, the average of technology sophistication in the establishment plus a constant (to be estimated) times the within-establishment dispersion in technology. We assume that the log of the number of employees proxies for the establishment characteristics X_j that influence differentially the marginal cost of sophistication across functions.

5.2 Analysis of Estimates

We implement this estimation procedure on the FAT dataset using as counterpart to $s_{f,j}$ the average of MAX and MOST in each establishment and business function (i.e., $(MAX_{f,j} + MOST_{f,j})/2$). The procedure yields estimates of the slopes of technology curves for each business function, $\widehat{\varepsilon}_f$, and of the TechFP for each establishment, $\widehat{a}_j$. Table C.12 and Table 9a report the estimates $\widehat{\varepsilon}_f$ and summary statistics of their distribution. Figure 8 and Table 9b present the histogram and summary statistics of the distribution of $\widehat{a}_j$ across establishments. For comparison purposes, Table 9c reports summary statistics of the (log) value added per worker across establishments.[47]

Consistent with the visual assessment made in Fact 7, Table C.12 and Table 9a show that there is a large dispersion in $\widehat{\varepsilon}_f$ across functions. For example, the ratio of the estimates of the technology elasticity in business administration and payments (the two GBFs with highest and lowest estimates) is 3.5. If we order all the business functions by their technology elasticity, the ratio of the estimates of the business functions in the $90th$ to the 10^{th} percentiles is 4.32.

There is also a large dispersion in TechFP across establishments. The standard deviation of $\widehat{a}_j$ represents 45% of the dispersion in value added per worker. The distribution of TechFP is more skewed than the distribution of labor productivity and both roughly have no excess kurtosis (i.e., kurtosis is around 3).

Fit of the technology curve

We use the estimates of (19) to study the sources of within-establishment variance in technology sophistication. Table 10 reports the variance of each of the components of $s_{f,j}$ in (17) as well as the within-establishment variance in technology sophistication. The variance of the establishment effects (κ_j) represents only 4.3% of the variance of technology sophistication net of function effects ($Var(s_{f,j} - \widehat{\kappa}_f)$). The component that reflects the function-specific

[46] Specifically, $\overline{\varepsilon} = \sum_{f=1}^{N_f} \omega_f \varepsilon_f$ with $\omega_f = \left(\xi_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma} s_f}$. $\varphi = \sqrt{Var(\omega_f)} * \sum_{f=1}^{N_j} Corr(\omega_f, s_{f,j})/N_j$.

[47] We exclude the top and bottom 1% of establishments in the distribution of value added per worker as we consider them as outliers.

marginal costs of sophistication ($Var(\widehat{\beta}_f * X_j)$) represents 6% of the within-establishment variance ($Var(s_{f,j} - \widehat{\alpha}_j - \widehat{\alpha}_f)$). With the estimates $\widehat{\varepsilon}_f$ and $\widehat{a}_j$ we compute the level of technological sophistication predicted by the technology curves ($\widehat{\varepsilon}_f * \widehat{a}_j$). The next Fact compares its variance to the within-establishment variance in technology sophistication.[48]

Fact 9.

The technology curves account for 43% of the within-establishment variance in technology sophistication.

It is quite remarkable that a simple linear representation such as the technology curve can accounts for such a significant fraction of the variance on technology sophistication, given the magnitude of the variation in technology sophistication across the business functions of an establishment (Fact 4). Next, we discuss the interpretation of our estimates.

5.3 Discussion

What drives the heterogeneity in the slopes of technology curves? Is it variation in ε_f across functions as implied by the model we have developed in section 4 or is it due to differences across functions in the technologies that form the grid?

We explore these questions in two different ways. We first focus on functions that are relevant in multiple sectors and allow the slopes of the technology curves to vary across sectors. Table C.13, in the appendix, reports the estimates of $\widehat{\varepsilon}_{f,s}$ for each of the seven general business functions in each of the 13 2-digit sectors (the 11 sectors with SSBFs, plus other manufacturing and other services), and for fabrication in the six manufacturing sectors. Table 11 reports, for each general business function, the difference between the highest and lowest estimates of $\widehat{\varepsilon}_{f,s}$ across the 2-digit sectors. Fact 10 reports our key finding.

Fact 10.

In all the business functions that are relevant across multiple sectors, we find significant differences across sectors in the estimates of $\varepsilon_{f,s}$.

The difference between the highest and lowest estimates of $\varepsilon_{f,s}$ across sectors range from 1.1 for business administration to 0.18 for fabrication, with an average difference for the general business functions of 0.78. This magnitude is comparable with the difference between

[48]Note that we do not attribute any part of the covariance between the technology curve and residual technology sophistication to the technology curve. In this respect our statistic of the contribution of the technology curve is quite conservative.

the highest and lowest estimates of ε_f across the general business functions in the baseline reported in Table C.12 where ε_f is restricted to be the same across sectors.

Fact 10 sheds light on the origin of the variation in the slopes of technology curves. Since for each business function relevant in multiple sectors the technologies in the grid are kept constant, the cross-sector variation in $\widehat{\varepsilon}_{f,s}$ can only reflect sectoral differences in how the importance of the function changes with a_j.

Complementarily, we can understand the sources of variation in the slopes of technology curves by explicitly introducing business function-specific technological change in the model developed in section 4.[49] In particular, consider the following implicit aggregator of the establishment technology choices across functions, $s_{f,j}$

$$\sum_{f=1}^{N_f} \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}\gamma_f s_{f,j}} = 1 \tag{21}$$

where γ_f is the rate of technical change in function f. Without lost of generality, we assume that the average of γ_f across business functions is 1, but it can differ across functions to reflect cross-function differences in the productivity embodied in more sophisticated technologies.

The optimal technology choices of the establishment, after making a first order approximation around $\gamma_f = 1$, can be expressed as

$$s_{f,j} \simeq \widetilde{\kappa}_f + \left(\widetilde{\kappa}_j + \varepsilon_f a_j - \sigma ln(C_{f,X})\right)\left(1 - (1-\sigma)\left(\gamma_f - 1\right)\right). \tag{22}$$

The main difference between (22) and (11) is that now there is an interaction between the establishment random effect ($\widetilde{\kappa}_j$) and function-specific factors, in particular, the rate of technological progress (γ_f). This interaction affects the estimate of the slope of the technology curve, $\widehat{\rho}_f$ that now is given by

$$\widehat{\rho}_f = \varepsilon_f \left(1 - (1-\sigma)\left(\gamma_f - 1\right)\right) - \frac{Cov(\overline{a}_j, \widetilde{\kappa}_j)}{Var(\overline{a}_j)} (1-\sigma)\left(\gamma_f - 1\right). \tag{23}$$

As in the baseline model, the first term of (23) reflects the fact that optimal technology use dictates establishments that want to attain a greater TechFP to choose disproportionately higher sophistication levels in business functions with greater ε_f. But now, due to the interaction between $\widetilde{\kappa}_j$ and γ_f in (22) and the positive covariance between $\overline{a}$ and $\widetilde{\kappa}_j$, $\widehat{\rho}_f$ has a second term that reflects that establishments with higher $\widetilde{\kappa}_j$ want to use relatively less sophisticated technologies in functions with faster technological progress (γ_f). As a result

[49] See Appendix B.3 for the derivations.

the estimated slope of the technology curve is lower than in the baseline in functions with $\gamma_f > 1$ and higher in functions with $\gamma_f < 1$.

Which of these two mechanisms in (23) is driving the variation in the estimated slopes of technology curves across functions? Consider business administration and payments, with the slope in the technology curve of the former being 3.5 times larger than the slope of the later. Is that difference due to the fact that, in establishments with higher TechFP, the sophistication of technologies in business administration is more important than the sophistication of payment technologies(i.e. higher ε_f)? Or is it because the productivity embodied in more sophisticated technologies in payments is greater than in more sophisticated technologies in business administration (i.e., lower γ_f)? In this case, it seems to us clear that the steeper slope of the technology curve for business administration reflects its higher ε_f rather than a lower γ_f, as the productivity embodied in payments is clearly smaller than in business administration.[50]

More generally, two observations from our estimates limit the magnitude of the second term in (23). First, the fact that the variance of $\widehat{\kappa}_j$ is small (Table 10).[51] Second, the fact that the point estimates of the slopes of the technology curves ($\widehat{\rho}_f$) for all functions are positive.[52]. Therefore, it seems reasonable to conclude that the heterogeneity in the slopes of technology curves is largely driven by cross-function variation in ε_f as predicted by our baseline model.

The estimate of TechFP in establishment j is $\widehat{a}_j = Cov(\widehat{\rho}_f D_j, s_{f,j})/Var(\widehat{\rho}_f)$ where D_j is a dummy for establishment j. In the extended model, $\widehat{a}_j$ is equal to:

$$\widehat{a}_j = a_j * (\omega_1 + \omega_3) + \frac{\widetilde{\kappa}_j}{Cov(\overline{a}_j, \widetilde{\kappa}_j)} * (\omega_2 + \omega_3), \tag{24}$$

where $\omega_1 = \frac{Var(\varepsilon_f(1-(1-\sigma)(\gamma_f-1)))}{Var(\widehat{\rho}_f)}$, $\omega_2 = \frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)^2(1-\sigma)^2 Var(\gamma_f)}{Var(\widehat{\rho}_f)}$, and $\omega_3 = 1 - \omega_1 - \omega_2$.[53]

There are two differences between (B.32) and the estimates of TechFP in the baseline model. First, there is an attenuation bias reflected in the coefficient $(\omega_1 + \omega_3)$ for a_j smaller than 1 which originates from the fact that the covariance between $\widehat{\rho}_f$ and $\varepsilon_f * a_j$ is lower than

[50] In general, it is not easy to rank functions by γ_f. Most functions in the grid start with relatively manual processes and transition to more digital technologies as we move to more sophisticated technologies in the grid.

[51] We formalize this argument in Part A of Proposition 3, below.

[52] See Table C.12.

[53] With $\omega_1 > 0$ and $\omega_2 > 0$ and

$$Var(\widehat{\rho}_f) = Var(\varepsilon_f (1 - (1-\sigma)(\gamma_f - 1))) + (Cov(\overline{a}_j, \widetilde{\kappa}_j)(1-\sigma))^2 Var(\gamma_f) \tag{25}$$
$$-Cov(\overline{a}_j, \widetilde{\kappa}_j)(1-\sigma) Cov(\gamma_f, \varepsilon_f (1 - (1-\sigma)(\gamma_f - 1))). \tag{26}$$

in the baseline because of the second term in (23). Second, there is a second term that reflects the covariance between $\widehat{\rho}_f$ and $(\gamma_f - 1) * \widetilde{\kappa}_j$. As a result of the second term, the estimate of TechFP does not reflect only a_j but also the establishment random effect $\widehat{\kappa}_j$. Proposition 3 states sufficient conditions for the variance of $\widehat{a}_j$ to be equal to the cross-establishment variance of actual TechFP.

Proposition 3 *As $Var(\widetilde{\kappa}_j)$ tends to 0,*

A. $Var(\widehat{\rho}_f)$ tends to $Var(\varepsilon_f (1 - (1-\sigma)(\gamma_f - 1))$,
B. $E(\widehat{a}_j|a_j)$ tends to a_j,
C. $Var(\widehat{\rho}_f * \widehat{a}_j)$ tends to $Var(\varepsilon_f (1 - (1-\sigma)(\gamma_f - 1) * a_j)$.

Proposition 3 shows that, as the variance of the establishment effects, $\widetilde{\kappa}_j$, becomes small, the slopes of the technology curves reflect only the increase in technology sophistication, $s_{f,j}$, associated with an increase in establishment-level TechFP, a_j. Given that, the estimate of TechFP is unbiased and the variance of $\widehat{a}_j$ converges to the variance of a_j. Finally, the within-establishment variance in technology sophistication explained by the estimated technology curves is equal to the variance in sophistication induced by TechFP.[54]

Table 6 and Table 9b shows that the estimated variance of $\widetilde{\kappa}_j$ is 15 times smaller than the variance of $\widehat{a}_j$, and therefore, the condition in Proposition 3 roughly holds. This implies that the heterogeneity in the slopes of technology curves is largely driven by variation in ε_f, and that the estimates of TechFP do largely reflect the technological index of the establishment.

6 TechFP

The estimates of TechFP allow us to investigate three important question. First, how much do we miss by proxying TechFP by traditional measures of technology. Second, what establishment characteristics are correlated with TechFP. Third, what fraction of the cross-establishment dispersion in productivity can be accounted for by TechFP.

6.1 TechFP and other technology measures

We compare TechFP with various measures of technology at the establishment, T_j, that differ in their comprehesiveness and detail in terms of the number of technologies, whether they reflect just the presence of technology or incorporate information on the intensity of use, and whether they connect the technologies to the business functions where they are

[54]Trivially, the same results hold if $Var(\gamma_f)$ tends to 0, as this brings us back to the baseline model.

used. To explore how much of the cross-establishment variation in TechFP is captured by each of these measures, we estimate the following specification

$$\widehat{a}_j = \alpha_0 + \alpha_1 * T_j + u_j. \tag{27}$$

Our first technology measure reflects the standard approach of measuring establishment technology by the presence/absence of some general purpose technologies. We replicate this approach by constructing a binary measure that is equal to 1 if the establishment has computers, access to electricity and internet, and zero otherwise. Then we move to study the relationship between TechFP and sophistication measures in specific business functions, business administration and payments. Finally, we study which dimensions of our comprehensive information on technology sophistication are more relevant to explain the cross-establishment variation in TechFP, $MOST_j$ vs MAX_j, average technology sophistication vs. dispersion in sophistication across business functions, technology sophistication in general vs. sector-specific business functions. Table 11 reports the estimates, and Fact 11 summarizes our key findings.

Fact 11.

A. Measures of the presence of general technologies in the establishment such as computers, internet or electricity explain a small part (30%) of the cross-establishment variance in TechFP.

B. The cross-establishment variance of TechFP accounted for the variance in measures of technology sophistication in one business function differs very much across business functions.

C. Statistics that reflect the first and second moments of technology sophistication across all the business functions of the establishment jointly account for between 73% and 82% of the cross-establishment variance in TechFP.

D. MOST is more relevant than MAX for the cross-establishment variance in TechFP, while average technology sophistication measures are more relevant than within-establishment dispersion in technology sophistication.

E. The technological sophistication of general business functions accounts for a larger share of the cross-establishment variance in TechFP than the sophistication of sector-specific business functions.

F. However, the relevance of sophistication in sector-specific vs. general business functions for TechFP varies across broad sectors. Sector-specific technologies are most relevant in agriculture, while general business functions are most relevant in services and manufacturing.

Fact 11 shows the relevance of using comprehensive technological measures to accurately reflect the technological landscape of a establishment. Additionally, parts B and E show the importance of connecting technology to the business functions where they are used as the relevance of technology sophistication for TechFP varies across functions, and to not restrict attention to just a few business functions. Part C highlights the importance of going beyond first or second moments as the aggregators that produce TechFP measures from the vector of function-level technology sophistication are very non-linear. Finally, part F shows shows that there are important differences in the relevance of business functions across sectors.[55]

6.2 TechFP and establishment characteristics

We next explore the association between technological factor productivity and firm characteristics by running the following regression:

$$\widehat{a}_j = \alpha_s + \alpha_c + \beta * X_j + u_j \tag{28}$$

where α_c and α_s are country and sector dummies and the vector of establishment characteristics, X_j, includes the fraction of workers with college education, an index of management practices following Bloom et al. (2019),[56] the multi-establishment, exporter and multinational status, and age categories. Table 13 reports the estimates for both estimates of technological productivity.

Fact 12.

> Technological factor productivity is positively associated with the share of workers with college education, the index of good management practices, exporting, multi-establishment and multi-national status. TechFP is insignificantly associated with establishment age.

6.3 Technology and Productivity

We conclude our exploration by studying the cross-establishment association between labor productivity and our estimates of technological factor productivity.

[55] This is consistent with variation in the slopes of technology curves we find across sectors in business functions that are relevant in different sectors (e.g., fabrication, or packaging). See Table C.12.

[56] See ?? of the appendix for more details.

To this end we estimate versions of the following productivity regression:

$$y_j = \alpha_s + \alpha_c + \beta * k_j + \rho * h_j + \gamma * \widehat{a}_j + \theta * X_j + u_j \tag{29}$$

where the dependent variable is the log of nominal value added per worker[57], k_j is the log of the book value of capital per worker, h_j is the percentage of workers in the establishment with a college degree, α_s, and α_c are sector and country dummies, $\widehat{a}_j$ is the TechFP of establishment j, X_j is a vector of additional controls, and u_j is classical measurement error. Table 14 reports the estimates. In column 1 we only include sector and country fixed effects, and physical and human capital. Column 2 adds TechFP. Column 3 excludes the country effects. Column 4 the relationship between TechFP and establishment productivity to differ across 1-digit sectors, and Column 5 introduces other potential drivers of labor productivity such as the management score and the computer-internet-electricity dummy.[58]

Fact 13.

A. There is a strong association between productivity and TechFP across establishments.

B. The association is strongest in agricultural establishments and weakest for those in services.

C. The association is robust to controlling for measures of physical capital, human capital and management practices in the establishment.

We use the estimates from equation (29) to assess the fraction of the cross-establishment variance in productivity that can be accounted by TechFP. [59] Fact 14 reports our findings.

Fact 14.

A. TechFP accounts for 15% of the dispersion in productivity between establishments at the 10th and 90th percentiles in the distribution of productivity. As a way of comparison, management practices account for 5%.

[57]Nominal value added are sales minus the costs of materials.

[58]In regressions reported in the online appendix we have also controlled for other characteristics such as the exporter, multi-national and multi-establishment status. The estimates are robust to these additional controls.

[59]We first compute the residual productivity for all firms by regressing productivity on the country and sector dummies, and the measures of physical and human capital, and then computing the residual. We do the same for TechFP and the management practices scores. For each of these three residuals, we calculate the gap between the 10^{th} and 90^{th} percentiles. We then multiply the coefficients of TechFP and management in column 5 of Table 14 times the 10-90 gaps in the residuals for each variable and divide that by the residual of labor productivity. The resulting number is the percentage of the cross-establishment dispersion in productivity account for by TechFP and management practices, respectively. We conduct a similar exercise using column 6 to assess the contributions of TechFP to the dispersion in productivity by one-digit sector.

B. There are differences across sectors in the contribution of TechFP to establishment productivity. In agricultural establishments it accounts for 30% of the dispersion in productivity, in manufacturing for 21% and in services for 13%.

Fact 14 shows that TechFP is an important driver of cross-establishment variation in productivity. Interestingly, we find that TechFP is most relevant for establishments that operate in agriculture, and least for establishments in the service sector.

7 Conclusions

In this paper, we have taken a deep dive into technology. Our exploration involves several distinct layers. First, we have created a new, comprehensive dataset that thoroughly characterizes the technology used by establishments in each of the key business functions. Key to the assembly of the FAT dataset is the construction of the grid that specifies the key functions for an establishment based on its sector and the range of technologies that can be used to complete the tasks associated to each business function.

Second, we have studied the technologies used by an establishment in a business function by constructing measures of the sophistication (relative to the frontier) of some key technologies employed by the establishment: the technology it uses most widely in the function and the most sophisticated among the technologies it uses in the function.

Third, we have studied the dispersion in technology sophistication across the business functions of an establishment. A key discovery is that there is larger variance in technology sophistication within an establishment than across establishments. Additionally, the within-establishment variance of technology sophistication is strongly positively associated with the average sophistication of technologies in the establishment, while other establishments characteristics such as size or age are uncorrelated with the variance across business functions of an establishment.

Fourth, we have developed a model to better understand the drivers of technology adoption across business functions. The model allows for both heterogeneity in the marginal value and marginal costs of implementing more sophisticated technologies to affect technology choices. Additionally, the model introduces the concept of technological factor productivity, which is an index that aggregates all the technologies used by the establishment across functions and that is a sufficient statistic of the relationship between technology and productivity at the establishment level.

The model predicts the existence of a parsimonious cross-establishment relation between the sophistication of technology in a business function and the technological factor productivity of the establishment. We have named this relation the technology curve. Fifth, exploiting the heterogeneity across functions in the slope of technology curves, we have developed an structural estimation strategy that allows us to identify the technological factor productivity using information of the technology choices of each establishment across its business functions.

Sixth, we have analyzed the estimates of technological factor productivity across establishments and documented its relation to various moments of the distribution of technology sophistication measures in the establishment as well as to establishment characteristics.

Seventh, we have studied the relationship between value added per worker and technological factor productivity across establishments, and have shown that variation in technological factor productivity accounts for 15% of the cross-establishment variance in value added per worker observed in the data, but this share is as high as 31% in agricultural establishments.

References

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica*, 60(2): 323–351.

Aguiar, Mark, and Mark Bils. 2015. "Has Consumption Inequality Mirrored Income Inequality?" *American Economic Review*, 105(9): 2725–56.

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.

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw. 2007. "How Does Information Technology Affect Productivity? Plant-level Comparisons of Product Innovation, Process Improvement, and Worker Skills." *The Quarterly Journal of Economics*, 122(4): 1721–1758.

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.

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 Management 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): 1648–83.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen. 2016. "International Data on Measuring Management Practices." *American Economic Review*, 106(5): 152–56.

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 Weerdt. 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, Marcio Cruz, and Kyung Min Lee. 2020. "Anatomy of Technology in the Firm." *NBER Working Paper Series, n. 28080.*

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. 2000. "An Uncertainty-Driven Theory of the Productivity Slowdown in Manufacturing." PhD diss. Harvard University.

Comin, Diego, and Bart Hobijn. 2010. "An Exploration of Technology Diffusion." *American Economic Review*, 100(5): 2031–59.

Comin, Diego, and Martí Mestieri. 2018. "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *American Economic Journal: Macroeconomics*, 10(3): 137–78.

Davies, Stephen. 1979. *The Diffusion of Process Innovations.* Cambridge University Press.

Gary, Pike R. 2007. "Adjusting for Nonresponse in Surveys." In *Higher Education: Handbook of Theory and Research.* 411–449. Springer.

Griliches, Zvi. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica*, 25(4): 501–522.

Gupta, Apoorv, Jacopo Ponticelli, and Andrea Tesei. 2020. "Information, Technology Adoption and Productivity: The Role of Mobile Phones in Agriculture." National Bureau of Economic Research.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review*, 109(3): 1032–79.

Hubbard, Thomas N. 2003. "Information, Decisions, and Productivity: On-board Computers and Capacity Utilization in Trucking." *American Economic Review*, 93(4): 1328–1353.

Jorgenson, Dale W, Mun S Ho, and Kevin J Stiroh. 2008. "A Retrospective Look at the US Productivity Growth Resurgence." *Journal of Economic Perspectives*, 22(1): 3–24.

Jorgenson, Dale W, Mun S Ho, Kevin J Stiroh, et al. 2005. *Productivity, Volume 3: Information Technology and the American Growth Resurgence.* MIT Press Books.

Lagakos, David, and Michael E Waugh. 2013. "Selection, Agriculture, and Cross-country Productivity Differences." *American Economic Review*, 103(2): 948–80.

Mansfield, Edwin. 1963. "Intrafirm Rates of Diffusion of an Innovation." *The Review of*

Economics and Statistics, 348–359.

Oliner, Stephen D, Daniel E Sichel, and Kevin J Stiroh. 2007. "Explaining a Productive Decade." *Brookings Papers on Economic Activity*, 2007(1): 81–137.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of political Economy*, 98(5, Part 2): S71–S102.

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.

Schroeder, Carsten, and Shlomo Yitzhaki. 2017. "Revisiting the Evidence for Cardinal Treatment of Ordinal Variables." *European Economic Review*, 92: 337–358.

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.

Van Ark, Bart, Mary O'Mahoney, and Marcel P Timmer. 2008. "The Productivity Gap Between Europe and the United States: Trends and Causes." *Journal of Economic Perspectives*, 22(1): 25–44.

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.

World Bank. 2009. "Enterprise Survey and Indicator Surveys: Sampling Methodology." *The World Bank, Manuscript.*

Figures

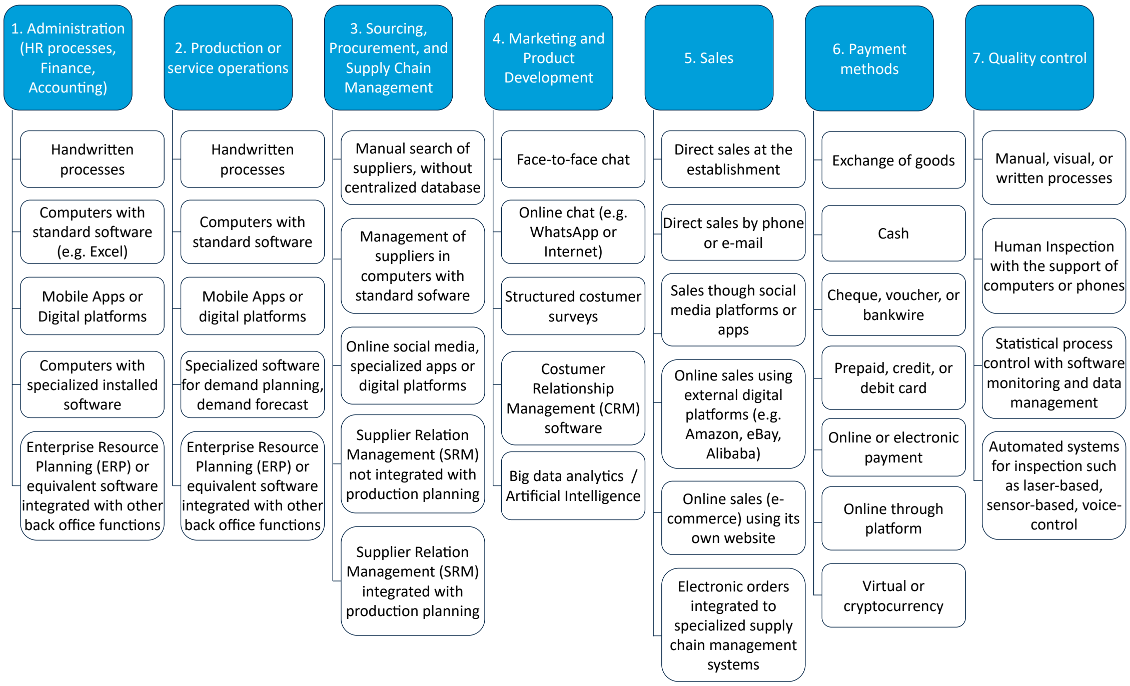


Figure 1: General Business Functions and Their Technologies

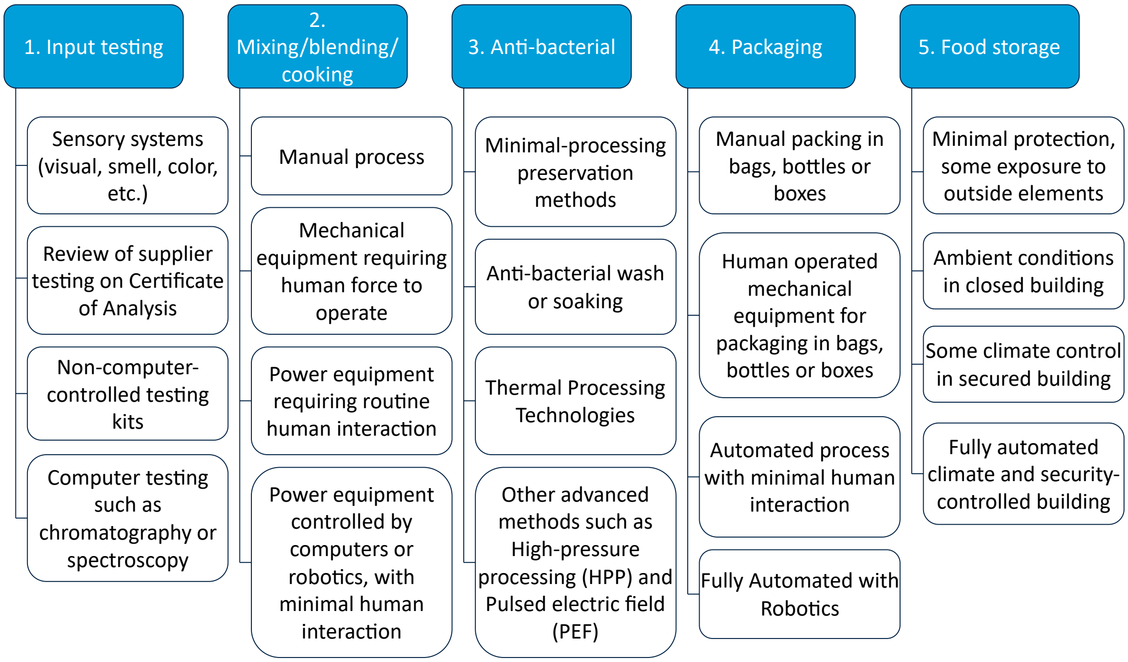


Figure 2: Sector Specific Business Functions and Technologies in Food Processing

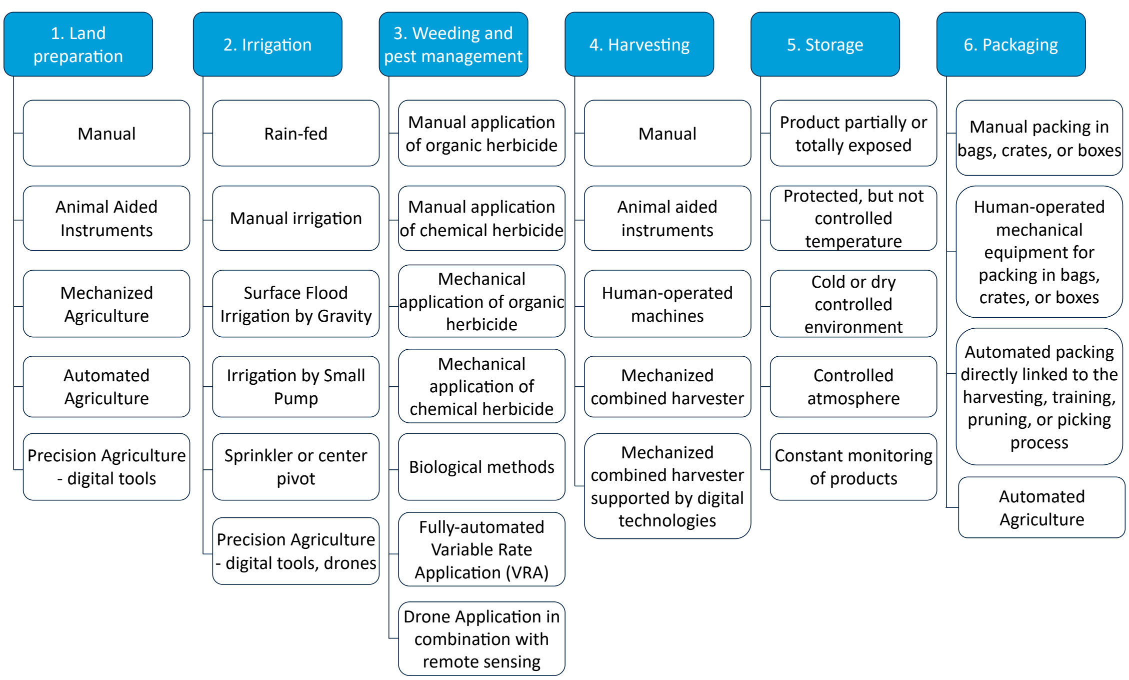
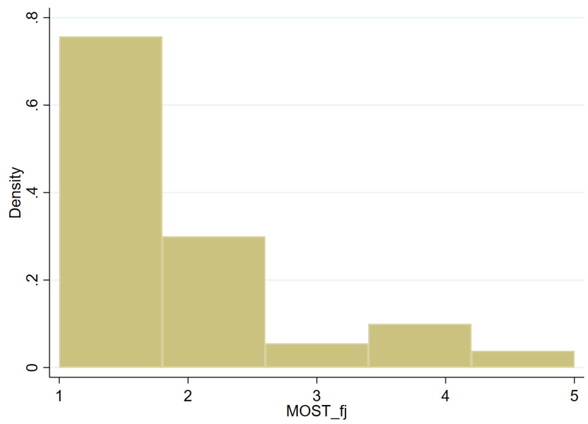
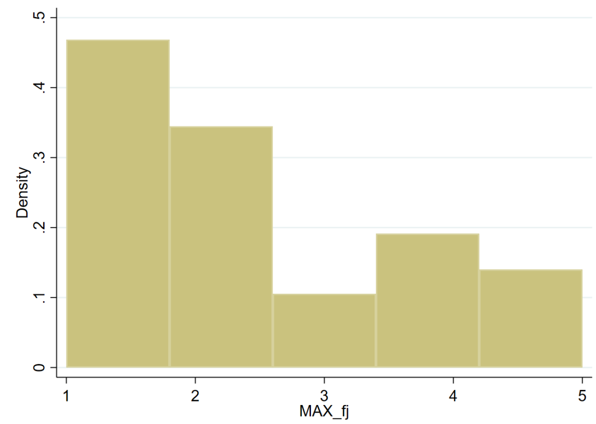


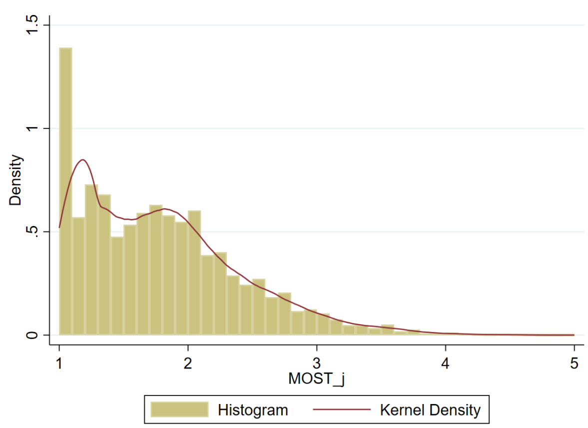
Figure 3: Sector Specific Business Functions and Technologies in Agriculture



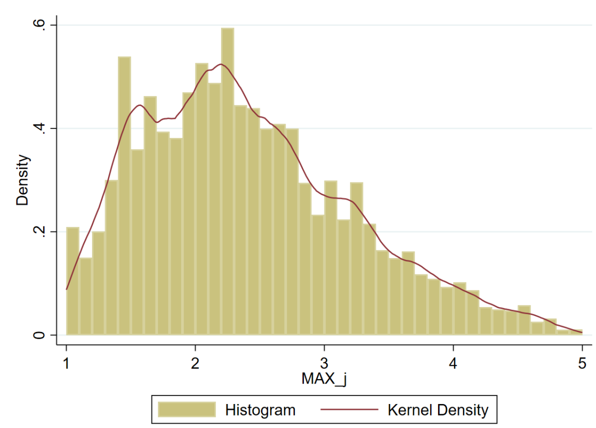
(a)



(b)



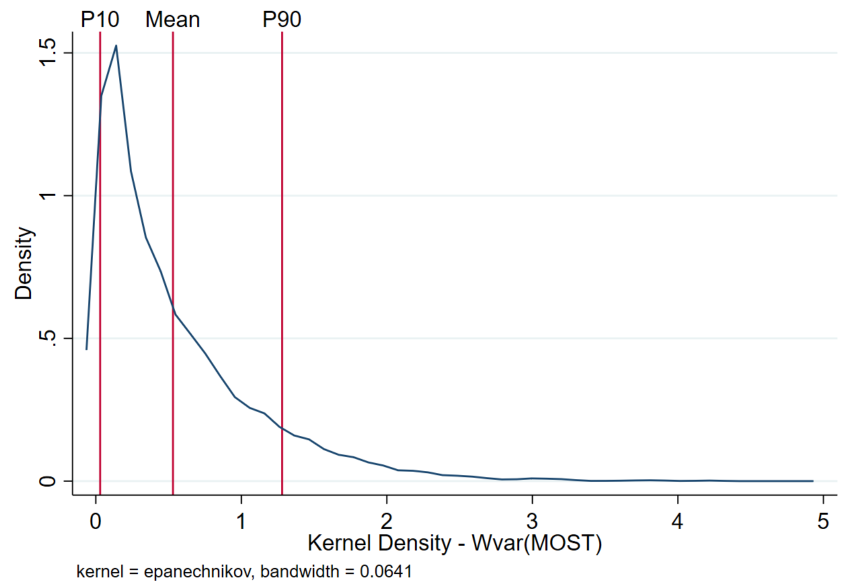
(c)



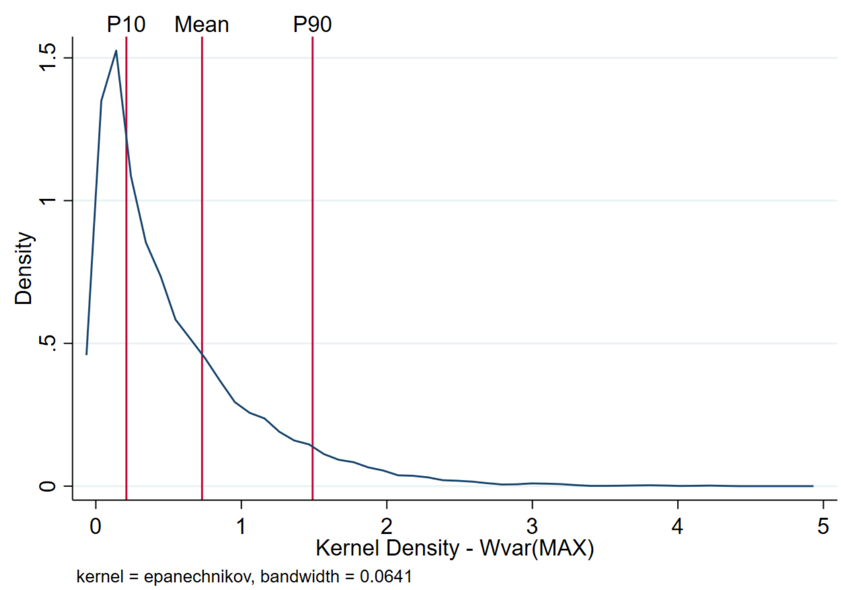
(d)

Figure 4: Distribution of MOST and MAX at the business function and establishment level

Note: For panels (a) and (b), densities are computed using weights for business function/establishment. For panels (c) and (d), densities are computed using establishment weights.



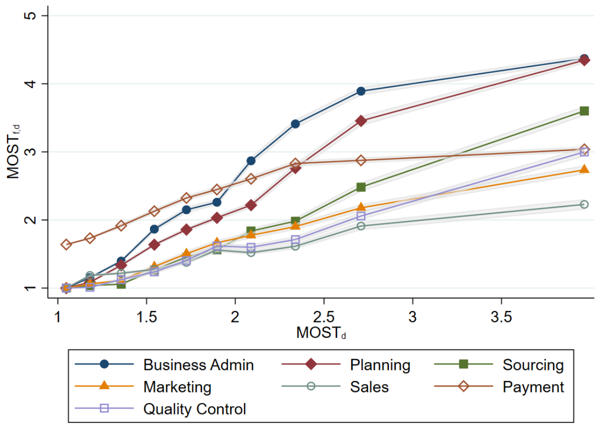
(a)



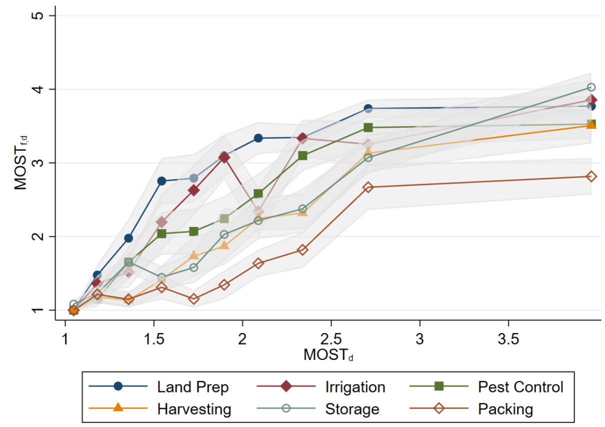
(b)

Figure 5: Distribution of Wvar across establishments

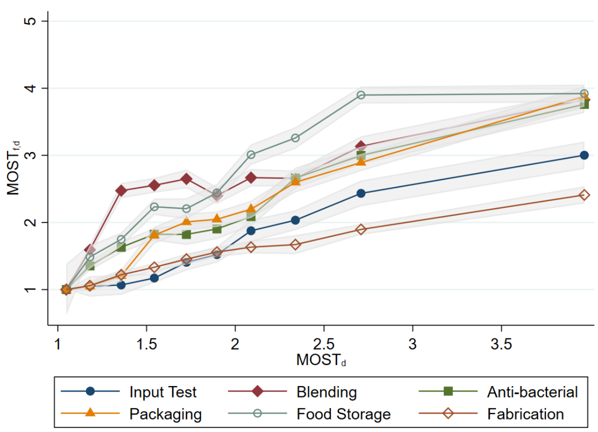
Note: Estimates using sampling weights for the establishments.



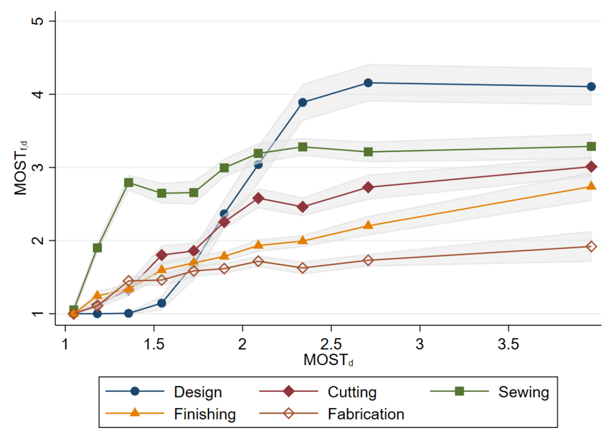
(a) General Business Functions



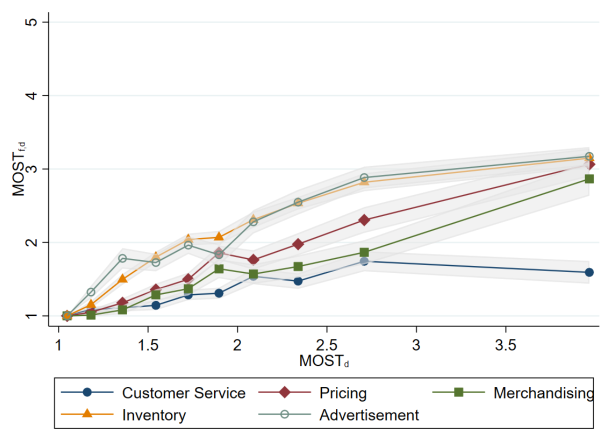
(b) Agriculture (crops)



(c) Food Processing



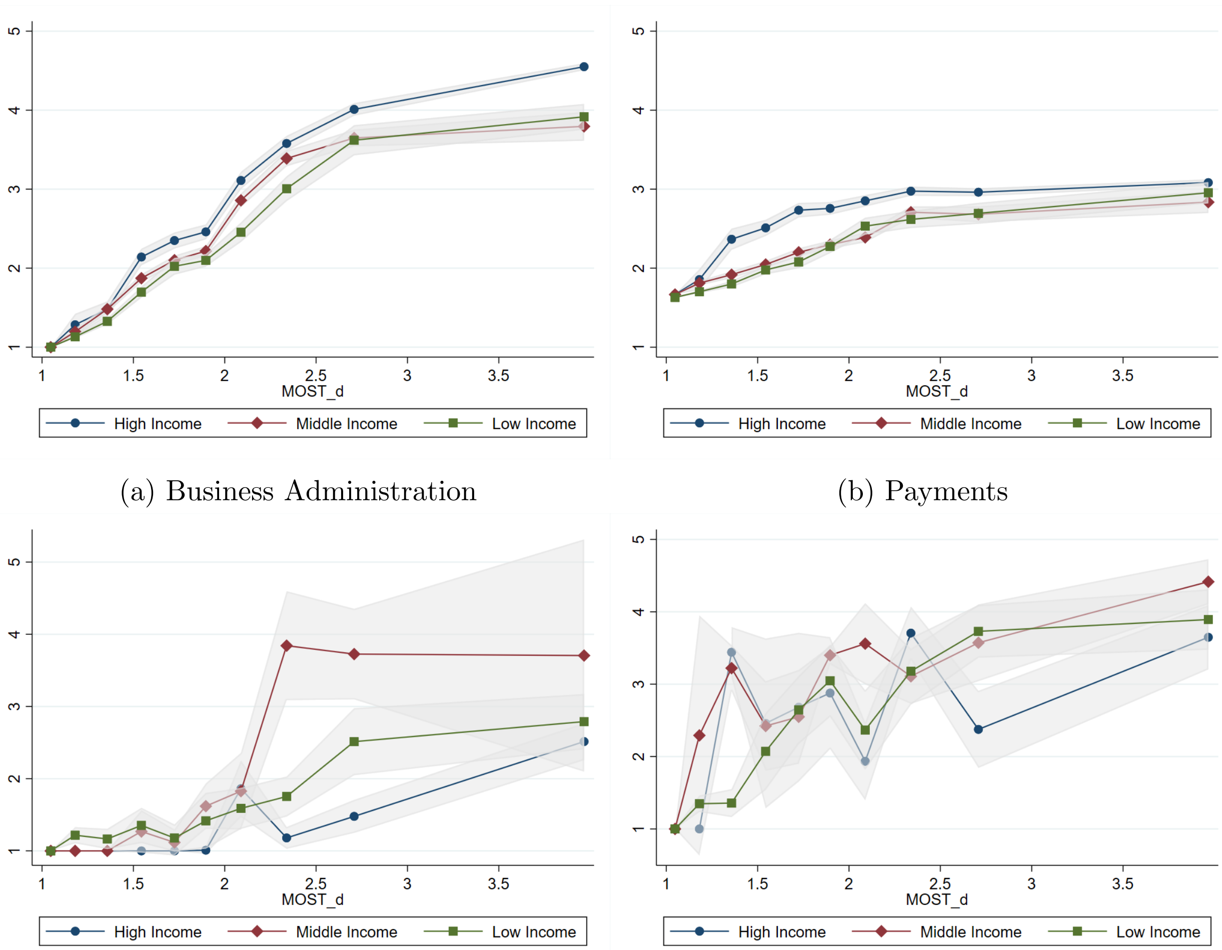
(d) Wearing apparel



(e) Retail

Figure 6: The Technology Curve, $MOST_{f,j}$ vs. $MOST_j$ by Deciles

Note: $MOST_{f,d}$ (vertical axis) is the average value of $MOST_{f,j}$ for the establishments in the d decile of $MOST_j$. $MOST_d$ (horizontal axis) is the average value of $MOST_j$ for the establishments in the d decile of $MOST_j$. All averages computed using establishment sampling weights.



(a) Business Administration

(b) Payments

(c) Packing

(d) Irrigation

Figure 7: Technology curves for three country groupings based on per capita income
Note: Each panel plots the technology curves for one business function. In each panel there are three curves that correspond to high- medium and low-income country groups. The shaded area represent the 90% confidence interval for each curve.

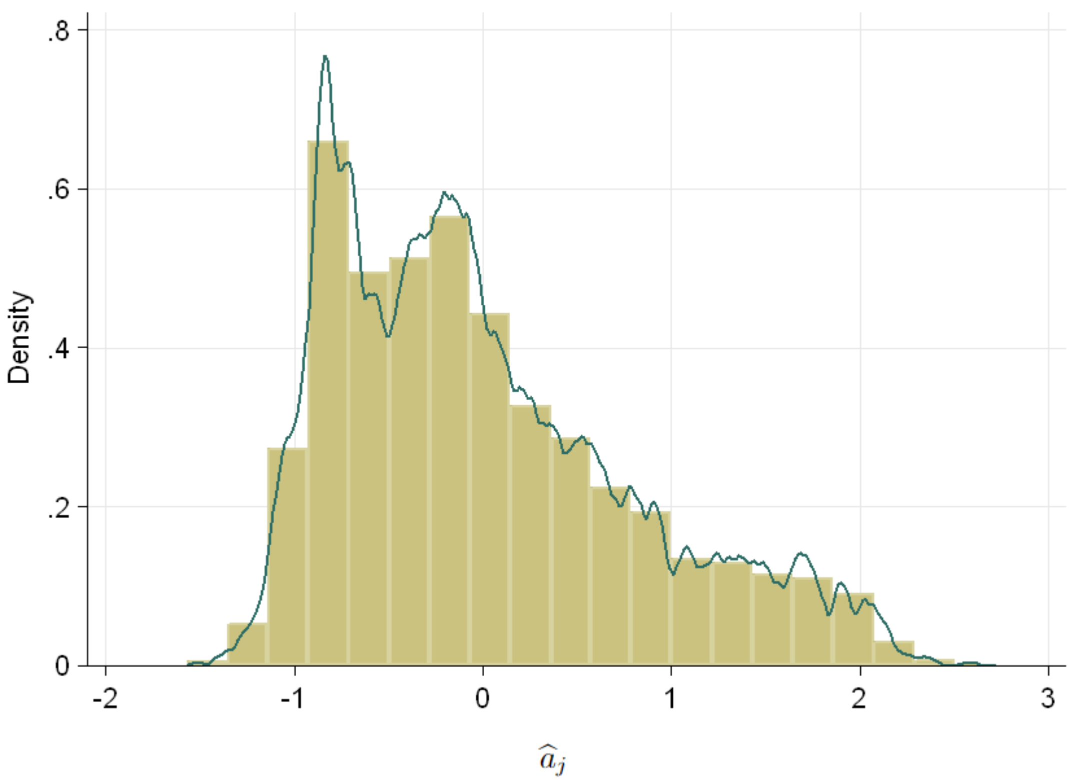


Figure 8: Distribution of $\widehat{a}_j$

Note: Densities computed using establishment sampling weights.

Tables

Table 1: Number of establishments in the sample, by sector and size groups

	Total	**Agricult.**	**Manufact.**	**Services**	**Small**	**Medium**	**Large**
Bangladesh	903		903		361	232	310
Brazil (Ceará)*	711	72	387	252	205	322	184
Burkina Faso	600	80	140	380	335	187	78
Ghana	1,262	85	275	902	774	382	106
India**	1,519		791	728	629	598	292
Kenya	1,305	155	335	815	499	421	385
Korea	1,551	129	652	770	656	569	326
Poland	1,500	90	607	803	779	394	327
Senegal	1,786	204	679	903	1,219	395	172
Vietnam	1,499	110	806	583	774	426	299
Total	**12,636**	**925**	**5,575**	**6,136**	**6,231**	**3,926**	**2,479**

Note: * Brazil refers to state of Ceará; **States of Tamil Nadu and Uttar Pradesh in India. The survey does not cover agriculture in Bangladesh and India, nor services in Bangladesh. Table A.2 provides the total number of firms in the universe covered by the survey, by country, sector, and size groups.

Table 2: Summary Statistics

	All			Small (5-19)		Medium (5-19)		Large (100+)	
	N	Mean	p50	N	Mean	N	Mean	N	Mean
Total # of employees	11124	98.13	20	5451	9.82	3531	41.05	2142	416.96
% of workers with college	11208	22.79	10	5474	20.83	3521	24.92	2213	24.23
Management practice (z-score)	12626	0.10	0.35	6225	-0.09	3925	0.17	2476	0.47
Log value added per worker	8258	10.01	9.96	3915	9.80	2671	10.09	1672	10.37
		Share			Share		Share		Share
Multi-establishment		21.3%			13.6%		23.5%		36.9%
Multinational		11.5%			7.3%		13.0%		19.8%
Exporter		21.3%			10.7%		21.1%		48.5%
Age groups:									
1-5 years		14.0%			17.8%		12.4%		7.1%
6-10 years		20.2%			23.2%		18.7%		15.2%
11-15 years		18.7%			20.5%		17.0%		16.7%
16+ years		47.1%			38.5%		51.9%		61.0%
Has electricity, computer and internet		76.0%			64.1%		84.0%		94.4%
Sector:									
Agriculture		5.5%			5.2%		5.6%		6.0%
Livestock		1.8%			1.9%		1.9%		1.6%
Food Processing		10.2%			9.0%		11.1%		11.7%
Apparel		8.9%			8.2%		8.1%		11.9%
Motor vehicles		3.2%			1.9%		3.9%		5.2%
Pharmaceuticals		3.3%			1.7%		3.9%		6.4%
Wholesale or retail		15.1%			19.4%		12.3%		8.7%
Financial services		4.3%			4.4%		5.0%		2.8%
Land transport		3.9%			3.4%		4.4%		4.1%
Health services		4.2%			1.7%		5.8%		7.7%
Leather		4.3%			3.9%		4.1%		5.6%
Other Manufact.		14.3%			15.7%		12.9%		13.3%
Other Services		21.2%			23.8%		21.0%		15.0%

Note: Statistics computed weighting establishment by sampling weights.

Table 3: Descriptive statistics of technology sophistication at the business function and establishment level

	Mean	SD	p10	p50	p90	Skewness	Kurtosis
MOST_{fj}	1.77	1.02	1.00	1.67	3.67	1.47	4.56
MAX_{fj}	2.39	1.28	1.00	2.00	4.33	0.59	2.17
MOST_{j}	1.79	0.61	1.10	1.71	2.64	0.92	3.67
MAX_{j}	2.41	0.82	1.43	2.28	3.58	0.61	2.86

Note: Statistics are computed using establishment and function/establishment weights for establishment-level and function/establishment-level measures. pX denotes the Xth percentile and SD standard deviation.

Table 4: Technology sophistication and establishment characteristics

VARIABLES	(1) MOST_j	(2) MAX_j
Medium	0.21***	0.35***
	(0.01)	(0.01)
Large	0.51***	0.80***
	(0.02)	(0.02)
Age 6 to 10	0.02	-0.02
	(0.01)	(0.02)
Age 11 to 15	0.05***	0.01
	(0.01)	(0.02)
Age 16+	0.04***	0.01
	(0.01)	(0.02)
Multi-establishment	0.14***	0.37***
	(0.01)	(0.02)
Foreign owned	0.18***	0.36***
	(0.02)	(0.03)
Exporter	0.20***	0.33***
	(0.01)	(0.02)
Observations	12,408	12,408
R-squared	0.45	0.41

Note: Observations are weighted by establishment weights. Regressions include country and one-digit sector dummies. Medium and Large are dummies that correspond to establishments with 20-99 employees and 100+ employees, respectively. Multi-establishment, Foreign owned, and Exporter are dummies that reflect whether the establishment is part of a multi-establishment firm, belongs to a multi-national, and exports, respectively.

Table 5: Technology sophistication across countries

	lnGDP	MOST All	MOST Agr	MOST Mfg	MOST Svc	MAX All	MAX Agr	MAX Mfg	MAX Svc
Korea	10.65	2.32	2.62	2.26	2.34	2.63	2.90	2.58	2.66
Poland	10.38	2.17	2.44	2.19	2.17	2.76	2.84	2.77	2.79
Brazil	9.59	2.33	2.43	2.11	2.40	2.96	2.90	2.78	3.04
Vietnam	8.92	1.95	2.07	1.83	2.00	2.60	2.61	2.44	2.67
India	8.81	1.67	-	1.67	1.65	2.48	-	2.39	2.53
Ghana	8.58	1.66	1.87	1.57	1.68	2.52	2.28	2.25	2.64
Bangladesh	8.44	1.54	-	1.49	-	1.90	-	1.80	-
Kenya	8.35	1.56	1.87	1.62	1.56	2.39	2.34	2.42	2.43
Senegal	8.09	1.38	1.36	1.35	1.41	1.92	1.71	1.83	1.99
BurkinaFaso	7.66	1.31	1.31	1.37	1.33	1.90	1.95	1.87	1.93
Corr	.	0.93	0.95	0.97	0.91	0.76	0.89	0.80	0.74
SD	0.97	0.38	0.49	0.34	0.40	0.38	0.45	0.37	0.36
Cov	.	0.34	0.50	0.32	0.37	0.28	0.43	0.29	0.27

Note: Table reports coefficients of country and country-sector dummies in establishment-level regressions of technology sophistication on establishment characteristics and dummies (see Table 4). LnGDP denotes log of per capita income from Penn World Tables in 2019. SD denotes standard deviation of column. Corr reports correlation of column with log per capita income. Cov denotes covariance of column and log per capita income.

Table 6: Variance decomposition of technology sophistication

	MOST	MAX
Var(S_{fjc})	1.05	1.64
Var(α_{fc})	0.31	0.42
Var(α_j)	0.22	0.48
Var(u_{fj})	0.47	0.66
Contribution of Wvar	45%	40%

Note: Table reports variance of the components of technology sophistication in business function f of establishment j (X_{fj}) in country c ($S_{f,j,c}$) as defined in equation (4). Contribution of Wvar is calculated as the ratio of Var(u_{fj}) over Var(X_{fjc}). Observations are weighted by function/establishment weights.

Table 7: Moments of the distribution of within-establishment variation across establishments

	Mean	SD	p10	p50	p90	Skewness	Kurtosis
Wvar(MOST)	0.54	0.55	0.04	0.36	1.30	1.73	6.73
Wvar(MAX)	0.74	0.55	0.22	0.57	1.51	1.53	5.91

Note: Wvar(X) denotes the within-establishment variance of technology sophistication measure X. In total, our sample has 12,394 observations. Mean denotes the average Wvar across establishments, SD the standard deviation, pX the value of Wvar in the Xth percentile of the distribution. Statistics computed using establishment sampling weights.

Table 8: Wvar and establishment characteristics.

VARIABLES	(1) Wvar(MOST)	(2) Wvar(MAX)
S_j	1.24***	1.31***
	(0.03)	(0.03)
S_j^2	-0.15***	-0.18***
	(0.01)	(0.01)
Medium	-0.02***	0.01
	(0.01)	(0.01)
Large	0.01	0.10***
	(0.02)	(0.02)
Age 6 to 10	0.01	-0.03**
	(0.01)	(0.01)
Age 11 to 15	-0.01	-0.05***
	(0.01)	(0.01)
Age 16+	0.02**	-0.04***
	(0.01)	(0.01)
Multi-establishments	-0.03***	-0.02**
	(0.01)	(0.01)
Foreign-owned	-0.04**	-0.02
	(0.02)	(0.02)
Exporters	-0.02*	-0.03**
	(0.01)	(0.01)
Number of BF	0.01***	0.00
	(0.00)	(0.00)
Manufacturing	-0.01	0.04
	(0.02)	(0.03)
Services	-0.00	0.01
	(0.02)	(0.03)
Country Fixed Effects	Yes	Yes
Observations	12,169	12,169
R-squared	0.49	0.33

Note: Dependent variables are the within-establishment variance of technology sophistication at the establishment level. Technology sophistication is measured either with MAX or with MOST. S_j=$MOST_j$, MAX_j denotes the average technology sophistication in establishment j. All regressions include country and one-digit sector fixed effects. Observations are weighted by establishment weights.

Table 9: Descriptive statistics of estimates of technology elasticity, TechFP, and log value added per worker

(a) $\widehat{\varepsilon}_f$

	Mean	SD	CV	p10	p90	p90/p10
$\widehat{\varepsilon}_f$	0.82	0.40	0.49	0.31	1.34	4.32

(b) $\widehat{a}_j$

	Mean	Median	SD	p10	p90	p90−p10	Skewness	Kurtosis
$\widehat{a}_j$	0	-0.16	0.80	-0.87	1.25	2.13	0.79	2.90

(c) (Log) Value added per worker

	Median	SD	p10	p90	p90−p10	Skewness	Kurtosis
ln(VAPW)	9.90	1.77	7.88	12.15	4.27	0.16	3.44

Note: Statistics of $\widehat{\varepsilon}_f$ are unweighted. Statistics of $\widehat{a}_j$ and log value added per worker are weighted using establishment weights. Value added per worker excludes establishments at top and bottom 1% of distribution.

Table 10: Variance decomposition of $s_{f,j}$

Var($\widehat{\kappa}_j$)	0.04
Var($\widehat{\beta}_f * X_j$)	0.06
Var($\widehat{\varepsilon}_f$*$\widehat{a}_j$)	0.44
Var($s_{f,j}$-$\widehat{\kappa}_f$−$\widehat{\kappa}_j$)	1.02
Var($s_{f,j}$-$\widehat{\kappa}_f$)	0.93
Contribution of technology curve to within-establishment variance	43%

Note: Estimates from specification (17) using the two-step approach described in the text. Observations are weighted by function/establishment weights. Contribution of the technology curve to WVAR is computed as the ratio of the third to the fourth line times 100.

Table 11: Technology elasticities across sectors and functions

Business Function	Max-Min across sectors	Baseline
Business Administration	1.10	1.25
Production Planning	0.67	1.18
Sourcing	0.44	0.85
Marketing	0.74	0.71
Sales	0.98	0.5
Payment	1.03	0.36
Quality Control	0.48	0.8
Average (Max-Min) across GBFs	0.78	
Fabrication	0.18	
Max-Min across functions		0.89

Note: Max-Min across sectors reports the gap between the highest and lowest estimate of $\varepsilon_{f,s}$ across sectors for each business function f. Fabrication reports the difference between the highest and lowest estimates of the technology elasticities of fabrication across the six manufacturing sectors where it is relevant. Baseline reports the estimates of ε_f in the baseline specification (17) where the technology elasticities of each GBF are restricted to be the same across sectors. Max-Min across functions reports the difference between the highest and lowest estimates in the second column.

Table 12: TechFP and other technology measures

VARIABLES	(1)	(2)	(3)	$\widehat{a}_j$ (4)	(5)	(6)	(7)
Computer/Electricity/Internet	0.941***						
	(0.020)						
Sophistication in Business Administration		0.515***					
		(0.007)					
Sophistication in Payments			0.297***				
			(0.020)				
s_j				0.978***			
				(0.016)			
$\sqrt{WVAR_j}$				0.171***			
				(0.040)			
MOST$_j$					0.646***		
					(0.021)		
MAX$_j$					0.421***		
					(0.012)		
$s_{SSBF,j}$						0.033**	
						(0.015)	
$s_{GBF,j}$						0.994***	
						(0.011)	
$s_{SSBF,j}$*Agriculture							0.379***
							(0.042)
$s_{SSBF,j}$*Manufacturing							-0.012
							(0.024)
$s_{SSBF,j}$*Services							0.046**
							(0.021)
$s_{GBF,j}$*Agriculture							0.458***
							(0.050)
$s_{GBF,j}$*Manufacturing							0.963***
							(0.020)
$s_{GBF,j}$*Services							1.023***
							(0.014)
No SSBFs						-0.243***	-0.475***
						(0.039)	(0.048)
Observations	10,812	10,833	10,986	11,114	11,114	10,958	10,958
R-squared	0.297	0.691	0.075	0.736	0.739	0.814	0.818
Sector FE	No	No	No	No	No	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates of specification (27). Dependent variable is TechFP in the establishment ($\widehat{a}_j$). Independent variables are listed in the first column. Computer/Electricity/Internet is a binary variable that takes value of 1 if the establishment has these three technologies and 0 otherwise. Sophistication in Business Administration and in Payments are the sophistication levels ($s_{f,j}$) in these functions, where $s_{f,j} = 1/2*(MOST_{f,j} + MAX_f, j)$; s_j is the average of $s_{f,j}$ across all the business functions of the establishment; $s_{SSBF,j}$ ($s_{GBF,j}$) is the average of $s_{f,j}$ across all the sector-specific (general) business functions of the establishment; Agriculture, Manufacturing and Services are dummies that take the value of 1 if the establishment belongs to that 1-digit sector. No SSBFs is a dummy that is 1 if the establishment belongs to a sector where FAT does not record any information on the technologies used in sector-specific functions. Observations weighted by establishment sampling weights.

Table 13: Technological productivity and firm characteristics

VARIABLES	$\widehat{a}_j$
h_j	0.005***
	(0.000)
Management	0.171***
	(0.015)
Multi-establishment	0.292***
	(0.035)
Multinational	0.215***
	(0.040)
Exporter	0.251***
	(0.031)
6 to 10 years	-0.009
	(0.032)
11 to 15 years	0.006
	(0.033)
16+ years	0.014*
	(0.033)
Observations	10,320
R-squared	0.485
Country; sector FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates from equation (28). h_j is the share of college workers in the establishment; management is the z-score. Observations weighted by establishment sampling weights.

Table 14: Productivity and technology

	Value added per worker				
VARIABLES	(1)	(2)	(3)	(4)	(5)
k_j	0.304***	0.287***	-0.032***	0.286***	0.283***
	(0.023)	(0.022)	(0.009)	(0.022)	(0.022)
h_j	0.005***	0.003***	0.015***	0.003***	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\widehat{a}_j$		0.306***	0.475***		0.260***
		(0.042)	(0.041)		(0.046)
$\widehat{a}_j$*Agriculture				0.790***	
				(0.160)	
$\widehat{a}_j$*Manufacturing				0.364***	
				(0.049)	
$\widehat{a}_j$*Services				0.268***	
				(0.053)	
Management					0.087***
					(0.031)
Computer					0.108
					(0.067)
Observations	6,839	6,839	6,839	6,839	6,812
R-squared	0.483	0.495	0.214	0.496	0.498
Country FE	Yes	Yes	No	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates from equation (29). k_j refers to the log of book value of capital per worker; h_j is the share of college workers in the establishment; Computer is a dummy variable that takes the value of one if the establishment has access to computer, electricity, and internet; Management is the z-score following **?**.

Contents (Appendix)

A The FAT survey and section 2 results

This section provides more details on the Firm-level Adoption of Technologies (FAT) survey and its implementation. We start with a description of the grid of technologies in FAT. Then we describe the sampling frameworks used and the construction of sampling weights. We finalize describing all the tests conducted to minimize potential biases, including validation exercises ex post implemented with with external data sources.

A.1 The survey

The FAT survey is a multi-country, multi-sector, representative firm-level survey. It collects information on the technologies used by firms in specific business functions that encompass the key activities that each firm conducts. Compared to existing firm-level surveys, the FAT survey covers a significantly larger number of technologies and business functions (Table A.1), and a wider range of sectors; for example, it covers agriculture distinguishing between crops and livestock.

Table A.1: Coverage of Firm-Level Technology Surveys

Surveys	# of Technologies	# of Business Functions	Includes Firms in Agriculture
Firm-level Adoption of Technology Survey	287	59	Yes
Survey of Advanced Technology (SAT)	57	3	No
Community Survey on ICT Usage and E-Commerce in Enterprises	9	0	No
Information & Communication Technology Survey (ICTS)	4	0	No
Annual Business Survey (ABS) 2019	5	0	No

Note: The Number of technologies and business functions are computed by authors.

The FAT survey addresses important knowledge gaps compared to other surveys measuring technology at firm-level surveys. For starters, the number of technologies covered is rather limited when compared to how many technologies are involved in production processes. Second, their focus on the presence of advanced technologies makes impossible to understand how production takes place in companies without such advanced technologies. This concern is most relevant in developing countries where advanced technologies have diffused less. Third, since their unit of analysis is the firm, existing studies are not designed to study what business functions benefit from each technology. This drawback is particularly problematic for general technologies that can be relevant for multiple business functions. Finally, existing surveys largely omit questions about how intensively a technology is employed in the firm, and therefore, they do not reveal whether a technology that is present is widely utilized or just marginally.

Specifically, the FAT survey comprises five sections:

- Module A– Collects general information about the characteristics of the establishment; such as sector, multi-establishment and ownership.
- Module B – Covers the technologies used in eight generic business functions.
- Module C – Covers the use of technologies for functions that are specific to each of ten industry and services sectors
- Module D – Includes questions about the drivers and barriers for technology adoption.
- Module E – Collects information on employment, balance sheet and performance, which allow us to compute labor productivity and other measures at the company level.

A.1.1 The Grid. Business functions and relevant technologies

We construct a technology grid that identifies the main business functions and the key technologies used to carry out the tasks of each business function. To design modules B and C, the survey draws upon the knowledge of experts in production and technology in various fields and sectors. These experts provided their knowledge on: i) what are the key general and sector-specific business functions, ii) what are the different technologies used to conduct the main tasks in each function, and iii) how are the different technologies related, both in terms of their sophistication and the degree of substitutability between them.

First, we started with a desk research revising the specialized literature identifying business functions and technologies across the value chain.[60] Second, for each sector, as well as for the general business functions, we hold meetings with private sector specialists at the World Bank Group to validate the initial findings and start to define the key business functions and technologies. Third, we hold meetings with Lead and Senior Economists across the World Bank Group, including the International Finance Corporation (IFC), from different fields of specialization and wide experience with sectoral projects in several countries (e.g. agriculture, manufacturing, retail, transport, health, etc.). Fourth, we hold meetings and validation exercises with external senior consultants, with wide experience on the field (e.g. at least 15 years), including experience with firms in developing countries as well as advanced economies.

The source of external senior consultants in the last layer of quality control varied across sector. For agriculture and livestock, the validation exercise was conducted with agricultural engineers and researchers from Embrapa, an agricultural research institution from Brazil. For food processing, wearing apparel, pharmaceutical, transport, and retail, as well as for

[60] This process involved the revision of peer-review journals and reports from international organizations and industry associations.

the general business functions, the team hired external consultants through a large management consultant organization. For automotive sector, the team has hired a senior consultant directly. For health, the team invited directly five physicians with different field of specializations and practical experience in hospitals in clinics in the United States and low income countries in Saharan Africa.

The validation exercise with sector specialists were organized as follows. First, the team would explain the purpose of the project, present the initial findings, and share a draft with identified business functions and technologies associated with them. The sector specialists would have between one and two weeks to reflect on the material to validate them or propose a new combination of business functions and technologies associated with them. After receiving the revised material, a second meeting with sector specialists would be organized with the FAT survey team to discuss the proposal and converge towards an updated combination of business's functions and technologies.

In what follows we describe the grids for both types of business functions.

A.1.2 General Business Functions

Figure 1 shows the 7 general business functions in FAT and the possible technologies used to conduct them. The business functions identified are: administration (HR processes, finance,..), production planning, Procurement and supply chain management, marketing and product development, sales, payment methods, and quality control. These are business functions that in addition to be central in the functioning of the firm, are also retained in some capacity (or some tasks) within the firm, and where there are some known and off-the-shelf technologies, often ICT technologies, that can be adopted to implement the tasks needed in these business functions. For example, for administrative processes, these range from handwritten processes (the least sophisticated) to the use of enterprise resource planning which are software that allow for real time, integrated management of the main business processes. With the help of management consultants, we identify the technologies feasible for each business function and develop similar rankings of sophistication based on the consultants understanding of the number of tasks and complexity that the technologies can handle.

One important characteristic of the grid is that the sophistication rankings are not fully hierarchical for all business functions. In the case of sales, for example, firms can use various modes, and while online sales are more sophisticated technologies that on the phone on email, there is no clear sophistication ranking between sales on the company's website or using online platforms; both are complementary. A similar example occurs with payment methods, where firms may use a variety of them, often depending on the existence of payments systems in the country.

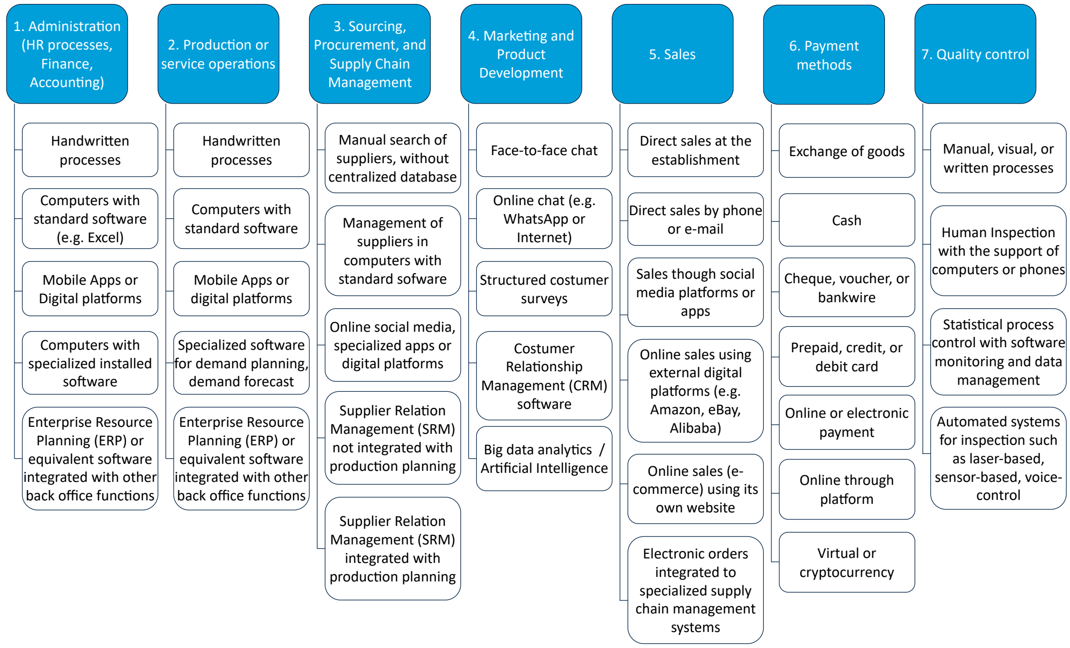


Figure A.1: General Business Functions and Their Technologies

A key advantage of the grid structure is that it allows to accommodate the use of more than one technology by business function. The survey questionnaire is implemented so respondents are asked about the use of each of the technologies in the grid. In addition, for those technologies selected in each business function, the respondent is asked to identify the one that is more intensively used in implementing the tasks of the business function; and when using one of the most advanced technologies also the year of adoption. This allows to uncover new facts about technology adoption and use, by allowing to build new measures of technology sophistication at the business function level based on extensive measures, the most sophisticated technology, and intensive measures, the technology used more intensively; and also calculate diffusion lags for several advanced technologies from firm level data.

A.1.3 Sector Specific Business Functions

For the sector-specific technologies, a similar approach was used to identify key business functions and associated technologies in 12 sectors of activity across agriculture, manufacturing, and services (including agriculture-crops; livestock; food processing; wearing apparel; leather and footwear; automotive; pharmaceutical; wholesale and retail; transportation; financial services; health services; accommodation). One business function, fabrication, was also included for all manufacturing sectors. The identification of key business functions and the frontier in each sector required a significant interaction with several sector specialists. These functions tend to be associated with sector-specific production processes.

Here, we present all sector specific business functions and associated technologies covered by the FAT survey in the first and second phase of data collection. These figures complement the information provided in section 2, particularly Figures 1 and 2, which describe the functions and associated technologies for GBFs and food processing, among SSBFs. The complementary information is provided for all SSBFs (Agriculture - Crops (Figure A.2), Livestock (Figure A.3), Food Processing (Figure A.4), Wearing Apparel (Figure A.5), Leather and Footwear (Figure A.6), Automotive (Figure A.7), Pharmaceutical (Figure A.8), Wholesale and Retail (Figure A.9), Transportation (Figure A.10), Financial Services (Figure A.11), Health Services (Figure A.12), and Other Manufacturing (Figure A.13)). [61]

[61] As the survey is rolled out in other countries, the number of additional sectors included in the survey is also increasing.

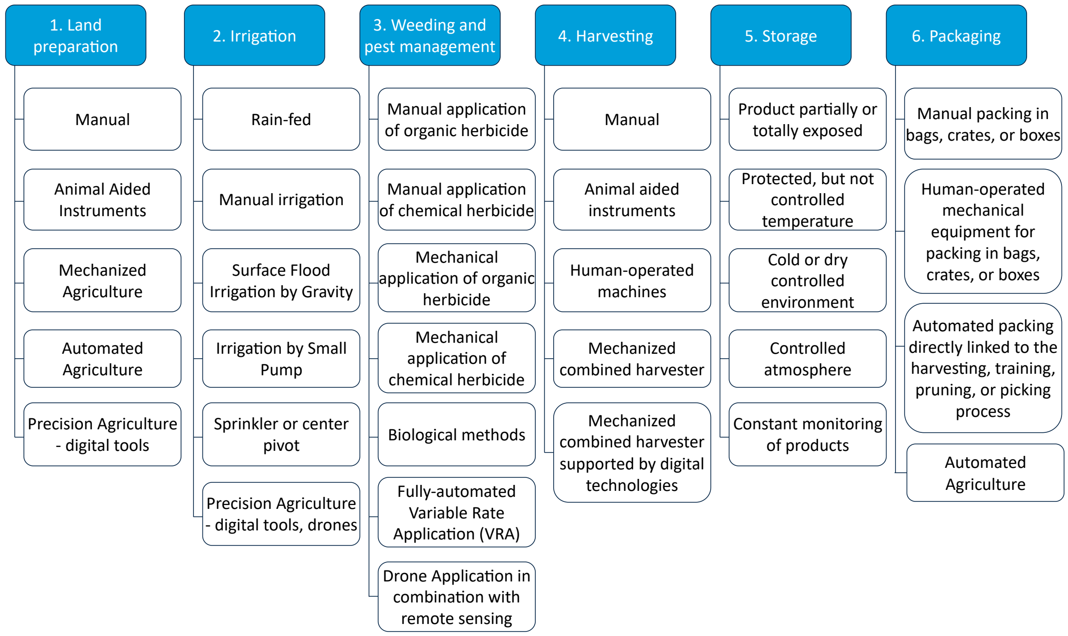


Figure A.2: Sector Specific Business Functions and Technologies in Agriculture - Crops

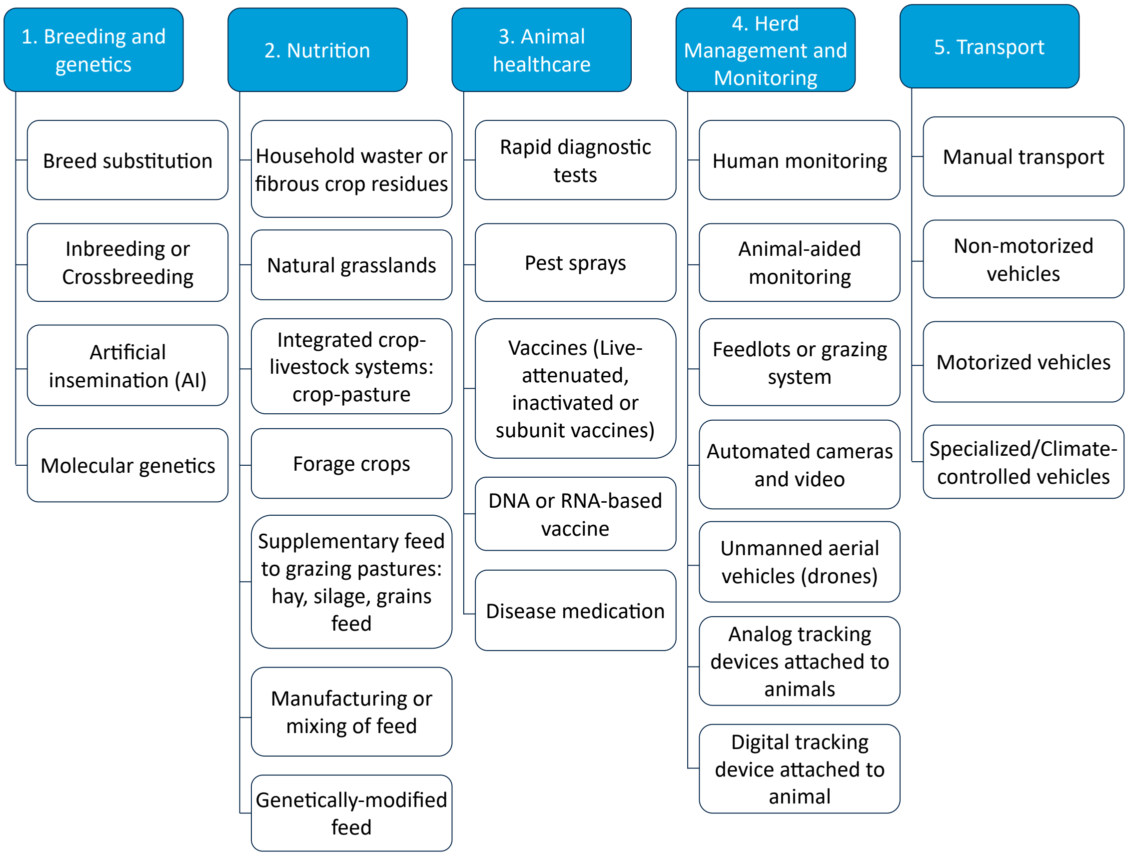


Figure A.3: Agriculture - Livestock: Business Functions and Technologies

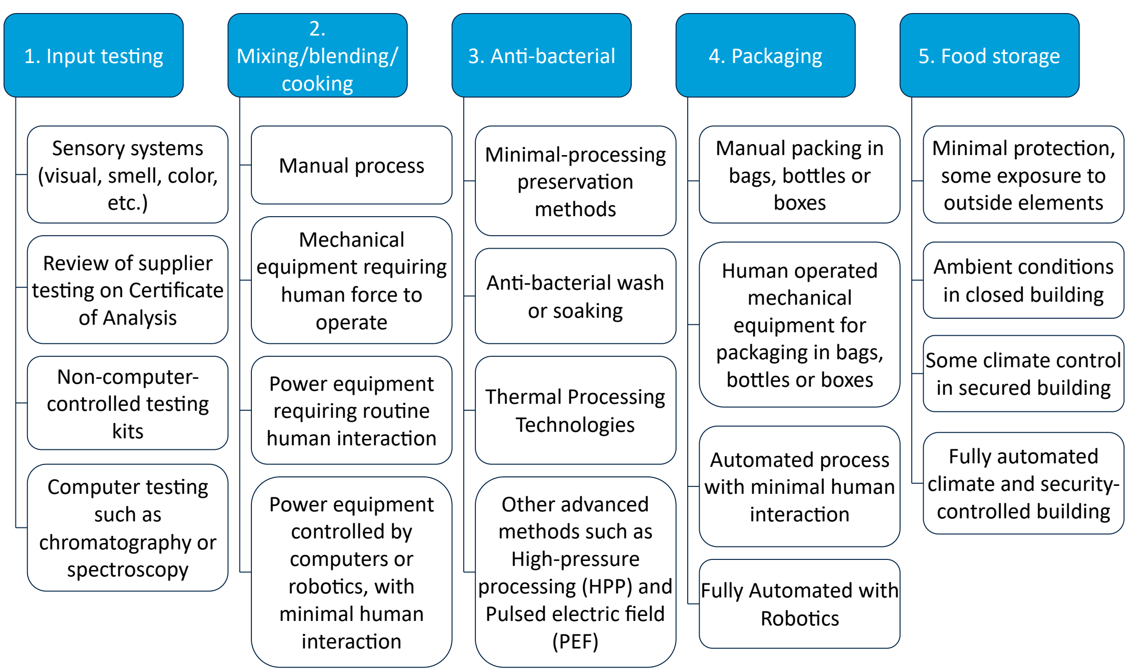


Figure A.4: Food Processing: Business Functions and Technologies

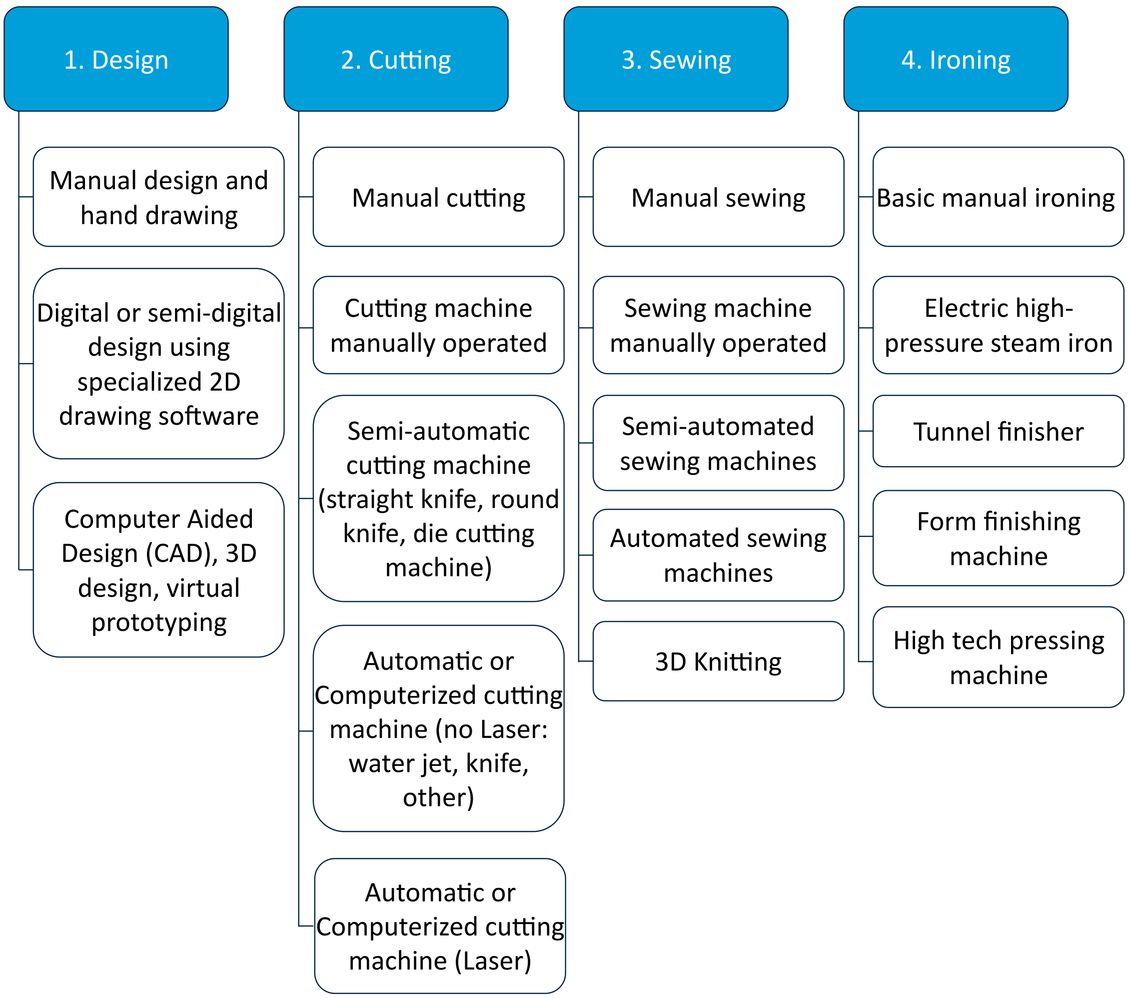


Figure A.5: Wearing Apparel: Business Functions and Technologies

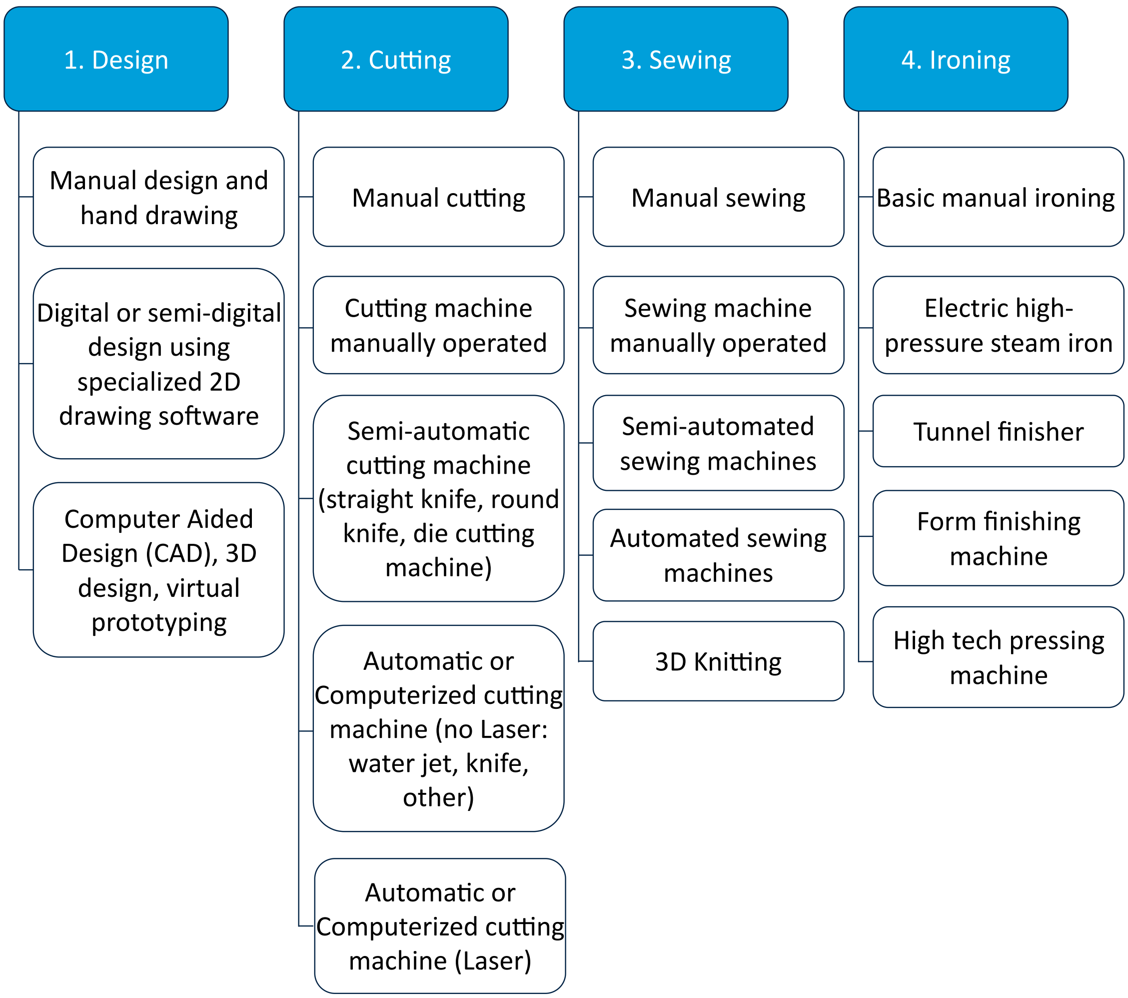


Figure A.6: Leather and Footwear: Business Functions and Technologies

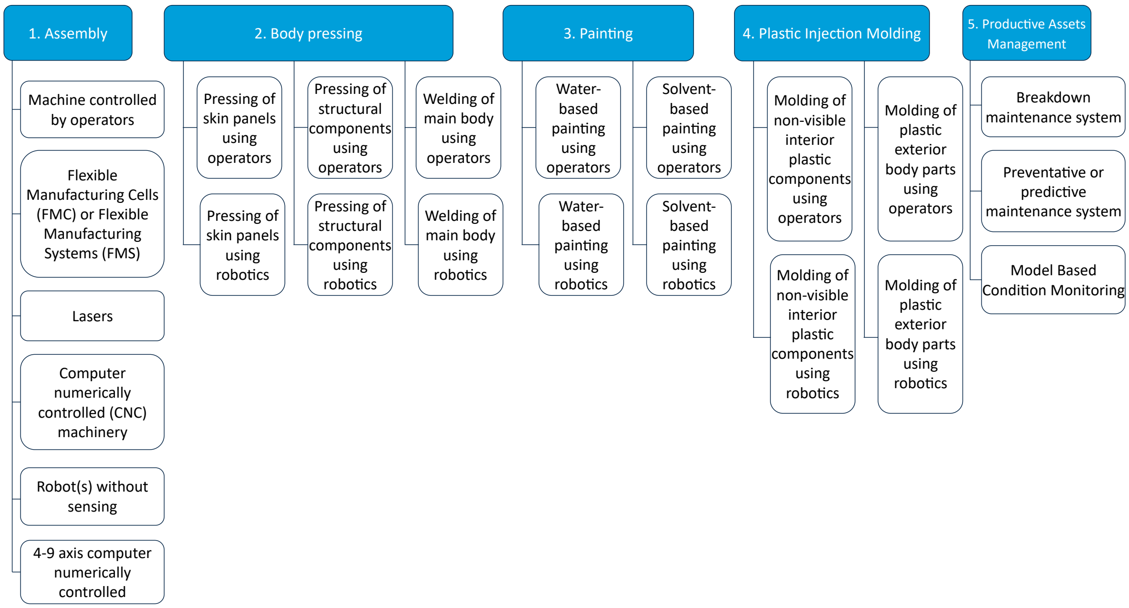


Figure A.7: Automotive: Business Functions and Technologies

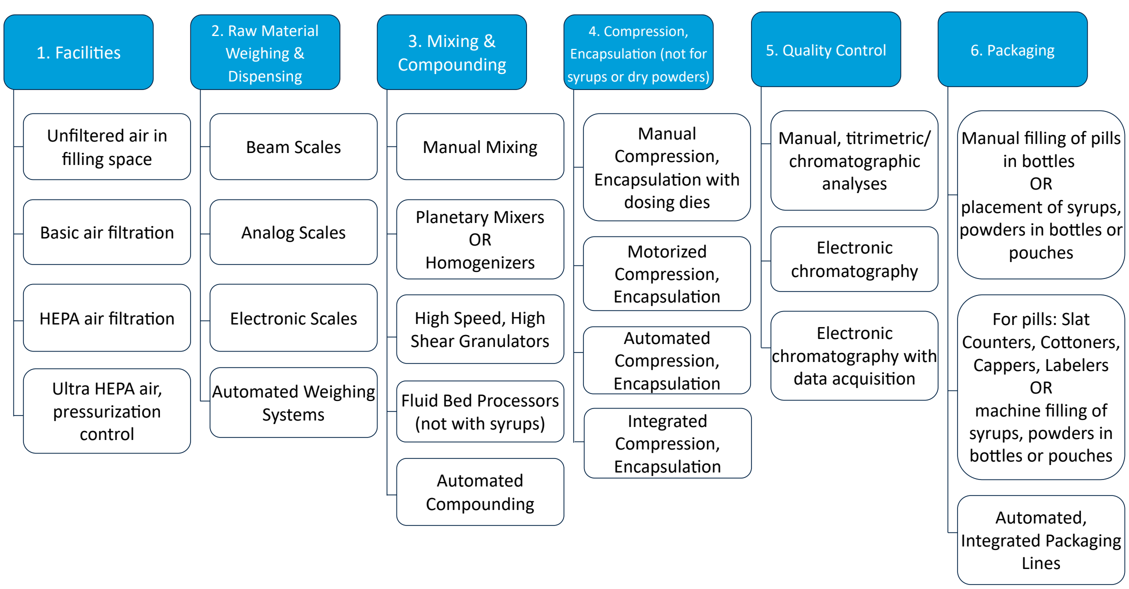


Figure A.8: Pharmaceutical: Business Functions and Technologies

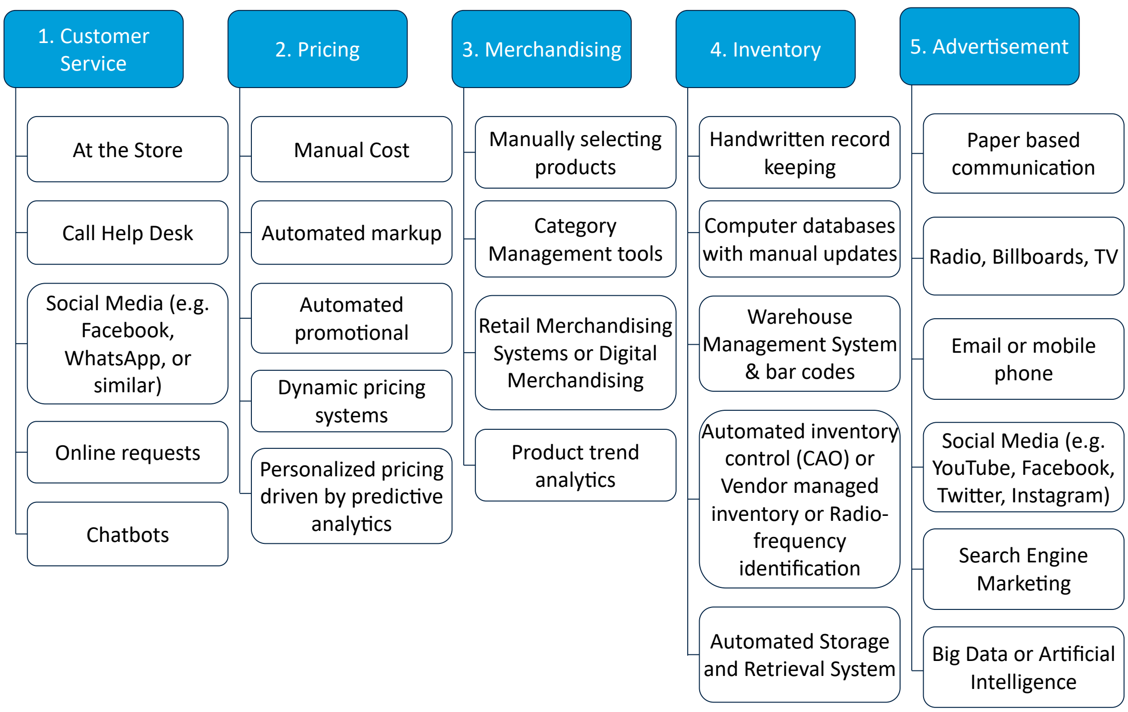


Figure A.9: Wholesale and Retail: Business Functions and Technologies

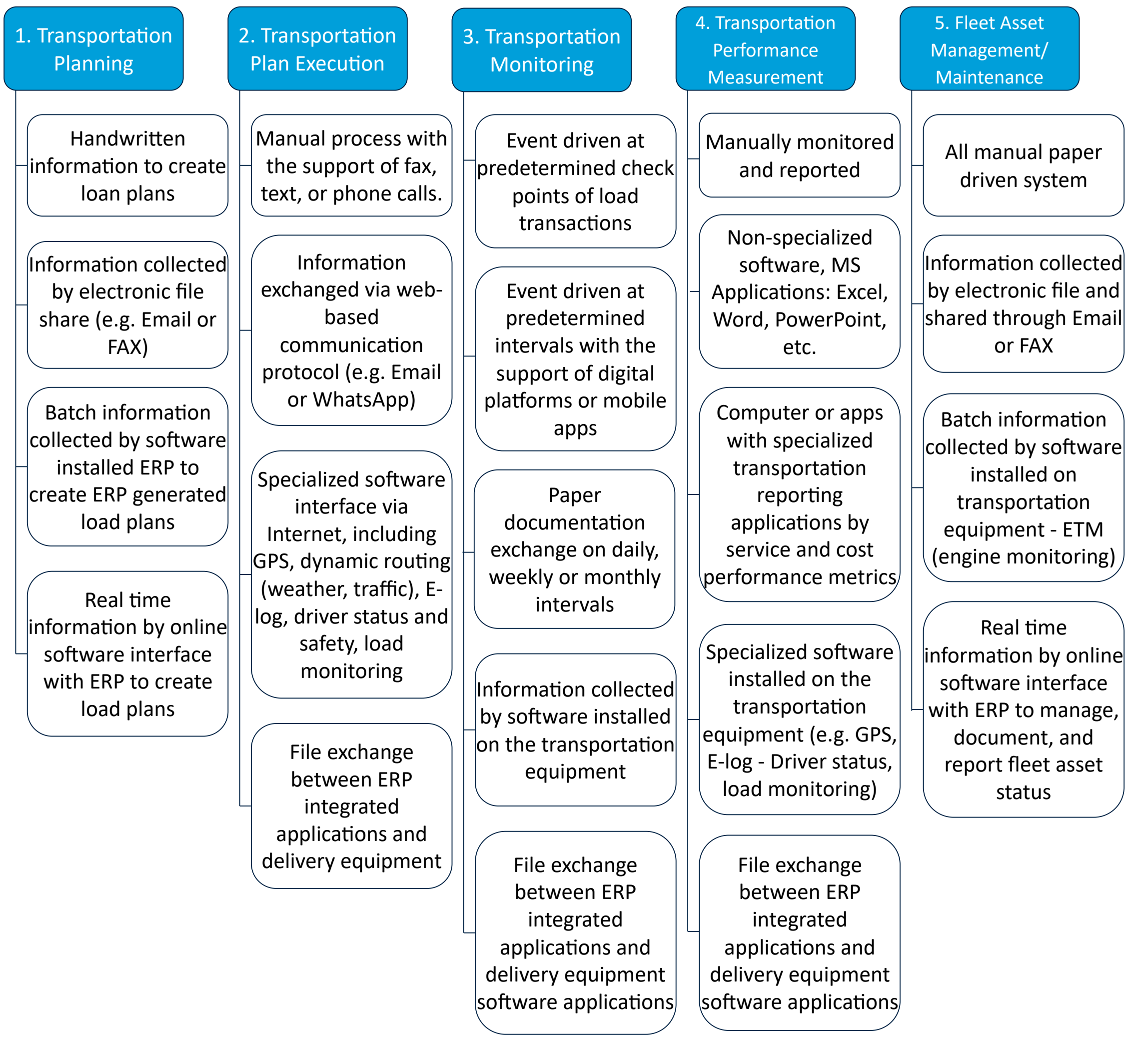


Figure A.10: Land Transportation: Business Functions and Technologies

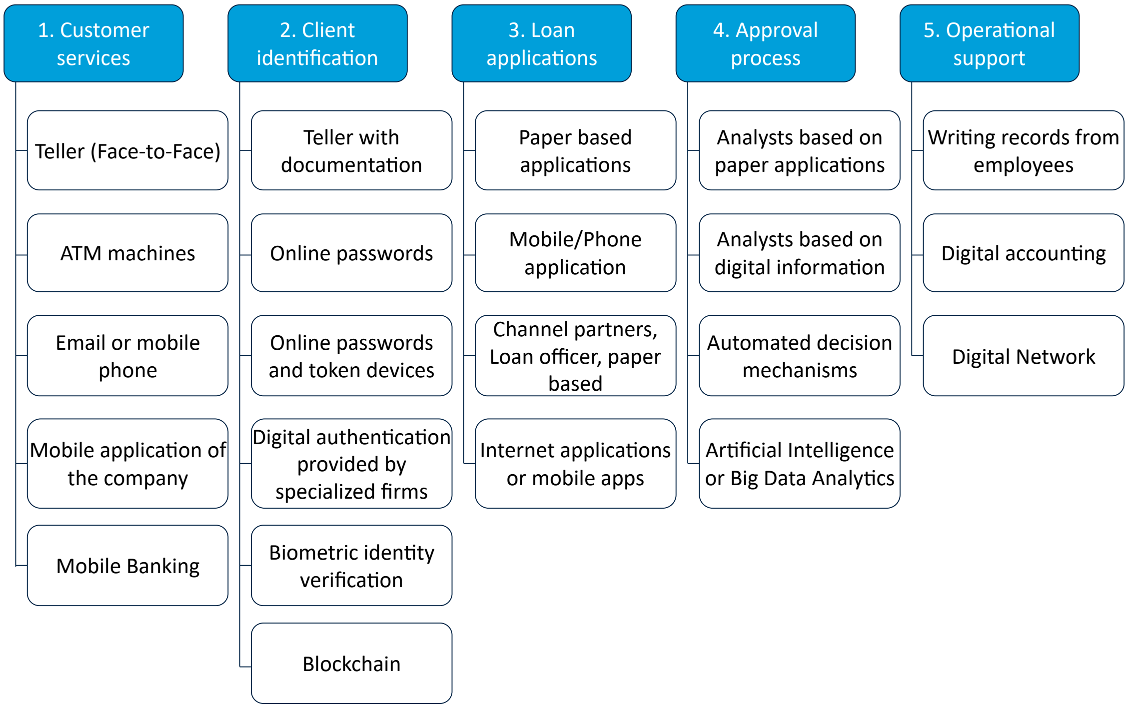


Figure A.11: Financial Services: Business Functions and Technologies

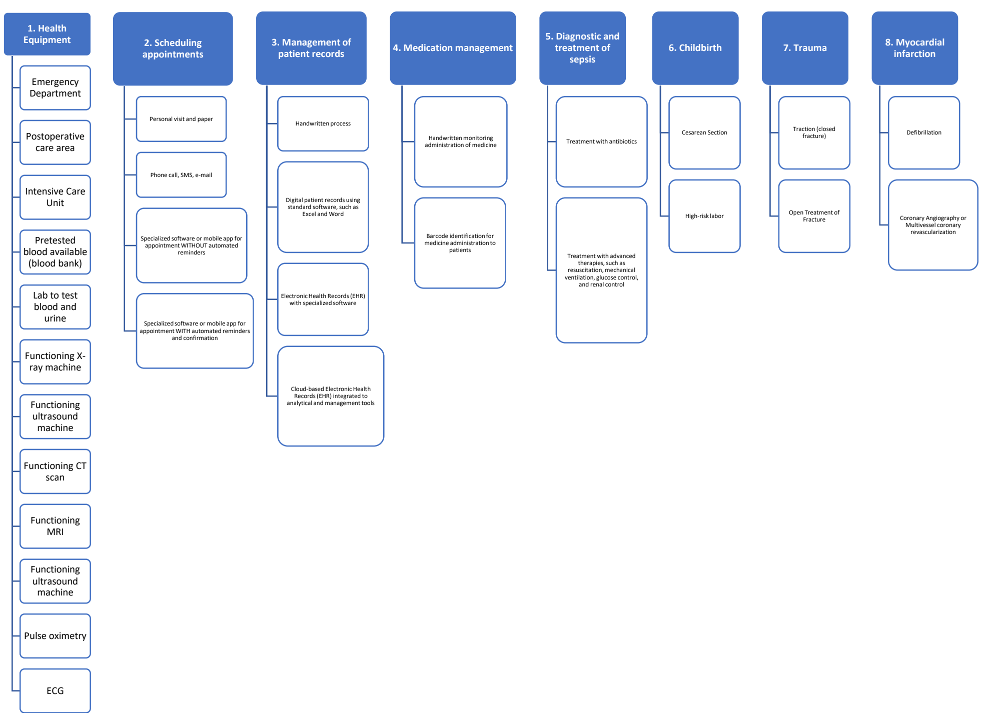


Figure A.12: Health Services: Business Functions and Technologies

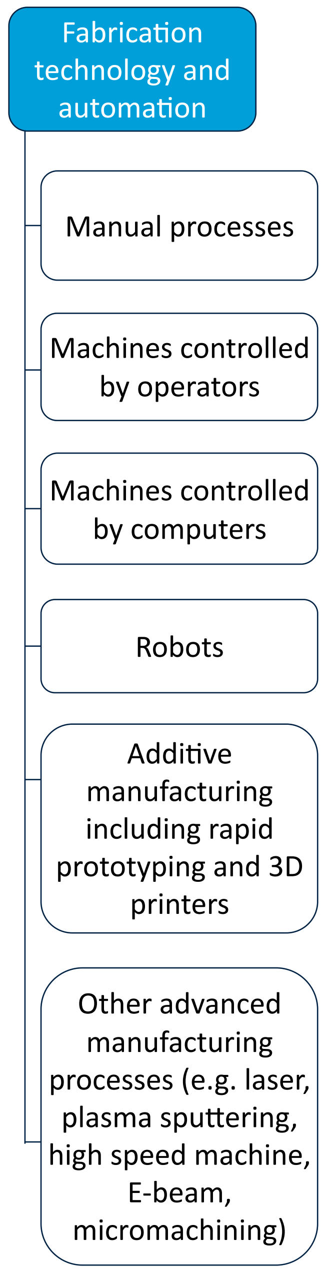


Figure A.13: Other Manufacturing: Business Functions and Technologies

For sector-specific business functions, digital technologies tend to be embedded in other technologies that are usually at the frontier. This is a common feature, particularly in agriculture and manufacturing, and has important implications in terms of the costs of adoption and the importance of network effects. For example, among methods commonly used by agricultural firms to perform harvesting, the most basic option is to harvest manually, followed by animal-aided instruments, human-operated machines, or a single tractor with one specific function (such as a single-axle tractor), a combined harvester (machines or tractors that combine multiple functions fully operated by the worker), and combined harvester using the support of digital technologies (such as global positioning systems [GPS] or computing systems integrated with the tractor). Unlike GBFs, the application of digital technologies for harvesting requires other sophisticated equipment or machines.

In addition to the possibility of computing different measures of technology sophistication for sector specific business functions, an important feature of the sector specific grid is the fact that it includes screening questions that allow for the fact that not all the business functions are carried out within the establishment. In other words, not all entries in the grid need to be implemented at the establishment or at the firm. While the tasks of most general business functions related to management and organization are usually carried out within the boundaries of the firm - either in the same establishment or in another establishment of the firm if multi establishment - some sector specific business functions can be carried out in another establishment within the same firm (insourcing), or they can be (outsourced) to a different firm. Our approach is, therefore, rooted in a view of the firm similar to Coase (1937), where firms are agents coordinating and implementing some tasks. The advantage of this approach is twofold. In addition to the fact that this approach allows a better identification of technology and its use described above, it allows to study critical questions such as the organization of the firm and tasks (Williamson, 1979), and more importantly the relationship between organizational modes, transaction costs and technological choice (Williamson, 1988).

After finalizing the FAT questionnaire, we pre-piloted it in Brazil and Senegal. We personally conducted the face-to-face interviews, in collaboration with enumerators and supervisors trained to conduct data collection with firms from different sectors and size groups. In the pre-pilot stage, we tested if the business functions and technologies covered by the questionnaire were comprehensive and clearly understood by respondents, through detailed discussions and follow up questions with representative of firms, which led us to make the necessary adjustments to the survey. For example, we experimented with survey designs that asked about the fraction of time/output/processes that were conducted with each of the technologies in the business function. We decided against using this approach to reflect the intensity of use of technologies because it was harder for respondents to answer precisely, and as a result led to a more subjective interpretation, which made the comparability of answers across business functions and companies harder to interpret.

A.1.4 Barriers and Drivers

In addition to the information on the technologies used by firms, the survey also collects information on potential drivers of and barriers to technology adoption. First, the surveys asks whether the firm acquired new machines, equipment, and software in the last three years, and in the case of machine whether leased, new or secondhand. the survey also asks questions on links to larger firms and multinationals, either via value chain links as supplier or buyer, or via the CEO previous experience working in a MNE or a large firm.

The survey also asks questions about access to finance and trade status. The first question is about having secured a loan in the previous three years for purchasing equipment, machinery or software. On more general access to finance, the survey asks how many times the establishment needed to borrow money to expand production but could not get it.On trading status, the survey asks whether the firm is an importer and exporter, and if an exporter, what is the share of sales that is exported.

A key complementary factor for technology adoption is the quality of management. The survey pays special attention to management by collecting information on the top manager's background and on management practices. Specifically, FAT asks about the level of education attainment of the top manager in the establishment, whether she has studied abroad, and whether she has experience in multinationals. In addition, the survey contains four questions about management practices. These include four questions from MOPS (Bloom et al., 2016) on the number of KPIs, the frequency with which they are monitored, the horizon of production targets and a question on the use of formal incentives. Though the information we collect on management practices is more restricted than the sixteen questions in MOPS, we have used information from the Mexico ENAPROCE survey and show that the index that emerges from the small number of variables collected is highly correlated with the full MOPS index and it captures a large fraction of the cross-firm variance in the quality of management practices.[62]

To investigate also the potential role of policy, the survey asks questions about awareness about public programs to support technology upgrading and whether the firm is a beneficiary of a program, and if so, what type of support the firm received.

[62]Specifically, we use data from Mexico ENAPROCE survey and calculate the correlation between a management quality index with the 4 questions in FAT and the overall index using all questions of MOPS that are in ENAPROCE. The correlations are 0.74 for 2015 survey and 0.73 for 2018; which suggests that with less questions we are still able to capture most of the variation in management quality.

While the approach of the survey is as much as possible to ask factual questions, it is still important to understand the perceptions that entrepreneurs and managers have on what are the main barriers and drivers of the decisions to adopt or not to adopt technology. To this end, the questionnaire asks the respondent to select the most important obstacle and driver for adoption from a closed list of options. As barriers we include, lack of information, uncertainty, cost, finance, regulations or lack of infrastructure. As drivers, we include competition, adoption by other firms, production of new products, accessing new markets, cost reductions or adjusting to regulations. A final set of questions to gather the role of beliefs of the main managers in technology adoption decisions try to measure over confidence.

A.1.5 Balance Sheet

In addition to the information on the technologies used by firms, the survey also collects balance sheet information, information on the business owners, employees, and on potential drivers of and barriers to technology adoption.

Balance sheet. The survey asks the establishment about its total sales, material inputs, replacement value of capital stock, energy consumption, wages and employment. This allows to construct measures of nominal value added per worker, and capital per worker.

Employment. Beyond the number of employees, the survey asks questions that provide information on the education of the workers (share of workers with primary, secondary and tertiary education), and about the occupation composition of the labor force (share of Managers, Professionals, and Technicians; Clerical support workers and sales workers; Production workers and Service workers).

A.2 Sampling frame

The sampling frames were based on the most comprehensive and latest establishment census available from national statistical agencies or administrative business register. Table A.1 provides the main data sources used in the sample frame for each country.

The universe of study includes establishment with 5 or more employees in agriculture, manufacturing and services. The sector classification is based on the International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4. More specifically, our sample includes firms from the following ISIC rev 4 sectors: Agriculture (ISIC 01, from Group A); All manufacturing sectors (Group C); Construction (Group F), Wholesale and retail trade (Group G), Transportation and storage (Group G), Accommodation and food service activities (Group I), Information and communication (Group J), Financial and insurance activities (Group K), Financial services (ISIC, 64), Travel agency (ISIC 79, from group N),

Table A.1: Sampling frame by country

Country	Source	Sampling frame	Year
Bangladesh	Bangladesh Bureau of Statistics.	Est. census, 2013	2019
Brazil	Ministry of Labor	Employer census, RAIS, 2018	2019
Burkina Faso	Business Registry	Business Registry	2021
Ghana	Ghana Statistical Service	Est. census, 2013–18	2021
India	Central Statistics Office of India	Est. census, 2013–17	2020
Kenya	Kenya National Bureau of Statistics	Est. census, 2017	2020
Korea, Rep.	Statistics Korea	Est. census, 2018	2021
Poland	Statistics Poland	Est. census, 2020	2021
Senegal	National Agency for Statistics (ANSD)	Est. census, 2016	2019
Vietnam	General Statistics Office of Vietnam	Est. census, 2018	2019

Health services (ISIC 86, from group Q), and Repair services (ISIC 95, from Group S).

Table A.2: Total number of firms in the universe covered by the survey

Country	**Total**	**Sector**			**Firm size**		
		Agri.	**Manuf.**	**Serv.**	**Small**	**Medium**	**Large**
Bangladesh	15,358	-	15,358	-	4,164	3,425	7,769
Brazil (Ceará)	23,364	392	4,758	18,214	12,771	8,955	1,638
Burkina Faso	57,328	4,808	7,493	45,027	40,189	13,284	3,855
Ghana	42,165	880	10,284	31,001	30,133	10,070	1,962
India*	92,061	-	44,015	48,046	47,319	37,413	7,329
Kenya	74,255	3,680	5,407	65,168	50,584	16,676	6,995
Korea, Rep.	545,515	1520	167,466	376,529	450,264	82,403	12,848
Poland	244,983	3,021	52,340	189,622	198,107	37,799	9,077
Senegal	9,583	1,051	4,069	4,463	7,805	1,414	364
Vietnam	179,713	1,080	45,805	132,828	135,046	33,107	11,560
Total	**1,284,325**	**16,432**	**356,995**	**910,898**	**976,382**	**244,546**	**63,397**

Note: * Brazil refers to state of Ceará; **States of Tamil Nadu and Uttar Pradesh in India. The survey does not cover agriculture in Bangladesh and India, nor services in Bangladesh. Table 1 provides the distribution of the number of firms sampled in each country, by sector and firm size group.

We exclude micro-firms with fewer than 5 employees. Micro firms, particularly in developing countries, are more likely to be informal (Ulyssea, 2018), making them less likely to be captured in the sampling frame; and this would require further adjustment in the survey instrument and sampling design.[63] This size threshold is aligned with other firm-level standardized surveys with comparability across countries. The World Bank Enterprise Survey

[63] In addition, establishments below this threshold often lack the organizational structure to respond to some of the questions.

(WBES) also uses a threshold of 5 employees. The World Management Survey (WMS) uses a threshold of 50 employees.

We stratified the universe of establishments by firm size, sector of activity, and geographic regions. Our sample is representative across these dimensions. In the firm size stratification, we have three strata: small firms (5-19 employees), medium firms (20-99 employees), and large firms (100 or more employees). Regarding sector, for all countries, we stratified at least for agriculture (ISIC 01), food processing (ISIC 10), Wearing apparel (ISIC 14), Retail and Wholesale (ISIC 45, 46 and 47), other manufacturing (Group C, excluding food processing and apparel), and other Services (including all other firms, excluding retail). We use this sector structure of the data for most of the analysis in this paper. Additional sector stratification that were country specific included: Motor vehicles (ISIC 29); Leather (ISIC 15), Pharmaceutical (ISIC 21), and Motor vehicles (ISIC 29); and Land transport (ISIC 49), Finance (ISIC 64), and Health (ISIC 86).[64] In the geographic stratification, we use sub-national regions.

To calculate the optimal distribution of the sample, we followed a similar methodology as described by the World Bank (2009). The sample size for each country was aligned with the degree of stratification of the sample.

The data used in this paper corresponds to the first and second phase of the survey implementation. The surveys were administered between June 2019 and the end of 2021 by the World Bank in partnership with public or private local agencies across ten countries: Bangladesh, Brazil (the state of Ceará), Senegal, and Vietnam in the first phase until January 2020. In the second phase, conducted during the COVID-19 pandemic, after January 2020, included Burkina Faso, India (the states of Tamil Nadu and Uttar Pradesh), Ghana, Kenya, Poland, and the Republic of Korea. The mode of data collection was face-to-face before the pandemic and mostly on the telephone during the pandemic.

[64] These specific stratifications were taken into consideration when determining sampling weights.

Table A.3: Year and mode of data collection

Country	Year	Mode
Bangladesh	2019	Face-to-face
Brazil	2019	Face-to-face
Burkina Faso	2021	Telephone
Ghana	2021	Telephone
India	2020	Face-to-face
Kenya	2020	Telephone
Korea, Rep.	2021	Telephone
Poland	2021	Telephone
Senegal	2019	Face-to-face
Vietnam	2019	Face-to-face

A.3 Survey Weights

We construct the sampling weights of establishments in two steps. First, we compute design weights as reciprocals of inclusion probabilities. Then, to mitigate the risk of non-response bias, we adjust the design weights for non-response.

We adopt a stratified one stage element sampling design and randomly select establishments with equal probabilities within strata. Therefore, the inclusion probability of establishment k, within stratum isr (identified by industry i, size s, and region r), is:

$$\pi_{isr\ k} = \frac{n_{isr}}{N_{isr}} \tag{A.1}$$

where n_isr is the number of establishments targeted by the survey for stratum isr, and N_isr is the number of establishments in the sampling frame for the same stratum. Accordingly, the design weights of establishments are:

$$d_{isr\ k} = \frac{1}{\pi_{isr\ k}} = \frac{N_{isr}}{n_{isr}} \tag{A.2}$$

To adjust the design weights in equation A.2 for non-response we follow a simple Response Homogeneity Groups (RHG) approach (Särndal, Swensson and Wretman, 1992), with the groups determined by the strata. In other words, we assume that establishment response probabilities are the same within each stratum, but differ across different strata. Under the RHG approach assumptions, response probabilities can be estimated using the observed response rates within each group, and bias protection is obtained by dividing design weights by group-level response rates.

Denoting with the estimated response probability in stratum isr, and with m_isr the number of respondent establishments in the stratum (so that m_isrn_isr), the non-response adjusted weights can thus be written as follows:

$$w_{isr\ k} = \frac{d_{isr\ k}}{\hat{\theta_i sr}} = \frac{d_{isr\ k}}{m_{isr}/n_{isr}} = \frac{N_{isr}/n_{isr}}{m_{isr}/n_{isr}} = \frac{N_{isr}}{m_{isr}} \tag{A.3}$$

Note that the adjusted weights in equation A.3 are such that the distribution of our respondent sample across strata exactly matches the distribution of establishments in the sampling frame:

$$\sum_{k\ \in R_i sr} w_{isr\ k} = N_{isr} \tag{A.4}$$

where R_isr denotes the respondent sample for stratum isr.

Because of the different number of establishments in each country, when computing global statistics, we re-scale weights so that all countries are equally weighted.

A.4 Measures to minimize bias and measurement error during survey design and implementation

During the design of the survey questionnaire a number of good practices were considered in order to minimize different types of potential biases. The literature on survey design has identified three types of potential bias and measurement errors. These depend on whether they originate from the non-response, the enumerator or the respondent (Collins, 2003). In this section we describe all the steps taken in the design and implementation of the FAT survey to minimize these errors.

Non-response bias. A critical potential bias is associated with non-response in particular questions or non-participation in the survey (Gary, 2007). When this non-response follows a pattern that can be linked to factors correlated to the measured object, this non-response is associated with biases. For example, if more technology sophisticated firms refuse to participate because of fear to reveal commercial information, this would result in significant downward bias in estimating the level of technology sophistication. To minimize this risk, we try to maximize participation in the survey and follow three steps. First, we partner with national statistical offices and industry associations to use the most comprehensive and updated sampling frame available, as well as their experience on data collection, which are supported by endorsement letters from local institutions.[65] Having up to date contact details significantly improves response and minimized contact fatigue. Second, we follow a standard

[65]These procedures are in line with suggestions of good practice for implementation by Bloom et al. (2016).

protocol in which each firm is contacted several times to schedule an interview. We split the sample in different batches, following the order of randomization within stratum, and provide contact information of subsequent batches only after interviewers have shown evidence that they have exhausted the number of attempts to complete the initial list. Third, we monitor the implementation, validation of skip conditions and outliers (e.g. balance sheet information) in real time using standard survey software, and request that any missing information are completed through a follow up call, checked by supervisors. This minimizes risks that enumerators skip the order of their randomly assigned list of firms.

Enumerator bias and error counts. Minimizing cognitive biases in respondents in face to face and phone interviews starts with making sure that enumerators are able to implement the survey in a clear and consistent manner. To this end, the survey, training, and data collection processes are largely designed to minimize enumerator biases and data collection errors. First, to reduce the likelihood of coding errors, we use closed-ended questions, which make coding the answers a mechanical task, eliminating the reliance on the enumerator's interpretation of the answer and subjective judgement to code them, as it is the case with open-ended questions (Bloom et al., 2016). Second, to make sure that implementation is consistent across enumerators within and between country surveys, we implement the same standardized training in each country with enumerators, supervisors, and managers leading the data implementation. The training is led by team members directly involved in the elaboration of the questionnaire and implemented in local languages - English, French, Portuguese and Vietnamese,[66] and they include vignettes to ensure that enumerators understand the specific technologies they are asking about. The two to three days training consists of one general presentation about the project, covering the main motivation, relevance, coverage, and protocols that should be used to approach the interviewees and the review of the full questionnaire (question by question). The training material includes pictures of each technology mentioned in the survey both in general and sector-specific business functions, which are shared with enumerators. After going over the full questionnaire and clarifying any questions that emerge, the participants of the training conduct a mock interview using CAPI, under the supervision of our team.

Third, to guarantee that translations use words that are understood by firms managers, in each country we conduct a pre-test pilot of the questionnaire with firms out of the sample. A pilot of the questionnaire is implemented in each country with firms out of the sample. This allows to fine-tune questions to the local language, finalize the translation and select the most relevant examples in each question. After the pilot, our teams have the opportunity to discuss with the managers implementing the questionnaires and clarify any potential question

[66]In the case of Vietnamese, we used translation services support.

over the implementation process.

Fourth, to attain greater quality control during the data collection process, enumerators record the answers via *Computer-Assisted Personal Interviews* (CAPI) and *Computer-Assisted telephone Interviews* (CATI) software.[67] Using CAPI/CATI has clear advantages. First, it allows the use of logical conditions and skips which prevent data inputting errors and omitting questions, and also reduces the potential for abnormal values or non-response to specific questions. Second, it reduces substantially the time of implementation of the survey, increasing the quality of responses and minimizing survey fatigue. Supervisors are assigned to review all interviews, identifying missing values and abnormal responses. In addition, the CAPI/CATI system can identify when enumerators complete the survey too fast and other abnormal issues that can raise concerns about the quality of the interview. Finally, CAPI/CATI also allows for the core team to regularly monitor the data collection process and use standard algorithms to analyze the consistency of the data at different stages of data collection and by watches, thus providing continuous feedback and quality control.

Respondent bias. Perhaps the most important type of bias relates to cognitive biases from respondents. These biases can be large in surveys with open ended questions or where concepts can be largely subjective. Specifically, two broad groups of factors can trigger response errors: *cognitive*, which affect the comprehension of the questions, and *framing*, which may cause biased answers due to the perceived socially (un)desirability of the answers (Bertrand and Mullainathan, 2001). We take several steps to minimize this respondent bias. First, surveys need to be responded by the appropriate person in the firm that has all the information needed to respond. During the implementation of the screening process we ensure that the interview is arranged with the appropriate person or persons (Bloom et al., 2016). Senior managers (and in larger firms other managers such as plant managers) are asked to respond to the sections that cover the technologies used, and HR managers are asked to respond the questions on employment. Second, when possible use face-to-face interviews, which lead to higher response rates and lower respondent bias and measurement errors than web-based interviews.Only during the pandemic and due to existing mobility restrictions, we implemented surveys on the phone. Third, as discussed above, the use of a closed-ended design in the questionnaire reduces measurement error in the answers as the respondent is questioned about specific technologies (one at a time), and only when the presence of each of the possible technologies is established, the question about the most widely used technology is triggered. While this increases the length of the interview, it also increases the reliability

[67] Randomized survey experiments with household survey has demonstrated that a large number of errors observed in *Pen-and-Paper Personal Interview* (PAPI) data can be avoided in CAPI (Caeyers, Chalmers and De Weerdt, 2012).

of the data collected. Fourth, also as discussed above, we pre-pilot the questionnaire in each country to ensure that questions are clear in their wording in the specific geographical and cultural contexts, simple, and objective, so that the response does not require any subjective judgement (Bertrand and Mullainathan, 2001). Fifth, and more importantly, to avoid *social desirability bias*, by which respondents may overstate the use of more sophisticated technologies, the survey avoids the words "technology" and "sophistication" and employs more neutral terms such as "methods" and "processes". In addition, the survey is administered so that the respondent does not know all the possible technologies in a business functions before asserting whether a technology is used in the firm.[68] This reduces the risk that managers are framed to bias responses to the more advanced (socially desirable) technology, since they don't know what they will be asked in advance. Finally, when possible, enumerators are instructed to visually verify the information provided during the interviews. For example, in the case of use of a sophisticated production technology that can be visually identified in the shop floor.

A.5 Ex-post checks and validation exercises

While we apply best practices in survey design and implementation, it is important to perform validation checks once the data is collected. this allows us to measure the effectiveness of all these efforts to minimize bias and measurement error. In what follows, we describe some of these tests.

Minimizing potential non-response bias Our survey implementation was designed to minimize non-response through the use of well-prepared agencies and institutions to administer the survey and the presentation of adequate supporting letters to encourage participation. Table A.4 shows response rates by country, firm size group and sector. Response rates vary between 24% in Korea and 80% in Vietnam.

These are unweighted response rates calculated as the ratio between firms that responded the survey and the total number of firms in the sample which we attempted to conduct the interview. The high response rate for Vietnam is associated with the fact that the survey was implemented by the national statistical office. In most cases, these response rates are high relative to typical response rates in firm-level surveys, which for the U.S. are around 5 to 10 percent, and are consistent with response rates observed for WMS and MOPS (Bloom et al., 2016).[69]

To minimize potential non-response bias, we adjusted the sampling weights for unit

[68] It also allows for "don't know" options.

[69] The average response rate for the WMS is around 40 percent.The response rate for MOPS, implemented by the United States Census Bureau, was around 80 percent.

Table A.4: Response rates (by country)

Country	Response rate
Bangladesh	30%
Brazil	39%
Burkina Faso	45%
Ghana	49%
India	37%
Kenya	77%
Korea*	24%
Poland	47%
Senegal	57%
Vietnam	80%
Average across countries	49%

non-response. The non-response 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.

More importantly, to check the reliability of the instrument we implemented a series of ex-post tests in the first phase of the survey, focusing on countries we implemented the survey first. First, we study whether, in the sample of contacted firms, there are significant differences between those that responded and those that declined participating or could not be reached. The only information available in all firms we attempted to contact in the three sampling frames is the number of employees. Table A.5 tests whether there are differences in employment between the respondent and non-respondent groups, controlling for characteristics used for stratification. We find no significant differences in firm size between respondents and non-respondents in any of the three countries.

Second, under the premise that any systematic relationship between firm characteristics and participation is continuous in their reluctance to participate in the survey, we can learn about sample differences between respondents and non-respondents by comparing firms across different percentiles of the distribution of the number of attempts it took for them to respond the survey.[70] For Senegal, we explore whether after controlling for observable characteristics, there are significant differences in average technology sophistication in GBFs between firms that required a larger number of attempts to be contacted (top quartile) and those that did not. Table A.6 shows that there are no statistically significant differences in technology sophistication between the two groups.

Third, we compare firms that were in the first sample list provided to enumerators and

[70] Behaghel et al. (2015) infer the reluctance to participate in the survey from the number of attempts that it take for a firm to accept the request.

Table A.5: Comparison of firm size between respondents vs non-respondents

VARIABLES	Brazil	Vietnam	Senegal
Respondents (FAT)	2.52	52.34	-4.92
	(22.19)	(80.27)	(6.63)
Observations	1,754	1,500	3,075
R-squared	0.129	0.172	0.237
Controls:			
Sector	√	√	√
Size group	√	√	√
Region	√	√	√

Note: *** p<0.01, ** p<0.05, * p<0.1. Data are from the list of firms contacted by the enumerators. For each country, the level of employment was regressed on a dummy for respondent while controlling for stratification such as sectors, size groups (small, medium, and large), and regions. Estimates for Vietnam are based on the original list of 1500 firms, with 1346 respondents and 154 non-respondents. Robust standard errors in parenthesis.

Table A.6: Comparison of technology sophistication between high and low number of attempts

VARIABLES	Senegal	Senegal
Top quartile of attempts (4 or more)	-0.021	-0.027
	(0.020)	(0.019)
Observations	1,753	1,666
R-squared	0.377	0.437
Controls:		
Sector	√	√
Size group	√	√
Region	√	√
Age		√
Exporter		√
Foreign owned		√

Note: *** p<0.01, ** p<0.05, * p<0.1. Data are from the Senegal FAT survey with information on the number of attempts to complete interview at the firm level. Technology sophistication is regressed on a dummy for the top quartile of the number of attempts (4 or more) with controls for the stratification (sectors, size groups, and regions) and/or firm characteristics (age groups, exporter, and foreign owned). Robust standard errors in parenthesis.

those in subsequent lists. Table A.7 show that there are no statistically significant differences between the two groups.

In each of these exercises, we find no statistical difference in the number of employees, technology sophistication, wages, and share of workers by skill and education between firms in the group that proxies for the response sample and the group of firms that proxies for the

Table A.7: Comparison of technology sophistication between original and replacement sample

VARIABLES	Brazil	Brazil	Vietnam	Vietnam	Senegal	Senegal
Original sample	-0.014	-0.037	0.030	0.043	0.021	0.028
	(0.048)	(0.047)	(0.050)	(0.048)	(0.018)	(0.018)
Observations	638	637	1,484	1,484	1,753	1,666
R-squared	0.299	0.335	0.262	0.320	0.377	0.437
Controls:						
Sector	√	√	√	√	√	√
Size group	√	√	√	√	√	√
Region	√	√	√	√	√	√
Age		√		√		√
Exporter		√		√		√
Foreign owned		√		√		√

Note: *** p<0.01, ** p<0.05, * p<0.1. Data are from the Brazil, Vietnam, and Senegal FAT surveys. For each country, technology sophistication ($MOST_j$) is regressed on a dummy for the original sampling list with controls for the stratification (sectors, size groups, and regions) and/or firm characteristics (age groups, exporter, and foreign owned). Robust standard errors in parenthesis.

non-response sample.[71]

Minimizing enumerator bias. To minimize the potential for enumerators to introduce biases when administering the survey, we conduct in each country standard training and piloting prior to going to the field, and we also evaluate the quality of the interviews. Specifically, we conduct ex-post tests on the differences in sophistication in abnormal interviews by running regressions of firm-level sophistication on enumerator dummies and firm controls as discussed in the text. Table A.8 shows that enumerator dummies are not significant for Brazil, Ghana, India, and Korea. For Bangladesh, Kenya, Senegal, and Vietnam, less than 12% of enumerator dummies are statistically significant. Table A.9 compares the average technology sophistication ($MOST_j$) for GBF, excluding the firms with abnormal enumerators and in the entire sample. We find no economic or statistical difference between mean sophistication in these countries.

Minimizing respondent bias. A critical factor to minimize respondent bias is to identify the right respondent. The protocol for the implementation of the survey required that the survey should be ideally answered by the top manager. About 47% of the survey was answered by the owner or CEOs, while the other responses included factory managers, other managers, administrative staff, and accountants. Almost 80% of the interviews were

[71]See Table A.5 to A.11 in Appendix A.

Table A.8: Analysis of enumerator bias

VARIABLES	Brazil	Vietnam	Senegal	Bangladesh
Share of Significantly Different Interviewers	0	0.09	0.08	0.11
Number of Significantly Different Interviewers	0	13	2	4
Number of Interviewers	8	145	25	37
	Ghana	India	Korea	Kenya
Share of Significantly Different Interviewers	0	0	0	0.006
Number of Significantly Different Interviewers	0	0	0	9
Number of Interviewers	44	18	9	450

Note: Data from the Firm-level Adoption of Technology (FAT) surveys in Brazil, Vietnam, Senegal, Bangladesh, Ghana, India, Korea and Kenya. Significantly different interviewers are identified from the regressions of employment on interviewer dummies with controlling for stratification information (e.g., sector, size, and region). For each country, the share of significantly different interviewers is computed by dividing the number of interviews conducted by significantly different interviewers by the total number of interviews.

conducted through one visit in person interview with the main respondent. In circumstances in which the main respondent did not have all the information about a general topic of the questionnaire, especially in modules B and C, they were requested to consult with other colleagues.

To assess the relevance of response bias, we conduct a parallel pilot in Kenya where we re-interview 100 randomly selected firms with a short version of the questionnaire. For those firms, we randomly select three business functions and ask about the presence of the relevant technologies.[72] Both the original and back-end interviews in the pilot are conducted by phone by different interviewers.

Despite using phone interviews which are subject to greater measurement error than face-to-face interviews, comparison of answers from the pilot reveals that 73% of the answers were the same across the two interviews.[73] We estimate a probit model to assess the likelihood of consistent answers between the original and the back-check interviews, controlling for firm-level fixed-effect. Reporting the use of a technology in the back-check interview is associated with 80.6% of 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

[72] The pilot coincided with the beginning of the data collection for phase two which includes new countries, and Kenya is one of them. Despite the fact that Kenya is not in the sample, the pilot is informative about the significance of response bias. The re-interviews produce 1,661 answers (106 interviews times 3 business functions times an average of 5.2 technologies per function).

[73] The consistency ranges 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.

Table A.9: Difference in technology sophistication in general business functions with and without outlying enumerators

	All Sample	Sample Without Different Enumerators	Difference
Vietnam			
Mean	1.934	1.947	-0.013
SE	(0.012)	(0.012)	(0.017)
Observations	1,499	1,341	
Senegal			
Mean	1.406	1.404	0.002
SE	(0.011)	(0.011)	(0.016)
Observations	1,786	1,784	
Bangladesh			
Mean	1.482	1.458	0.024
SE	(0.015)	(0.015)	(0.021)
Observations	903	798	
Kenya			
Mean	1.938	1.936	.002
SE	(0.020)	(0.020)	(0.029)
Observations	1305	1296	

Note: *** p<0.01, ** p<0.05, * p<0.1. Data from the Firm-level Adoption of Technology (FAT) surveys in Vietnam, Senegal, Bangladesh and Kenya. Brazil, Ghana, India and Korea are excluded because they do not include significantly different interviewers. The average of technology sophistication in general business functions ($MOST_j$) is compared between all sample and sample excluding significantly different enumerators. Standard errors in parenthesis.

with a 29.3% likelihood of being reported in the original survey.

RAIS validation exercise

Some final ex-post checks were conducted with the Brazil data and takes advantage of the fact that we have access to the RAIS administrative data, which is a matched employer-employee dataset that covers the universe of firms in the sampling frame. This allows us to compare variables in RAIS with variables we collected in FAT for the same firms.

First, we analyze the correlation between the value-added per worker and our technology measures (GBF) and (SBF) from FAT and average wages from RAIS. Table A.10 reports the point estimates of regressing firm-level FAT variables on the log or average wages per worker from RAIS and a set of firm-level controls. The FAT variables are log of value added per worker (column 1), and average technology sophistication (GBF, column 2, and SSBF, column 3). In all three cases we find strong positive associations between the FAT and the

RAIS variables.

Table A.10: Relationship between FAT survey variables and log of wages from administrative data for Brazil

Variable	(1) ln(VAPW)	(2) GBF	(3) SSBF
ln(Wage) RAIS	0.873***	0.507***	0.549***
	(0.200)	(0.121)	(0.138)
Observations	530	675	568
R-squared	0.217	0.354	0.230
Controls:			
Sector FE	$\surd$	$\surd$	$\surd$
Region FE	$\surd$	$\surd$	$\surd$
Size group	$\surd$	$\surd$	$\surd$
Age	$\surd$	$\surd$	$\surd$
Exporter	$\surd$	$\surd$	$\surd$
Foreign owned	$\surd$	$\surd$	$\surd$

Note: *** p<0.01, ** p<0.05, * p<0.1. Average wage information for each establishment is obtained from the 2017 *Relação Anual de Informações Sociais* (RAIS) merged with the Firm-level Adoption of Technology (FAT) data used in this exercise, including value added per worker (VAPW), the technology adoption index ($MOST_j$) for GBF and SSBF, and firm characteristics used as controls. Robust standard errors in parenthesis.

Second, we compare the differences between labor-related indicators from a matched employer-employee administrative data for firms in FAT versus the universe of firms. To perform this comparisons we obtained the weighted average for firms in FAT, using the weights we constructed as described in section A3 and compare it with the average for all firms in RAIS that are part of our universe for the State of Ceará, in Brazil. [74] We then perform a t-test to compare the differences. Table A.11 shows that the differences are not statistically significant.

Overall, these ex post checks appear to validate the quality of the data collected.

[74]The variables are number of workers, average wages, share of workers with college degree, share of low skilled, and share of high-skilled workers, where high- and low-skilled workers are defined as in Autor and Dorn (2013).

Table A.11: Comparison between FAT sample and RAIS data (universe)

	Number of employees	Average wage	Share college	Share low-skill	Share high high-skill
FAT Average (weighted)	28.55	1,311.89	0.05	0.16	0.42
RAIS Average (universe)	23.85	1,349.29	0.05	0.17	0.39
Estimate (RAIS - FAT)	-4.70	37.40	0.00	0.00	-0.03
Standard Error	(3.08)	(29.77)	(0.01)	(0.01)	(0.02)
T-Statistic	-1.52	1.26	0.55	0.20	-1.64

Note: *** p<0.01, ** p<0.05, * p<0.1. Data from the 2017 *Relação Anual de Informações Sociais* (RAIS) and the Firm-level Adoption Technology (FAT) survey in Brazil. The estimates from RAIS data are unweighted, and those from FAT surveys are weighted by the sampling weights. Robust standard errors in parenthesis.

B Derivations

B.1 Quantification of stability of technology curves to country groupings

In this section, we provide details about the quantification of the relative variance in technology sophistication induced by differences across country groupings in the technology curve relative to the variance induced by the baseline technology curve. Let $S_{f,d}$ the average level of sophistication in function f across the establishments that are in the d decile of the distribution of average sophistication (e.g. $MOST_j$) and S_d is the average sophistication in the establishments in the d decile. The technology curve for function f plots $S_{f,d}$ against S_d. Now we consider three technology curves for each of the three country groupings. In the horizontal axis we still plot S_d, but in the vertical now we plot the average value of $S_{f,j}$ for the subsample of establishments in decile d that belong to the country grouping c. We denote this by $S_{f,c,d}$.The technology curves in Figure 7 plot $S_{f,c,d}$ against S_d.

We define the variance of the (baseline) technology curve for a function f as the variance of the vertical gaps between the technology curve and the 45-degree line. Adding up across functions, this can be expressed as $\sum_f \sum_d \omega_{f,d}(S_{f,d}-S_d)^2$, where $\omega_{f,d}$ is the weight of function f for the d decile. We define the variance across country groupings in the technology curves as $\sum_f \sum_c \sum_d \omega_{f,c,d}(S_{f,c,d} - S_{f,d})^2$. The sum of the variance in the technology curves and the variance across country groupings in the technology curves is equal to the total variance in the country-grouped technology curves which is given by $\sum_f \sum_c \sum_d \omega_{f,c,d}(S_{f,c,d} - S_d)^2$.

B.2 Model derivations

In this section we proof the theoretical results of 4. **Proposition 1** The first order conditions for the optimal adoption problem are

$$\Pi'_j(a_j)\frac{\frac{(1-\sigma)}{\sigma}\left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}s_{f,j}}}{\sum_{f=1}^{N_f}\frac{\varepsilon_f}{\sigma}\left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}s_{f,j}}} = C_j C_{f,X} e^{s_{f,j}} \tag{B.5}$$

Taking logs and isolating $s_{f,j}$, we obtain

$$s_{f,j} = \sigma ln(\frac{(1-\sigma)}{\sigma}) + ln(\Omega_f) - \sigma ln(C_{f,X})$$
$$+ \sigma ln(\Pi'_j(a_j)) - \sigma ln(C_j) - \sigma ln(\sum_{f=1}^{N_f} \frac{\varepsilon_f}{\sigma} \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma} s_{f,j}}) + \varepsilon_f a_j \quad \text{(B.6)}$$

Defining $\overline{\varepsilon}_j/\sigma \equiv \sum_f \frac{\widetilde{\varepsilon}_f}{\sigma} \omega_{f,j}$ with $\omega_{f,j} \equiv \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\gamma_f \sigma} s_{f,j}}$, replacing the RHS of B.6 for $s_{f,j}$ and solving for $\overline{\varepsilon}_j/\sigma$, we obtain

$$\overline{\varepsilon}_j/\sigma \equiv \sum_f \frac{\varepsilon_f}{\sigma} \omega_{f,j} = \left[\left(\frac{\sigma}{1-\sigma}\right) \sum_f \frac{\varepsilon_f}{\sigma} \left(\Omega_f e^{\varepsilon_f a_j} \left(\frac{C_{f,X} C_j}{\Pi'_j(a_j)}\right)^{1-\sigma}\right)\right]^{1/\sigma} \quad \text{(B.7)}$$

This proves part A. To proof part B, substitute (B.6) into the implicit definition of a_j (8) to obtain the expression in the proposition.

Corollary 1: We define the average establishment as one with average marginal costs of adoption $\overline{C} = \sum_j C_j/N_j$ and $C_{f,\overline{X}} = \sum_j C_{f,X}/N_j$, and marginal profits $\overline{\Pi'}(\overline{a}) = \sum_j \Pi'(\overline{a})/N_j$, where $\overline{a}$ and $\overline{\varepsilon}$ are defined by the following system:

$$\overline{\varepsilon}/\sigma = \left[\left(\frac{\sigma}{1-\sigma}\right) \sum_f \frac{\varepsilon_f}{\sigma} \left(\Omega_f e^{\varepsilon_f \overline{a}} \left(\frac{C_{f,\overline{X}} \overline{C}}{\overline{\Pi'}(\overline{a})}\right)^{1-\sigma}\right)\right]^{1/\sigma} \quad \text{(B.8)}$$

$$\sum_{f=1}^{N_f} \Omega_f e^{\varepsilon_f \overline{a}} \left(\frac{C_{f,\overline{X}} \overline{C} \overline{\varepsilon}}{\overline{\Pi'}(\overline{a}) \sigma}\right)^{1-\sigma} \left(\frac{1-\sigma}{\sigma}\right)^{\sigma} = 1. \quad \text{(B.9)}$$

We start by taking a first order expansion of (B.7) around the TechFP, marginal costs and weighted technology elasticity of the average establishment to obtain:

$$\sigma ln(\overline{\varepsilon}_j/\overline{\varepsilon}) \simeq (a_j - a) \left\{\overline{\varepsilon}_2 - (1-\sigma) \frac{\overline{\Pi''}(\overline{a})}{\overline{\Pi'}(\overline{a})}\right\} + (1-\sigma) \left(\sum_f \varpi_f^{\varepsilon} ln(C_{f,X}/C_{f,\overline{X}}) + ln(C_j/C) - ln(\Pi'_j(\overline{a})/\overline{\Pi'}(\overline{a}))\right) \quad \text{(B.10)}$$

where

$$\overline{\varepsilon}_2 \equiv \frac{\sum_f \varepsilon_f^2 \left(\Omega_f e^{\varepsilon_f \overline{a}} \left(C_{f,\overline{X}}\right)^{1-\sigma}\right)}{\sum_f \varepsilon_f \left(\Omega_f e^{\varepsilon_f \overline{a}} \left(C_{f,\overline{X}}\right)^{1-\sigma}\right)} \quad \text{(B.11)}$$

and

$$\varpi_f^{\varepsilon} = \frac{\frac{\varepsilon_f}{\sigma} \Omega_f e^{\varepsilon_f \overline{a}} C_{f,\overline{X}}^{1-\sigma}}{\sum_f \frac{\varepsilon_f}{\sigma} \Omega_f e^{\varepsilon_f \overline{a}} C_{f,\overline{X}}^{1-\sigma}} \quad \text{(B.12)}$$

A linear approximation of (13) yields:

$$\left(\bar{\varepsilon}\left(\frac{1-\sigma}{\sigma}\right)^{1+\sigma} - (1-\sigma)\frac{\overline{\Pi''}(\overline{a})}{\overline{\Pi'}(\overline{a})}\right)(a_j - \overline{a})$$
$$+ (1-\sigma)\left\{-ln(\Pi_j'/\Pi') + \sum_f \omega_f ln(C_{f,X}/C_{f,\overline{X}}) + ln(C_j/\overline{C_j}) + ln(\overline{\varepsilon}_j/\overline{\varepsilon})\right\} \simeq 0 \quad \text{(B.13)}$$

where

$$\omega_f = \frac{\Omega_f e^{\varepsilon_f \overline{a}} C_{f,\overline{X}}^{1-\sigma}}{\sum_f \Omega_f e^{\varepsilon_f \overline{a}} C_{f,\overline{X}}^{1-\sigma}} \quad \text{(B.14)}$$

Substituting in (B.10) and rearranging we obtain

$$(a_j - \overline{a})\left(\frac{1-\sigma}{\sigma}\right)\left(\bar{\varepsilon}\left(\frac{1-\sigma}{\sigma}\right)^{\sigma} + \bar{\varepsilon}_2 - \frac{\overline{\Pi''}(\overline{a})}{\overline{\Pi'}(\overline{a})}\right) \simeq$$
$$\left(\frac{1-\sigma}{\sigma}\right)\left(ln(\Pi_j'/\Pi') - ln(C_j/\overline{C_j}) - \sum_f \left(\sigma\omega_f + (1-\sigma)\varpi_f^{\varepsilon}\right) ln(C_{f,X}/C_{f,\overline{X}})\right) \quad \text{(B.15)}$$

Isolating a_j, we can express (B.15) as (15), and it follows from the fact that $\sigma < 1$ and that $\overline{\Pi''}(.) < 0$ that the signs of the coefficients in the approximation are as stated in Corollary 1.

Proposition 2: We rewrite (11) as

$$s_{f,j} = \sigma ln\left(\frac{1-\sigma}{\sigma}\right) + ln\left(\Omega_f\right) + \sigma\left(ln(\Pi_j'(a_j)) - ln(C_j)\right) - \sigma ln(\overline{\varepsilon}_j/\sigma)$$
$$+ \left(\varepsilon_f - \bar{\varepsilon}\right)\left(a_j - \bar{a}\right) - \bar{\varepsilon}\bar{a} + \bar{\varepsilon}a_j + \varepsilon_f\bar{a} - \sigma(ln(C_{f,X}) - ln(C_{f,\overline{X}})) - \sigma ln(C_{f,\overline{X}}) \quad \text{(B.16)}$$

Rearranging, we can express this as

$$s_{f,j} = \overbrace{\sigma ln\left(\frac{1-\sigma}{\sigma}\right) + ln\left(\Omega_f\right) + \varepsilon_f\bar{a} - \sigma ln(C_{f,\overline{X}})}^{\kappa_f}$$
$$+ \overbrace{\sigma\left(ln(\Pi_j'(a_j)) - ln(C_j)\right) - \sigma ln(\overline{\varepsilon}_j/\sigma) - \bar{\varepsilon}\bar{a} + \bar{\varepsilon}a_j}^{\kappa_j} + \left(\varepsilon_f - \bar{\varepsilon}\right)\left(a_j - \bar{a}\right) - \sigma(ln(C_{f,X}) - ln(C_{f,\overline{X}})) \quad \text{(B.17)}$$

We can now define $u_{f,j} = s_{f,j} - \kappa_f - \kappa_j$, and therefore,

$$Var(u_{f,j}) = a_j^2 Var(\varepsilon_f) + \sigma^2 Var(ln(C_{f,X})) - \sigma a_j Cov(\varepsilon_f, ln(C_{f,X})). \tag{B.18}$$

B.3 Business function-specific technical change

In this section, we consider an extension of (8) that introduces business function-specific technical change. This is captured by a parameter γ_f that may differ from 1. In some functions, an increase in sophistication is associated with a larger increase in TechFP than in others. not because the function is more relevant but because the technological gains that bring the change in sophistication are more significant. Formally, the new aggregator is

$$\sum_{f=1}^{N_f} \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}\gamma_f s_{f,j}} = 1 \tag{B.19}$$

with

$$\frac{\partial a_j}{\partial s_{f,j}} = \frac{\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}} e^{\frac{\sigma-1}{\sigma}\gamma_f s_{f,j}}}{\sum_{f=1}^{N_f} \frac{\varepsilon_f}{\sigma} \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}\gamma_f s_{f,j}}} \frac{(1-\sigma)}{\sigma}\gamma_f \tag{B.20}$$

which is increasing in γ_f.

The first order conditions now become

$$\Pi_j'(a_j)\frac{\partial a_j}{\partial s_{f,j}} = \frac{\partial C_j(s_{f,j})}{\partial s_{f,j}} \tag{B.21}$$

$$\Pi_j'(a_j)\frac{\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}} e^{\frac{\sigma-1}{\sigma}\gamma_f s_{f,j}}}{\sum_{f=1}^{N_f} \frac{\varepsilon_f}{\sigma} \left(\Omega_f^{\frac{1}{\sigma}} e^{\frac{\varepsilon_f a_j}{\sigma}}\right) e^{\frac{\sigma-1}{\sigma}\gamma_f s_{f,j}}} \frac{(1-\sigma)}{\sigma}\gamma_f = C_j C_{f,X} e^{s_{f,j}} \tag{B.22}$$

Taking logs and isolating $s_{f,j}$ we obtain:

$$\begin{aligned} s_{f,j} = \frac{\sigma ln\left(\frac{1-\sigma}{\sigma}\right) + ln\left(\Omega_f\right) + ln(\gamma_f)}{(\gamma_f(1-\sigma)+\sigma)} &+ \frac{\sigma\left(ln(\Pi_j'(a_j)) - ln(C_j)\right) - \sigma ln(\overline{\varepsilon}_j/\sigma)}{(\gamma_f(1-\sigma)+\sigma)} \\ &+ \frac{\varepsilon_f}{(\gamma_f(1-\sigma)+\sigma)}a_j - \sigma\frac{ln(C_{f,X})}{(\gamma_f(1-\sigma)+\sigma)} \end{aligned} \tag{B.23}$$

where

$$\frac{\overline{\varepsilon}_j}{\sigma} = \left[\sum_f \frac{\varepsilon_f}{\sigma} \left(\frac{\Omega_f}{\gamma_f^{(1-\sigma)\gamma_f}} \left(\frac{\sigma}{1-\sigma} \right) e^{\varepsilon_f a_j} \left(\frac{C_{f,X} C_j}{\Pi_j'(a_j)} \right)^{(1-\sigma)\gamma_f} \right)^{\frac{1}{\gamma_f(1-\sigma)+\sigma}} \right]^{(\gamma_f(1-\sigma)+\sigma)/\sigma} \tag{B.24}$$

It follows from (B.23) that $s_{f,j}$ can be expressed as

$$s_{f,j} = \widetilde{\kappa}_f + \frac{\widetilde{\kappa}_j}{\gamma_f(1-\sigma)+\sigma} + \frac{\varepsilon_f}{(\gamma_f(1-\sigma)+\sigma)} a_j - \sigma \frac{ln(C_{f,X})}{(\gamma_f(1-\sigma)+\sigma)} \tag{B.25}$$

Conducting a first order approximation of the last three terms around $\gamma_f = 1$,we obtain

$$s_{f,j} \simeq \widetilde{\kappa}_f + \widetilde{\kappa}_j - \widetilde{\kappa}_j (1-\sigma)(\gamma_f - 1) + \varepsilon_f a_j (1 - (1-\sigma)(\gamma_f - 1)) - \sigma ln(C_{f,X}) (1 - (1-\sigma)(\gamma_f - 1)). \tag{B.26}$$

which corresponds to (22). The specification we estimate in the first stage is

$$s_{f,j} = \widetilde{\kappa}_f + \widetilde{\kappa}_j + \rho_f \overline{a}_j + \beta_f * X_j + v_{f,j} \tag{B.27}$$

The estimate of ρ_f,$\widehat{\rho}_f = \frac{Cov(s_{f,j} - \widetilde{\kappa}_f - \widetilde{\kappa}_j - \beta_f * X_j, \overline{a}_j)}{Var(\overline{a}_j)}$ which yields

$$\widehat{\rho}_f = \varepsilon_f (1 - (1-\sigma)(\gamma_f - 1)) - \frac{Cov(\overline{a}_j, \widetilde{\kappa}_j)}{Var(\overline{a}_j)} (1-\sigma)(\gamma_f - 1) \tag{B.28}$$

which corresponds to equation (23).

Let $\widehat{s}_{f,j} \equiv s_{f,j} - \widehat{\kappa}_f - \widehat{\beta}_f * X_j = (\widetilde{\kappa}_j + \varepsilon_f a_j)(1 - (1-\sigma)(\gamma_f - 1))$.

$$\widehat{a}_j = \frac{Cov(\widehat{s}_{f,j}, \widehat{\rho}_f D_j)}{Var(\widehat{\rho}_f)} \tag{B.29}$$

where D_j is a dummy that takes the value of 1 if the establishment is j, and 0 for all the other establishments.

The denominator of (B.29) is

$$\begin{aligned} Var(\widehat{\rho}_f) =& Var(\varepsilon_f (1 - (1-\sigma)(\gamma_f - 1))) + \left(\frac{Cov(\overline{a}_j, \widetilde{\kappa}_j)}{Var(\overline{a}_j)} (1-\sigma) \right)^2 Var(\gamma_f) \\ & - \frac{Cov(\overline{a}_j, \widetilde{\kappa}_j)}{Var(\overline{a}_j)} (1-\sigma) Cov(\gamma_f, \varepsilon_f (1 - (1-\sigma)(\gamma_f - 1))) \end{aligned} \tag{B.30}$$

While the numerator is

$$
\begin{aligned}
Cov(\widehat{s}_{f,j}, \widehat{\rho}_f D_j) &= Cov\left(\widehat{s}_{f,j}, \left(\varepsilon_f(1-(1-\sigma)(\gamma_f-1)) - \frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}(1-\sigma)(\gamma_f-1)\right) D_j\right) \\
&= Cov(\widehat{s}_{f,j}, (\varepsilon_f(1-(1-\sigma)(\gamma_f-1))) D_j) - \frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}(1-\sigma) Cov(\widehat{s}_{f,j}, \gamma_f D_j) \\
&= a_j Var(\varepsilon_f(1-(1-\sigma)(\gamma_f-1))) + \widetilde{\kappa}_j \frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}(1-\sigma)^2 Var(\gamma_f) \\
&\quad -(1-\sigma)Cov(\gamma_f, \varepsilon_f(1-(1-\sigma)(\gamma_f-1)))\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}\left\{a_j + \frac{\widetilde{\kappa}_j}{\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}}\right\}
\end{aligned}
\tag{B.31}
$$

Combining (B.30) and (B.31) we obtain

$$
\widehat{a}_j = a_j * (\omega_1 + \omega_3) + \frac{\widetilde{\kappa}_j}{\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}}(\omega_2 + \omega_3) \tag{B.32}
$$

where $\omega_1 = \frac{Var(\varepsilon_f(1-(1-\sigma)(\gamma_f-1)))}{Var(\widehat{\rho}_f)}$, $\omega_2 = \frac{(Cov(\overline{a}_j,\widetilde{\kappa}_j)/Var(\overline{a}_j))^2(1-\sigma)^2 Var(\gamma_f)}{Var(\widehat{\rho}_f)}$, $\omega_3 = -\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)/Var(\overline{a}_j)(1-\sigma)Cov(\gamma_f,\varepsilon_f\left(1-(1-\sigma)(\gamma_f-1)\right))}{Var(\widehat{\rho}_f)}$ and $\omega_1 + \omega_2 + \omega_3 = 1$.

Combining (B.28) and (B.32) we obtain the following expression for $\widehat{\rho}_f \widehat{a}_j$:

$$
\begin{aligned}
\widehat{\rho}_f \widehat{a}_j &= \varepsilon_f\left(1-(1-\sigma)(\gamma_f-1)\right) a_j(\omega_1+\omega_3) - \frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}(1-\sigma)(\gamma_f-1) a_j(\omega_1+\omega_3) \\
&\quad + \varepsilon_f\left(1-(1-\sigma)(\gamma_f-1)\right) \frac{\widetilde{\kappa}_j}{\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}}(\omega_2+\omega_3) - \frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}(1-\sigma)(\gamma_f-1)\frac{\widetilde{\kappa}_j}{\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)}}(\omega_2+\omega_3)
\end{aligned}
\tag{B.33}
$$

Proof of Proposition 3:

A. The $Var(\widehat{\rho}_f)$ is given in expression (B.30). To study its limit as $Var(\widetilde{\kappa}_j)$ tends to 0, note that the first term is independent of $Var(\widetilde{\kappa}_j)$. $\frac{Cov(\overline{a}_j,\widetilde{\kappa}_j)}{Var(\overline{a}_j)} = \sqrt{\frac{Var(\widetilde{\kappa}_j)}{Var(\overline{a}_j)}} * Corr(\overline{a}_j, \widetilde{\kappa}_j)$. $Corr(\overline{a}_j, \widetilde{\kappa}_j)$ and $Var(\overline{a}_j)$ are independent of $Var(\widetilde{\kappa}_j)$. Therefore, as $Var(\widetilde{\kappa}_j)$ tends to 0, the second term in (B.30) tends to 0, and this proofs the statement.

B.

$$
E(\widehat{a}_j|a_j) = a_j * (\omega_1 + \omega_3) + E(\widetilde{\kappa}_j|a_j) * \frac{Var(\overline{a}_j)(\omega_2+\omega_3)}{Cov(\overline{a}_j,\widetilde{\kappa}_j)} \tag{B.34}
$$

$$E(\widehat{a}_j|a_j) = a_j * (\omega_1 + \omega_3) + \sqrt{\frac{Var(\widetilde{\kappa}_j)}{Var(\overline{a}_j)}} * Corr(\overline{a}_j, \widetilde{\kappa}_j) \frac{Var(\overline{a}_j)(\omega_2 + \omega_3)}{Cov(\overline{a}_j, \widetilde{\kappa}_j)} * a_j \quad \text{(B.35)}$$

By the same argument, one can show that ω_2 and ω_3 tend to 0, and ω_1 tends to 1 as $Var(\widetilde{\kappa}_j)$ tends to 0. Therefore, $Var(\widetilde{\kappa}_j)$ tends to 0, $E(\widehat{a}_j|a_j)$ tends to a_j.

C. To be written.

C Additional Figures and Tables

In this section we present additional figures and tables referred in the draft. **??** plots the technology curves using the MAX measure of technology sophistication as described in section 3.3.

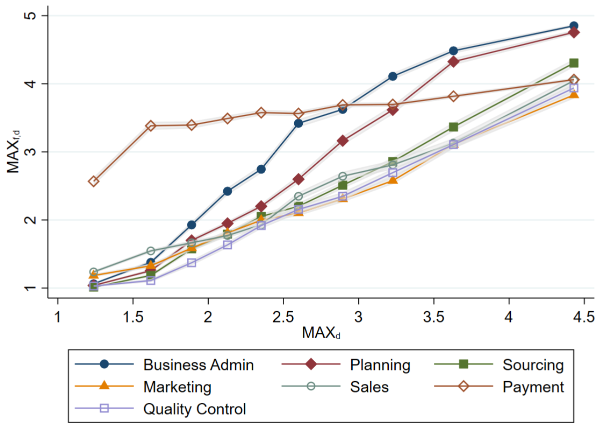
Table C.12 contains the estimates of the slopes of technology curves of each business function, $\widehat{\varepsilon}_f$. These are obtained from equation (18).

Table C.12: $\widehat{\varepsilon}_f$ by sector

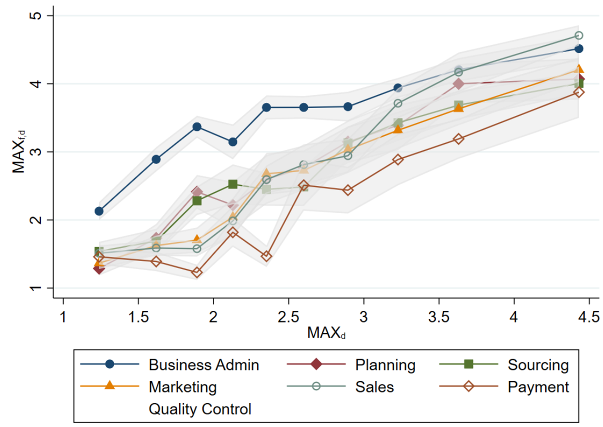
1-digit sector	**Sector**	**Business function**	$\widehat{\varepsilon}_f$	**1-digit sector**	**Sector**	**Business function**	$\widehat{\varepsilon}_f$
General	General	Business Administration	1.25	Manufactuyring	Car	Assembly	1.01
General	General	Production Planning	1.18	Manufacturing	Car	Pressing	0.41
General	General	Sourcing	0.85	Manufacturing	Car	Painting	0.29
General	General	Marketing	0.71	Manufacturing	Car	Plastic Injection & Molding	0.66
General	General	Sales	0.50	Manufacturing	Car	Management	1.54
General	General	Payment	0.36	Manufacturing	Car	Fabrication - Automotive	0.75
General	General	Quality Control	0.80	Manufacturing	Pharma	Facilities	0.96
Agriculture	Crops	Land Preparation	1.18	Manufacturing	Pharma	Weighing	0.61
Agriculture	Crops	Irrigation	1.43	Manufacturing	Pharma	Compounding	1.34
Agriculture	Crops	Pest Control	1.22	Manufacturing	Pharma	Encapsulation	0.93
Agriculture	Crops	Harvesting	1.10	Manufacturing	Pharma	Quality Control	1.73
Agriculture	Crops	Storage	1.18	Manufacturing	Pharma	Packaging	1.57
Agriculture	Crops	Packing	0.87	Manufacturing	Pharma	Fabrication - Pharma	0.84
Agriculture	Lifestock	Breeding	1.01	Services	Wholesale/Retail	Customer Service	0.31
Agriculture	Lifestock	Nutrition	0.31	Services	Wholesale/Retail	Pricing	0.56
Agriculture	Lifestock	Animal healthcare	0.64	Services	Wholesale/Retail	Merchandising	0.42
Agriculture	Lifestock	Herd management	0.44	Services	Wholesale/Retail	Inventory	0.72
Agriculture	Lifestock	Transport of Livestock	0.42	Services	Wholesale/Retail	Advertisement	0.88
Manufacturing	Food processing	Input Test	0.81	Services	Finance	Customer Service	0.61
Manufacturing	Food processing	Mixing Blending Cooking	0.27	Services	Finance	ID Verification	0.77
Manufacturing	Food processing	Anti-bacterial	0.76	Services	Finance	Loan Application	0.97
Manufacturing	Food processing	Packaging	0.74	Services	Finance	Loan Approval	0.95
Manufacturing	Food processing	Food Storage	0.91	Services	Finance	Operational Support Area	1.14
Manufacturing	Food processing	Fabrication	0.42	Services	Land transportation	Planning	1.30
Manufacturing	Apparel	Design	1.20	Services	Land transportation	Execution	0.90
Manufacturing	Apparel	Cutting	0.38	Services	Land transportation	Monitoring	1.39
Manufacturing	Apparel	Sewing	0.10	Services	Land transportation	Performance Measurement	1.19
Manufacturing	Apparel	Finishing	0.42	Services	Land transportation	Maintenance	1.11
Manufacturing	Apparel	Fabrication	0.11	Services	Health	Infrastructure and Machines	0.23
Manufacturing	Leather	Design	1.64	Services	Health	Scheduling Appointments	0.33
Manufacturing	Leather	Cutting	1.09	Services	Health	Management of Patient Records	0.71
Manufacturing	Leather	Sewing	0.83	Services	Health	Procedures	0.53
Manufacturing	Leather	Finishing	0.97				
Manufacturing	Leather	Fabrication - Leather	0.13				

Table C.13: Sector-specific estimates of slope of technology curves

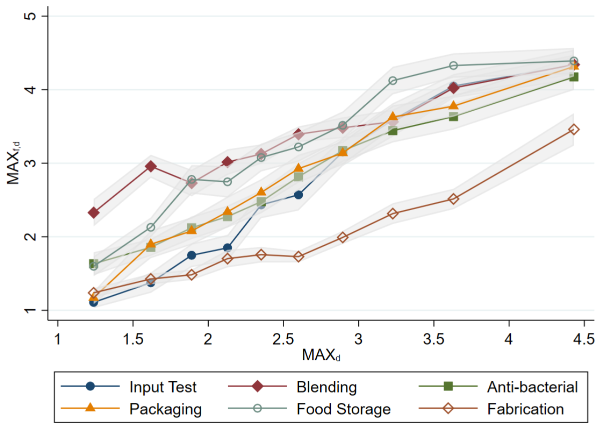
Sectors	Business Administration	Production Planning	Sourcing	Marketing	Sales	Payment	Quality Control	Fabrication
Crops	1.12	1.07	0.94	0.92	0.85	0.99	0.87	
Livestock	1.68	1.33	1.01	0.92	0.65	0.63	1.03	
Food Processing	1.40	1.25	1.03	0.91	0.87	0.80	1.02	0.74
Apparel	1.72	1.58	1.03	0.78	0.83	0.71	0.76	0.58
Leather	2.02	1.40	0.64	0.97	-0.06	-0.04	0.74	0.75
Car	1.00	1.13	0.97	1.51	0.89	0.40	0.99	0.57
Pharma	0.92	0.92	0.96	1.22	0.80	0.15	0.62	0.72
Other Manufacturing	1.49	1.36	1.08	0.84	0.92	0.65	0.89	0.62
Wholesale/Retail	1.58	1.50	1.08	0.97	0.66	0.53	1.10	
Financial Services	1.28	1.39	1.08	1.04	0.86	0.45	0.99	
Land Transportation	1.13	1.19	0.84	0.77	0.63	0.59	0.76	
Health	1.60	1.56	1.06	1.15	0.53	0.73	0.78	
Other Services	1.38	1.34	1.04	0.88	0.74	0.51	1.03	



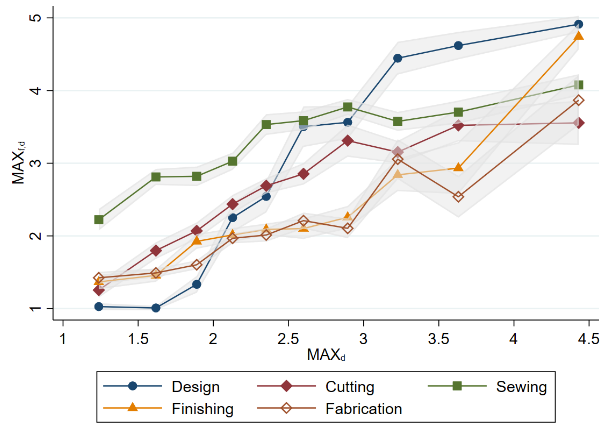
(a) General Business Functions



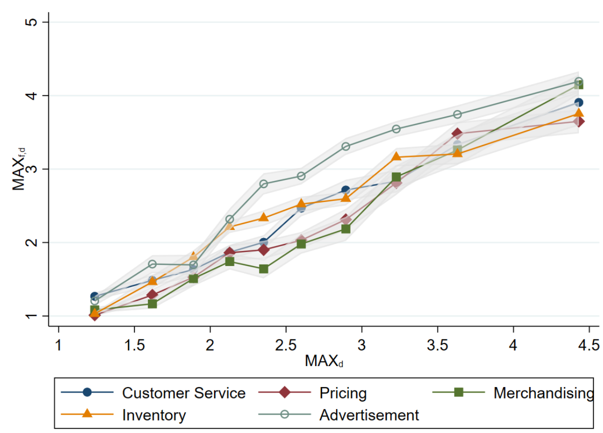
(b) Agriculture (crops)



(c) Food processing



(d) Waring apparel



(e) Retail

Figure C.14: The Technology Curve, $MAX_{f,j}$ vs. MAX_j by Deciles

Note: Note: $MAX_{f,d}$ (vertical axis) is the average value of $MAX_{f,j}$ for the establishments in the d decile of MAX_j. MAX_d (horizontal axis) is the average value of MAX_j for the establishments in the d decile of MAX_j. All averages computed using establishment sampling weights.

D Detailed acknowledgments

The preparation of the Firm-level Adoption of Technology survey questionnaire involved the contribution of several sector experts within and outside the World Bank.

First, we would like to thank the following World Bank Group colleagues: Erick C.M. Fernandes (Lead Agriculture Specialist), Holger A. Kray (Manager, Agriculture), Michael Morris (Lead Agriculture Economist), Madhur Gautam (Lead Agriculture Economist), Robert Townsend (Lead Agriculture Economist), Ashesh Prasann (Agriculture Economist), Aparajita Goyal (Senior Economist, specialized in Agriculture), Harish Natarajan (Lead Financial Sector Specialist), Erik Feyen (Lead Financial Sector Economist), Laurent Gonnet (Lead Financial Sector Specialist), Arturo Ardila Gomez (Lead Transport Economist), Victor A Aragones (Senior Transport Economist), Edson Correia Araujo (Senior Health Specialist), Irina A. Nikolic (Senior Health Specialist), Brendan Michael Dack (Chief Industry Specialist at IFC), Emiliano Duch (Lead Private Sector Specialist), Blair Edward Lapres (Economist), Kazimir Luka Bacic (Operations officer), Justin Hill (Senior Private Sector Specialist), Etienne Raffi Kechichian (Senior Private Sector Specialist), Justin Yap (Private Sector Specialist), Austin Kilroy (Senior Economist).

Similarly, we would like to thank several external experts. From Embrapa-Brazil, we would like to thank Alexandre Costa Varella (Head of Embrapa research unit in the South region, expert on livestock), Flávio Dessaune Tardin (Researcher, Maize and Sorghum), Edison Ulisses Ramos Junior (Researcher, Soybean), Isabela Volpi Furtini (Researcher, Rice and Beans), Alberto Duarte Vilarinhos (Researcher, Cassava and Fruits), Carlos Estevão Leite Cardoso (Researcher, Cassava and Fruits, Technology transfer), and other participants of the internal seminars to validate the sector-specific questionnaire for agriculture and livestock. We also would like to thank a group of senior consultants from Patina Solutions, who contributed with the preparation and validation exercise for several sectors, including Daren Samuels (Consultant, Manufacturing), Christ Baughman (Consultant, Food Processing), Sandra Aris (Consultant, Wearing Apparel and Lather), Steve Zebovitz (Consultant, Pharmaceutical sector), Shelly Wolfram (Consultant, Retail and Wholesale), James M. Keding (Consultant, Transport and Logistics), as well as Justin Barners (Consultant, Automotive). Finally, we would like to thank Sudha Jayaraman (Associate Professor, Department of Surgery, Virginia Commonwealth University), Christina Kozycki (Infectious Disease Fellow, NIH), Jonathan Skinner (Health Economist, Dartmouth College), and Elizabeth Krebs (Assistant Professor of Emergency Medicine at the Jefferson Center for

Injury Research and Prevention), for their comments and feedback on health services.

We also would like to thank the National Statistics Agencies and local partner institutions that provided access to the sample frame and/or support the implementation of the survey. They include: Statistics Korea, Statistics Poland, National Statistics and Demography of Senegal (ANSD), the Federation of Industries of the State of Ceará (FIEC), and the General Statistics Office of Vietnam (GSO), the Kenya National Bureau of Statistics (KNBS), the Bangladesh Bureau of Statistics, and the Central Statistics Office of India.