Within-firm services and firm capabilities*

Eric Mengus Tomasz Michalski

November 21, 2022

PRELIMINARY

Abstract

We document that more than one third of employment and two-fifths of the wage bill in French manufacturing firms is in jobs performing service functions—i.e., neither production nor management—and that these functions gained importance over the period 1999-2015. We underline the role of the *non-routine* functions that are more prevalent in larger firms. Conditioning on firm size, their higher shares in employment are correlated with innovation, intangible capital, product complexity, higher total factor productivity and profitability. This suggests that firms use non-routine services to generate within-firm knowledge and create firm capabilities. Consistently with empirical regularities, we model firms as organizations where production of higher-value added, complex goods requires within-firm knowledge workers to develop capabilities.

JEL Classification:

Keywords: firm organization, functions, productivity, knowledge generation, capabilities.

*Mengus: HEC Paris and CEPR. Email: mengus@hec.fr. Michalski: HEC Paris and GREGHEC. Email: michalski@hec.fr. We thank Charles Angelucci, Jaime Arellano Bover, Laura Boudreau, Giacomo Calzolari, Gregory Corcos, Jose de Sousa, Luigi Guiso, Johan Hombert, Raphaël Lévy, Claudio Michelacci, Denisa Mindruta, Ferdinando Monte, Tomasz Obloj, Luc Paugam, Franco Peracchi, Andrea Prat, Andrea Pozzi, Fabiano Schivardi, Carlos Serrano, Claudia Steinwender, Daniele Terlizzesse, David Thesmar, Laura Veldkamp, Eric Verhoogen, Birger Wernerfelt, Liangjie Wu, Riccardo Zago, Joseph Zeira, Luigi Zingales, Yanos Zylberberg and seminar participants at the 4th Joint Research Conference on "Firm organization, firm financing and firm dynamics" (Rome), EIEF, LMU, MIT Sloan, Paris Saclay and Université de Lille for helpful comments. All remaining errors are ours.

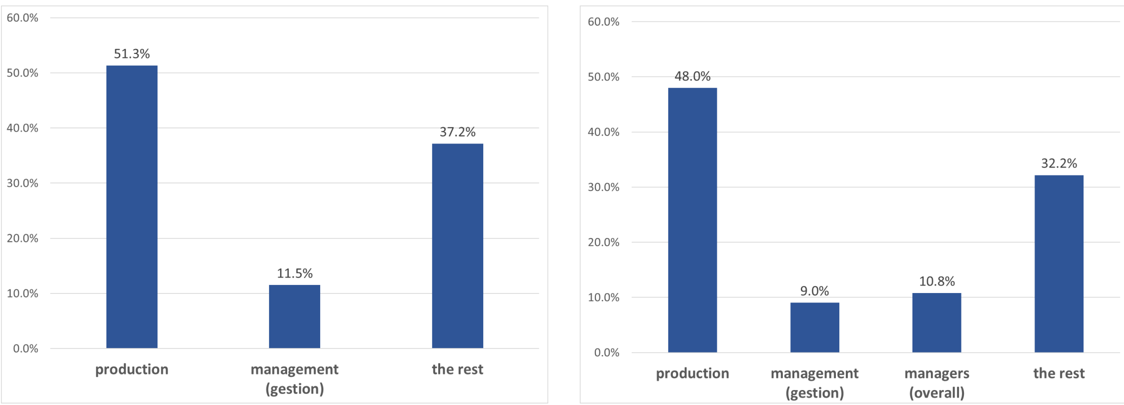
1 Introduction

[...] Just what, then, are the functions of the executives responsible for the fortunes of the enterprise? They coordinate, appraise and plan. They may, at the same time, do the actual buying, selling, advertising, accounting, manufacturing, engineering, or research, but in the modern enterprise the execution or carrying out of these functions is usually left to such employees as salesmen, buyers, production supervisors and foremen, technicians, and designers. [...]

Chandler (1962), p.8.

Using an official classification of jobs mapping into functions they fulfill in organizations, we identify that 37.2% of hours worked and 41.6% of wages paid *within* manufacturing firms in France in 2015 (see the left-hand panel of Figure 1) are in occupations performing functions other than production or management (administration).[1] This constitutes a large fraction of total hours worked and the total wage bill; hence the non-production and non-management functions must be important to successful firm operations. The share of these jobs is also growing. Hours worked and wages paid between 1999-2015 increased in our manufacturing sample by 7pp and 5.7pp respectively, predominantly at the expense of production-level jobs.

Figure 1: Shares of workers in different functions.



The left-hand panel shows the share of hours worked in production, management ("gestion" or administration in French) functions and the rest while the right-hand panel gives these shares and after isolating manager positions. Sample: firms in manufacturing with employment >50 workers in 2015.

To date, however, the connection between these jobs in non-management/non-production functions and the working of firms is still unexplored. In fact, the role of such jobs remains largely absent from economists' theories of the firm, even if they are mentioned in heuristic discussions or are invoked as

[1] The categorization of jobs into functions (see Table 1) is done by the INSEE, the national French statistical agency. Data for a sample of manufacturing firms with >50 employees. As the right-hand panel of Figure 1 shows, excluding jobs within service functions that have managerial tasks reduces the fraction of hours worked to 32.2% of hours. Data on such within-firm occupation details are not available for the U.S.

an explanation of empirical patterns (Atalay et al., 2014, —nonproduction workers in their language). As envisioned by Chandler (1962) in the quote above, we argue that these non-production and non-management functions escape the traditional hierarchical information processing (Radner, 1993), problem solving (Garicano, 2000) or vertical (i.e., concerned with incentives) views of the firm but are essential for carrying out vital firm functions. Also, as these functions (e.g., B-to-B, Intellectual services, Logistics, Maintenance, etc) provide services within the firm, we shall call them *service* functions.

In this paper, we present evidence and theory on French manufacturing firms that an important fraction of jobs dedicated to service functions allow to generate within-firm knowledge which allow firms to develop capabilities. By capabilities within a well defined market we mean firm productivity (the ability to produce cost-efficiently) and the complexity of goods offered (the ability to craft product characteristics to impact demand), both linked with profitability, consistently with Sutton (2012). Our main empirical finding is a tight connection between jobs in service functions, especially the *non-routine* ones, with measures of capability and knowledge generation. Importantly, we find that while R&D plays an important role within these service functions, other service functions are also critical, in particular B-2-B (purchases or sales) or Intellectual Services (legal services, IT, marketing among others). We rationalize these findings in a model with specialized labor necessary to generate knowledge in order to develop firm capabilities. Overall, our findings point to the critical role of service functions (along management) in knowledge generation within the firm and we discuss their implications for the borders of the firm and its market power.

To start with, we uncover new facts about functions in organizations using a sample of all French manufacturing firms above 50 employees in 2015. We focus on manufacturing because it offers a rather homogenous sector where the output is strongly linked to physical goods—hence it is easier to distinguish the "production" types of occupations from those performing service and management (administration) functions. To examine functions, we use the detailed classification ("PCS"), introduced by the French statistical office (INSEE), that allocates 486 different occupations exhaustively into 15 distinct functions.

First, we show that within-firm service functions correspond to an important share of employment—as mentioned, 37.2% of hours worked and 41.6% of wages paid. Furthermore, these functions are present in almost all firms. This holds particularly true for business relations (B-2-B), R&D, maintenance, transport and logistics and, to some extent, Intellectual Services (marketing, consulting, IT, legal services etc.). In addition, these jobs gained importance over time as their share increased within firms in the aggregate hours worked and wages paid between 1999-2015 by 7pp and 5.7pp respectively while that of production fell by similar magnitudes—consistently with overall trends such as offshoring, automation or outsourcing—and that of management remained constant.

Second, we show that the distribution of these service functions is not uniform across firms. The bulk of the corresponding heterogeneity pertains to a subset of these within-firm service functions: the ones that embody *non-routine* tasks according to the routineness measure from Autor et al. (2003). These functions are Business-to-Business purchases and sales (B-2-B); Research and Development (R&D) and Intellectual Services (other cognitive-intensive tasks such as lawyers, see Section 3.1). We show that larger firms are relatively more intensive in these non-routine service functions. The share of employment in non-routine services functions increases from an average of 12% for the smallest firms of our sample to more than 30% for the largest ones. In contrast, the shares of hours in routine service functions such as logistics or other functions (such as retail sales) are uncorrelated with size.

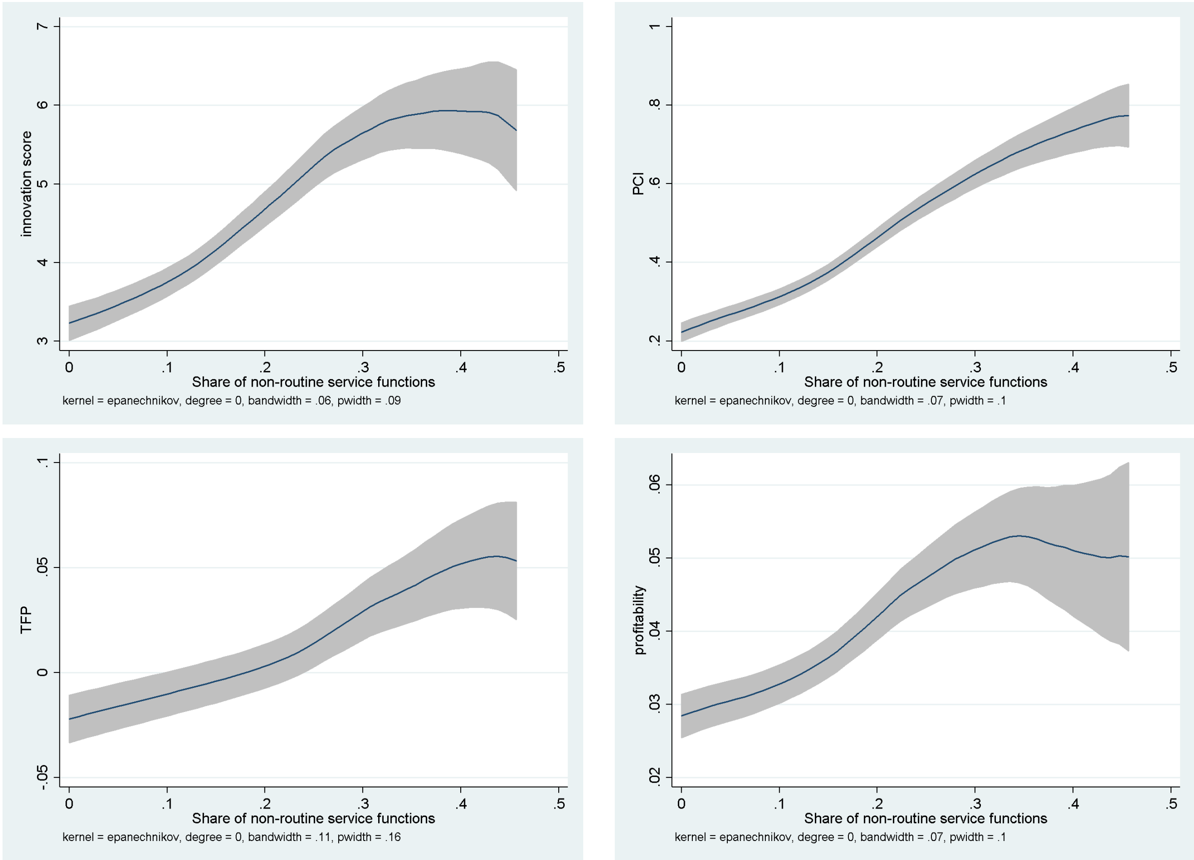
Third, we provide evidence suggesting that non-routine service functions are critical for building up capabilities. As Figure 2 illustrates, we show that firms with more of these functions also 1) generate more knowledge, 2) produce more complex goods, 3) have higher profitability,[2] and 4) are more productive. In particular, we show that higher shares of non-routine service workers are associated with (indirect) measures of knowledge production such as higher levels of self-reported product, process, marketing and logistics innovation and intellectual property protection but also more intangible capital (per hour worked or as a share of total capital). The share of non-routine service functions is also positively correlated with measures of product scope, complexity, higher sales volatility, product quality and firm profitability. We show that this connection holds true not only in the cross-section but also in the time-series, as rises in the firm-level share of non-routine service functions predict future improvements in the characteristics of the goods produced by firms. Finally, this share is also positively correlated with firm's TFP measures, even controlling for the share of the management/administration (as a function) and the share of managers overall and, also, on top of the (projected) skill composition of jobs.

Finally, as we noticed above, the role of non-routine functions include but does not limit to R&D. In particular, we obtain that all other non-routine functions such as B-2-B or Intellectual Services are positively correlated with product innovation measures, or e.g., marketing innovation, or measures of intangible capital, even when controlling for R&D. Zooming in further, the share of workforce active in input purchases or marketing are strongly correlated with firm productivity. The former is also strongly correlated with a higher share of outside purchases of goods and services in total operational costs.

To interpret these facts, we build a simple model of service functions and knowledge generation. Firms with heterogenous productivity levels optimize over the complexity (measured by the number of payoff-relevant states) of the goods they produce, their labor composition and the knowledge that they generate. To be profitable, the production of more complex goods requires knowledge generation. More specifically, in our model, firms observe their productivity draws and select whether to produce either a simple but low-value good or a high-value but complex good using a set of inputs that they source using labor. The production of the complex good requires the firm to select one specific input, in contrast with

[2]As documented by De Loecker et al. (2020), profitability of a firm is also related to its markup.

Figure 2: Shares of non-routine service workers, capabilities and outcomes.



The panels show from top clockwise the firm-level relation between the share of non-routine service workers employment and an innovation index, product complexity scores (PCI), total factor productivity (TFP), and profitability. Sample: firms in manufacturing with employment >50 workers in 2015 trimmed at top and bottom 2.5% in terms of measured non-routine service workers share.

the simple good which can be produced with any. However, firms do not know ex ante which input to select but they can make a more informed choice by acquiring this information. In our baseline model, firms employ specific labor internally to generate such knowledge.[3]

From this model, we derive empirical implications that are consistent with data. More productive firms, relative to less productive ones 1) have a higher share of labor specialized in knowledge generation, 2) generate more knowledge, 3) produce more complex goods, 4) are more profitable, and, to the extent that management is complement to knowledge generation, 5) have a higher share of labor in management. Our model, through various interpretations, indicates that B-2-B, Intellectual services and R&D functions should be especially important for within-firm knowledge generation and the building of capabilities.

Our findings that within-firm services, especially the non-routine ones, are crucial to building up firm capabilities have multiple implications.

[3]In an extension, we relax this assumption by endogeneizing the choice of firms of information production internally or externally through a firm specialized in knowledge generation. The tradeoff for the manufacturing firm is cheaper access to information vs. a more differentiated good if knowledge is produced internally. Notice also that our framework can be restated e.g. as firms searching for an "ideal" consumer variety or quality when they produce a complex good.

First, despite their roles in decision making, coordinating and incentivizing workers or information processing, management and CEO practices (Bandiera et al., 2020; Bloom et al., 2014; Dessein and Prat, 2022; Giorcelli, 2019) or hierarchies (Garicano, 2000; Caliendo et al., 2015, 2020), by themselves are not enough to create firm capabilities allowing to make products with attractive, complex characteristics at a low cost that permit to obtain high profitability—using a classical music example, a talented orchestra conductor needs gifted, specialized musicians as well to create a magnificent rendering of the *Eroica* symphony. This holds especially true in sectors like manufacturing where the set of tasks required to produce complex products requires expertise in multiple domains ranging from R&D through input procurement to marketing and advertising. To sum up, instead of a vertical/hierarchical view of knowledge generation, our results suggest the relevance of a functional view as summarized by Figure 3.

Second, it also implies that in-house production is not central for capabilities, in the spirit of "factoryless" firms as described by Bernard and Fort (2015). This is also in line with the idea that much of the labor force in multi-establishment firms is about creation and transfers of intangibles and not about the transfers of goods as emphasized by Atalay et al. (2014).

Third, the increasing importance of within-firms services that we document parallels the structural transformation of labor markets and complements the emergence of the sector of business services. In addition, with the exception of R&D, all non-routine service functions wages are parts of SG&A costs, which according to De Loecker et al. (2020), are critical drivers of the recent increase in markups: these non-routine service functions and their role to expand the firm's capabilities are then natural candidates to explain the rise of the most able firms' market power. We leave, however, the development of these observations to future research.

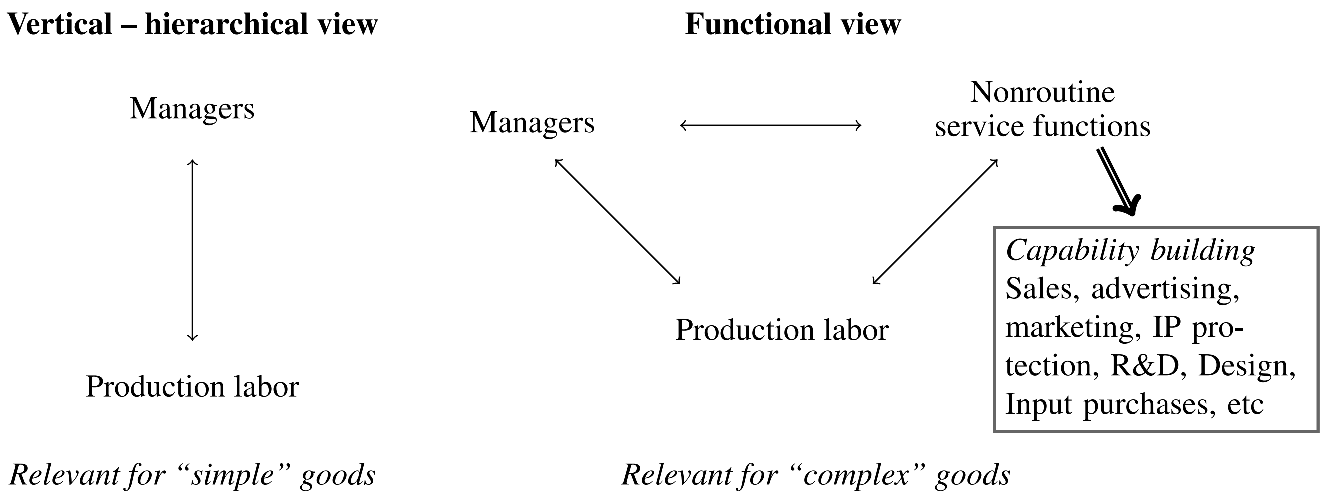


Figure 3: Relationship between different labor functions within firms and the complexity of products.

Related literature. Our paper is connected to different strands of the literature.

First, the role of service functions to produce knowledge within firms is connected to the resource-view of the firm (see Wernerfelt, 1984; Lockett et al., 2009, for a review), the dynamic capabilities theory of the firm (Teece, 1982; Teece et al., 1997) or the knowledge-based theory of the firm (Kogut and Zander, 1992). We share with the resource-view of the firm the approach that firms' capabilities cannot simply be bought 'off-the-shelf' on a market but have to be built up. Within this literature, we are closer to Grant (1996) – where the firm allows workers to coordinate the application of their knowledge – and Rajan and Zingales (1998) – who formulate a theory of the firm where workers make specific investment to work with the firm's critical resources. In particular, as in Grant (1996), our findings emphasize the key role of knowledge as a resource for the firm and the tight connection between knowledge and individuals[4] and, as in (Rajan and Zingales, 1998, 2001), some of the firm's workers are making specific investments, here in generating firm-specific knowledge.[5] In addition, the view that producing and sharing knowledge is key for the firms' borders is also put forth by Atalay et al. (2014). We contribute to this literature by providing evidence on the type of services (intangibles) that firms may produce and their relationship to different capabilities. Finally, our distinction between production, management and service functions is related to Porter (1985)'s value chain, that distinguishes between primary and support activities.[6] Although our approach builds on a different split of activities within the firms—based on the functions occupied by workers—, we share some of his insights, for example that capabilities, in his words the competitive advantage at the firm level, may result from all the components of the value chain.

In this regard, notice that knowledge generation does not necessarily require the acquisition of any firm-specific human capital as in Becker (1962) or Lazear (2009) but may well only require general human capital. In contrast, we concentrate on knowledge generation in the firm that allows the firm to create its competitive advantage as long as the created knowledge remains within the firm.[7] For example, AI modelers could be used by distinct firms to uncover knowledge needed in different areas: on exchange rate market fluctuations, programming robot arm movement, or optimizing supply chains. From this perspective, our paper is closer to Tambe et al. (2020)—who document the role of IT labor in accumulating digital capital—which is in turn a key determinant in future firm productivity.[8] Atkin et al. (2015) show correlations that marketing effort in establishing trade links may be important for generating higher markups, while more recently Patault and Lenoir (2021) investigate the role of sales managers for exporting. We add to these papers a systematic and exhaustive analysis of service functions from

[4]As Grant (1996) quotes Simon (1991): 'All learning takes place inside individual human heads; an organization learns in only two ways: (a) by the learning of its members, or (b) by ingesting new members who have knowledge the organization didn't previously have' (Simon, 1991: 125).

[5]As in Rajan and Zingales (1998), we do not explain the presence of service functions through the lens of the property-rights theory (Grossman and Hart, 1986; Hart and Moore, 1990), given that we are not concerned by the allocation of assets but of workers, on which the firm cannot have residual rights.

[6]Support activities include management and non-routine service functions as R&D and B-to-B. Primary activities include production but also service functions as logistics, marketing and sales.

[7]In particular, workers with the same type of general human capital could be used by different firms to generate disparate firm-specific knowledge required by the firm to produce complex goods. Workers obtain firm-specific knowledge that is potentially transferable to other firms if they leave so they become valuable to the company.

[8]A similar argument on the role of IT professionals on firm productivity is made also by Harrigan et al. (2021) for the French case.

administrative data and show the relative importance of all of these functions engaging in knowledge generation and its role for the production of complex and higher-markup goods.

The trade literature investigated capabilities at the country (Sutton and Trefler, 2016; Atkin et al., 2021) or the firm level (Bastos et al., 2018). As in this literature, capabilities are about the set and quality of products that can be produced and with what productivity. We link specific firm functions to capabilities at the same time measuring their labor content. We also share some of the complexity good measures that we take to qualify the difficulty to produce some goods.[9]

Our interpretation that a large fraction of service functions are about generating knowledge is related to the large literature about firm decision-making under different forms of uncertainty (e.g. Jovanovic, 1982; Zeira, 1987). The need for uncertainty reduction is also one of the interpretations of hierarchical firm organization theories as Radner (1993) or Garicano (2000).[10] In contrast, we study knowledge production not within a firm hierarchy involving different layers of management (that constitute knowledge hierarchies, solve team production or information processing problems) but through the non-routine service functions workers generating knowledge. We also show that these functions are able to produce multiple and potentially complementary capabilities such as different types of innovation. Our approach that knowledge generation is costly and has to be traded off with potential gains has also received specific attention by Dessein et al. (2016).

Finally, our approach on knowledge generation is connected to the literature on costly information acquisition following Townsend (1979). The use Shannon's entropy is connected with the literature on rational inattention as initiated by Sims (2003) (see Mackowiak et al., 2020, for an overview). In addition, the idea that outsourcing knowledge generation contributes to exacerbate competition by the firm is related to the literature on the strategic complementarities in information acquisition as, for example, Veldkamp (2006).

2 Data description and functions

In this section, we present our sources of data and how functions in firms are measured.

[9]Our paper is also connected with the literature on international trade and intra-firm trade with prominent examples as Helpman (1984) or Antràs (2003). This literature is particularly interested in the separation between headquarter services ("general purpose inputs") vs. production, and the fragmentation of production processes, also enabled by ICT (Fort, 2016). We aim to understand better such "headquarter" services themselves, as certain functions performed within firms – such as management (esp. by high-skilled workers)—received particular scrutiny in contemporary literature but the nature of the general purpose inputs produced within firms has so far received much less attention. Some attempts to study these phenomena were made by Bernard and Fort (2015) focusing especially on input outsourcing and imports and by Defever (2006) in the context of multinationals. Distinctively, our argument is not about the benefits of increased fragmentation of production allowing for specialization but rather about the importance of keeping particular non-routine service functions within-firms—the internal functional composition.

[10]Discussed e.g. by Garicano and Van Zandt (2013). See also Garicano and Wu (2012) for an overview and a connection to the strategy literature.

2.1 Data sources and sample

Sources. We rely on two main sources of data. We first use the French matched employer-employee data (DADS – "Déclarations Annuelles de Données Sociales") that gives worker-level information such as occupation, wage, hours for the current and the preceding year of each vintage of data. Occupations are coded following the 2003 PCS French classification at the 4-digit level.[11] Second, we use the FARE data set which is built from mandatory income statements of firms to tax authorities.From this database, we extract firm-level information such as capital, output, sales and value added or the stock of intangibles. As the base sample, we use the 2015 vintage of DADS-Postes (to obtain measures of compensation etc.) and FARE.

To generate additional measures on the complexity of production of firms we use the EAP data sets of the INSEE that track detailed quantity and value of firm sales at the product level. This allows us to link the characteristics of each firm's manufactures to other data sets such as the Eurostat's PRODCOM database (to obtain product-level sales growth volatility), the Product Complexity Index (PCI), or the Rauch (1999) classification. We also use the Community Innovation Survey for the study of firm-level innovation.

A detailed description of data sources and variables used is given in Appendix B.

Sample. We identify a firm as a legal entity with a unique SIREN number.[12] For our main results, we retain firms in the manufacturing sector (sectors in Section C according to the French NAF, rev. 2 classification), with employment in terms of hours from DADS-Postes (with full time equivalent of 1680 hours worked/year x 50) and FARE above 50 employees. One reason to apply such a threshold is to keep only the largest firms that may have a diversified workforce and hence multiple functions within firms. Another is that, given additional legal requirements on firms with more than 50 employees in France, there might be a discontinuity in productivity among firms while passing this threshold (see Garicano et al., 2016), confounding the analysis.[13] In the end, we are left with a sample of 6,715 firms. The sample statistics are shown in Table 2. The median firm has an employment of 126 full-time-equivalent (1680 hours/year) positions. We use 2015 as the base vintage but also use 1999, 2005 or 2010 for comparison and to analyze long-run changes.

[11] See Caliendo et al. (2015) among others for a use of this classification to study firm organization.

[12] We also repeated our analysis using a sample where some of these legal entities are consolidated at a group level ("enterprises profilées" in French) obtaining qualitatively and quantitatively similar results. Available upon request.

[13] We repeated our analysis with different thresholds at 10 or 25 employees, and also for 2010 with qualitatively similar results.

2.2 Functions

We study firms' different employment structures in terms of the functions that workers perform in the enterprise. To this purpose, we rely on the mapping of occupations into functions done by the French statistical agency, the INSEE. Based on the French PCS classification of occupations, this classification was developed since 1982 to study the different tasks conducted within firms from all sectors and which may involve different level of skills. More precisely, it allocates exhaustively 486 existing 4-digit occupation codes into 15 distinct functions such as e.g., production, management, transport and logistics, business-to-business sales and purchases. We provide a more detailed description of functions in Table 1 as well as examples of jobs corresponding to each of these functions.[14]

Such defined functions are transverse and do not overlap with industries: a research engineer can work within the same function (R&D) either in aircraft manufacturing or aluminum producing firms. In addition, they are not tied either to jobs' specific contractual terms (independent contractor, public or private entity, temporary or permanent employment). Importantly, they may combine very different levels of skills and distinct jobs focused on a particular function. For example, the function of "production" bundles together directly involved engineers (typically with college education), technicians (e.g., foremen, that might have some college and/or technical education) and skilled or unskilled blue-collar assembly line workers. Functions also cut through hierarchies. For example, "management" combines CEOs, managers of different levels, assistants, secretaries and regular office workers that perform tasks of managing the firm.[15]

3 The importance of service functions within firms

In this section, we document facts about the use of service functions within firms. We first show that the service functions are an important share of firms' employment and some service functions are close to ubiquitous. Second, service functions are heterogeneous across firms, with larger firms being more intensive in *non-routine* service functions. Third, we show that non-routine service functions gained in terms of importance over the period 1999-2015.

3.1 Service functions are an important component of firms

We start by investigating the importance of service functions for manufacturing firms as a whole.

[14] Some of these functions such as agriculture and fishing; health and social work or public administration will be less prevalent in the set of firms that we consider.

[15] Firms do not report the functions of their workers directly to the INSEE, but file contract-level social security declarations instead. The statistical institute INSEE attributes the functions to particular jobs and permits researchers later on to use this data. Thus, firms do not have an interest to strategically misreport this data to influence e.g. investors, as the data is available with a lag (of 2-3 years) to a limited group of researchers that are bound by secrecy oaths.

In Table 3, we report statistics on the distribution of functions across firms in terms of hours worked. The two main functions, typically considered in the literature, are clearly "management" and "production" that jointly account for more than 62.8% of hours worked in our sample and are present in more than 99% of the manufacturing firms studied. However, this also means that 37.2% of hours worked are unaccounted for by these base functions. Moreover, as shown in the last column, 41.6% of wages paid accrue to jobs that have neither non-management and non-production functions.

A close inspection of Table 3 reveals that some functions are close to ubiquitous, present in more than 80% of firms and account individually for more than 5% of total hours worked in manufacturing. This set of functions gathers handling business to business relations (B-2-B), R&D, maintenance and transport and logistics. We can add to this set "Intellectual Services" that are present in more than the majority of firms but correspond to a smaller share in total hours worked (2.4% of total). These functions group occupations that allow carrying tasks in the firm at different production stages, that could be associated as being transverse throughout the firm. They escape the traditional vertical management-production plant or hierarchical dichotomies. For example, Chandler (1962), p.8, dissociates "administration" functions (that would correspond to "management" in our classification) from those of "buying, selling, advertising, accounting, manufacturing, engineering, or research [...]" which would be encompassed by B-2-B, intellectual services, production and R&D functions. We shall call them thus service functions.

Consider the following examples of such service functions. B-2-B involves purchases of inputs (and managing effectively outsourcing) but also sales to other businesses. Maintenance involves functions such as servicing equipment and buildings, cleaning premises or the treatment of pollution. Transport and logistics involves warehousing and the movement of inputs, final goods or people—also within the firm. Intellectual services comprise of lawyers, marketing or IT professionals and different consultants.

Some other functions like construction and public works, culture and leisure, retail, health (e.g., company doctors) and social work or local services (e.g., cooks) are much less present in firms. Finally, public administration and agriculture and fishing are almost absent from our sample of manufacturing firms. Together they account only for 3.3% of total hours worked. We will group all these additional functions as "other" and disregard them in our analysis given their diverse nature.

Despite the fact that many service functions are present in a majority of firms, their employment shares within organizations differ greatly, as revealed by the coefficients of variation that are typically larger than 1. We turn next to functional firm heterogeneity.

Hierarchy and functions. Are service functions "management" in disguise? The INSEE classification that we use to discern between different functions or occupations performed by workers in firms does not inform precisely about within-organization hierarchy. It is important thus to inspect whether the service functions are not hiding a large fraction of workers that perform managerial roles.

We proceed in the following way. The DADS data at our disposal does not directly trace hierarchical ties in firms. But the detailed description of jobs by 4-digit occupation codes from the PCS along with the most typical job titles allow to approximate the share of managers in a given occupation. For each 4-digit occupation we code an index that is an equally weighted measure of two subindexes. The first is an indicator whether "manager" or similar role is mentioned in the occupation title—as an example, with this approach, foremen are coded as managers. The second is based on whether all (=1) or some (=0.5) of the typical jobs within the occupation have explicitly managerial title roles or not. This index takes thus values of $\{0, 0.25, 0.5, 075, 1\}$, and we apply it directly to the share of the workforce in a given occupation to approximate the share of managers.

In Table 3, lowest panel, we show the resulting shares of service and other functions within firms without managers. Even after our adjustment, non-manager and non-production workers constitute 32.2% of the hours worked in firms. Given with the share of 37.2% of hours in service functions, we conclude that a high fraction of hours worked in service and other functions is not related to the preponderance of managers. Employees performing non-management and non-production functions in firms are not merely "relabeled" managers dedicated to solving non-administration and non-production problems.

3.2 Service function heterogeneity

In this section, we first identify that service functions markedly differ in terms of routineness and, on top of management and production, we identify two groups of service functions that, accordingly, we label as routine and non-routine. Second, we show that larger firms are more intensive in non-routine service functions, while they have lower shares both in production and routine service functions and the share of management is relatively constant across firms of different sizes.

Routine and non-routine service functions. From the perspective of organization economics one important measure is that of routineness, that is the extent to which a given function involves tasks that are repetitive, standardized and can follow codified procedures (see e.g. Costinot et al. (2011) for such a link). Employees executing routine duties are much easier to manage, monitor and appraise. Conversely, the nature of non-routine tasks *par excellence* requires employees to have more own initiative, deal with non-standard problems, generate knowledge that may be specific to the unique issue at hand.[16]

To provide insight into the nature of functions, we turn to readily available measures of routineness—the Routine Task Intensity (RTI) index of Autor et al. (2003) that classify occupations according to the ease of their automation and transpose these measures into functions.[17]

[16] We do not address any potential incentive provision issues in this paper because of lack of suitable data.
[17] More details in the Online Appendix.

Functions differ markedly in such a measure of routineness (Table 4). Those requiring higher skills—such as Intellectual Services, R&D or B-2-B tasks—are typically much less routine with average RTI scores around -0.6. Management and production are the most routine among all functions with the former on average the most routine of them all. This may be surprising at first, but the bulk of hours worked in that function is performed by office workers (2-digit occupation code CS 54), among the most routine occupations. This is because the "management" function doesn't capture the hierarchical share of hours worked by managers in general, but the share of hours dedicated to administration tasks within each firm. Most of these are performed by workers that are at the bottom of the firm hierarchy.

Table 4 suggests a partition of the different functions considered into 4 major distinct categories. First, we want to treat separately management and production which are traditionally discussed in the literature. There is also a difference among the service functions among those that are more and less routine. The more non-routine service functions—intellectual services, R&D or B-2-B require experimentation and non-standard worker actions. We shall pay particular attention to these functions and label them *non-routine service functions*. We call the maintenance and transport and logistics as *routine service functions*. The summary statistics for such groupings of functions is shown in Table 3, lower panel.

Functions across firms. We now investigate how management, production, routine and non-routine functions are distributed as a function of firm size.

Figure 6 shows the extensive margins of each category (dichotomous variables whether they are present in firms or not) with firm size measured by hours worked.[18] It is clear from Figure 6 that the presence of different categories of functions is related with size: large firms have all types of functions. The shares of these functions within firms of different sizes (their intensity) are shown in Figure 7. In regressions weighted by firm size, there is a negative relationship between firm size and the share of employment in production. The share of management and routine service functions is constant while that of non-routine service functions increasing with firm size.

Given that the share of routine service functions will turn out not to be correlated with knowledge production measures or productivity, and for most firm sizes (e.g., below the 95% percentile their share is constant), we will concentrate in further investigations on non-routine service functions in contrast to management or production functions.

The fact that non-routine functions have lower shares in employment for smaller or less productive firms does not mean that e.g., R&D, purchases of inputs or sales, marketing etc. are not required by those firms in production.

[18] Figure 8, top panel, depicts the relation between the logarithm of the number of all 15 functions present in an organization and firm size.

The number of functions within firms. One explanation why small firms have lower shares of employment of explicitly non-routine functions is that workers in those firms perform many more unrelated tasks on their job that are not captured within the administrative data set. Indeed, e.g., managers in smaller firms may need to perform several functions at the same time that in a larger firm are taken care of by workers dedicated to them specifically, while retaining their base functions (in the words of Chandler) such as coordinating, appraising and planning. We do not see large or productive firms with only managers (which, according to the naive interpretation of our estimates would be the most productivity "enhancing" thing to do) or production workers, but a much more diverse workforce. This suggests complementarities among different functions as in our model. We thus control for firm size in what follows. However, it turns out (discussed further in Section 4.3) the number of functions in organizations is not correlated with their productivity (Figure 8, lower panel). But it is the functional composition that matters.

Firm size and outsourcing. Another explanation of this phenomenon could be that small or less productive firms use outsourcing (or offshoring) of these functions more. In particular, knowledge generation may well take place outside the firm. In favor of this possibility, we do observe outsourcing by firms not only in low- but also high-skilled (non-core in their terminology, i.e. with high codifiability and low weight in firm production) tasks as documented by Bergeaud et al. (2021) among others. However, in our data we also observe that outsourcing intensity[19] is in general positively correlated with measures of firm's size and TFP. This is shown in Online Appendix Figure C.3 for unconditional patterns and Table C.5 controlling for industry fixed effects and employment.

It is also possible that higher shares of non-routine service functions can simply reflect higher intensity of outsourcing or offshoring of production. This is clearly the case with the employment shares of some non-routine service categories such as B-2-B input purchasing—see last column of Table C.5. However, the correlations of production labor or non-routine service workers with outsourcing intensity measures in general do not yield to such an interpretation (see lower panels of Online Appendix Figure C.3 or Table C.5). The relationship between the shares of both types of labor within firms is hump-shaped, indicating more complex relationships at play, beyond the scope of this paper.

3.3 Time evolution of functions within manufacturing firms

Figure 5 presents the time evolution of functions within firms between 1999 and 2015.[20] Manufacturing firms changed considerably their workforce over this time period. In the sample of firms with more than 50 employees, the decline in the hours and wage shares of production workers was 6.8pp

[19]Unfortunately, we do not have a finer breakdown of exact firm input purchases by category nor function.

[20]Firms started reporting 4-digit PCS occupational hours that is the basis for constructing functions in the end of 1990s, but with imperfect coverage. 1999 is the first year where a large number of firms over 50 employees fully report their hours at the 4-digit level that we use in the tables here.

and 4.3pp respectively, implying relative contractions of 11.7% and 8.8% respectively. This was largely substituted by the increase in both the hours and the wage shares of non-routine services both by 5.4pp. Notably, the share of management function hours in total employment barely changed while, in terms of compensation, it fell by 1.5pp.

These changes mirror some of the trends in the overall economy as offshoring, automation of production or the development of market-based business services. Production and the more routine service functions can be more easily outsourced e.g., because of their standardized nature (see Bergeaud et al., 2021, among others). However, as we document, the increases in hours worked were tilted towards non-routine service functions—and not, for instance, management.

4 Knowledge generation and capabilities

The regularities we document raise the question why firms choose different functional employment shares. In this section, we provide evidence in the cross-section and in the time series of the connection between the use of service functions within firms—in particular non-routine service functions—and measures of firms' knowledge generation, capabilities and performance.

We first provide correlations between non-service functions and different measures capturing knowledge generation within the firms, as innovation (in product, marketing, processes, etc) and intangibles. Importantly, the role of non-routine service functions is not confined to R&D but concerns other non-routine service functions such as B-to-B or Intellectual Services.

Next, following Sutton (2012)'s definition of capabilities we document how the use of non-routine service functions is correlated in the cross-section of firms with measures of product characteristics such as complexity and quality. Using a long-difference regression, we also investigate which capabilities change with the increase in shares of the non-routine service functions.

Finally, we provide correlations in the cross section and in the time series of the shares of non-routine service functions with productivity and profitability that are the outcomes of firms being able to increase capabilities.

In the end, our findings can be summarized by Figure 4: non-routine service functions are critical to generate knowledge. This knowledge is at the core of the buildup of capabilities, which, in turn, are crucial to ensure firm's performance. We also map in this Figure the different observables that we are using to concepts to which we refer.

Non-routine service functions $\longrightarrow$ Knowledge $\longrightarrow$ Capabilities $\longrightarrow$ Firm performance

Observables			
	product, process and organizational innovations, intangible capital	production and product complexity, scope, quality, cost efficiency	profitability, high productivity measures

Figure 4: Mapping between non-routine services, knowledge creation, capabilities and firm performance.

4.1 Knowledge generation

In this subsection, we connect non-routine service functions with measures of knowledge generation, as measures of innovation activity or measures of intangibles.

Innovation. As classified, non-routine service functions perform non-routine tasks, related to experimentation and different types of knowledge generation—from understanding the needs of clients and suppliers, through technical R&D, design, and marketing product characteristics or targeted legal expertise. This should result in increased innovative activity.

We observe self-reported measures of different types of innovation for a subset of firms that responded to the 2018 Community Innovation Survey conducted by the INSEE. There are overall 15 questions that directly pertain to different types of innovation by firms in the period 2016-2018, grouped thematically: innovations in products and processes and the development of intellectual property. We correlate the share of non-routine service workers in 2015 (preceding the period 2016-2018 about which the CIS survey pertains to) with these measures. The results are reported in Table 5.[21]

We observe that the share of non-routine service workers in firms in 2015 is positively correlated with all innovation activity measures in the period 2016-2018 (overall index counting affirmative answers to types of innovation, product, intellectual property and process innovation in columns 1-4). We also distinguish between the two different types of non-routine functions (R&D and non-R&D) and both types of workers seem to be equally important in generating innovation when measured with composite indexes.[22]

[21] In the Appendix, Table C.4 provides more detailed process innovation measures.

[22] Further scrutiny of process innovation in the Appendix Table C.4 reveals that there exists innovation such as where R&D is not critical: the share of non-routine non-R&D service workers is positively correlated with marketing or logistics innovations, while that of R&D shares is respectively quantitatively weaker or not statistically significant.

Intangible capital production. We take intangibles as a proxy for the importance of knowledge as an input for the firm's production and so, how much knowledge-sensitive it is.[23] We report in Table 6 the outcomes of regressions of different measures of the intensity of intangibles: intangible capital per worker and the share of intangibles in total capital.

Conditional on firm size, the share non-routine service functions is strongly correlated with the different measures of intangible intensity, and so is management. In contrast routine service share is not correlated with level of intangible capital per worker and negatively correlated in measures of the share of intangibles in total capital. Importantly, this correlation is not driven solely by the R&D share. In the second column, non-routine service functions other than R&D are also strongly correlated with intangibles to the same extent as R&D. When considering alternative measures of intangible intensity as in the fourth or fifth columns (higher share of intangibles in PPE + intangible capital or total capital), we obtain that R&D is actually *less* important than other non-routine service functions (statistically significant at 2%) in creating intangible capital.

4.2 Product complexity and quality

In this subsection, we show that, both in the cross-section and in the time-series, firms that are relatively more intensive in non-routine service functions manufacture products that are more complex and of higher quality.

Measures. We capture the complexity of the product portfolio of a firm through five measures. First, we compute at the firm level their sales volatility weighted by shares of product sales at the 8-digit PRODCOM level. Greater uncertainty in product sales should require more specialized knowledge to perform experimentation and learning to cope with the arrival of new market information.[24] Second, we calculate a weighted product complexity measure based on the Product Complexity Index (PCI) from the Harvard Atlas of Economic Complexity.[25] Third, we compute firm-level sales-weighted share of differentiated products according to the Rauch (1999) classification. Differentiated products require inter alia R&D, marketing, advertising and legal services efforts to allow a firm to differentiate its products from other firms. As organized exchanges do not exist for them, they also require searching for and

[23] Intangibles are calculated by the INSEE as the cumulated sum of expenditures such as R&D, patents, brands, goodwill, etc. The exact list is what is registered as expenditures linked to intangible assets (category 20) as listed by the French generally accepted accounting principles or "Plan comptable général". Intangible capital values that we have are possibly imperfectly measuring the true extent of the concept as it retains only special investments within firms from the accounting perspective. However, an alternative measure considered in the literature based on accounting for R&D personnel expenditures would be clearly circular.

[24] In information theory, in dynamic information acquisition, higher volatility of observed signals is associated with faster depreciation of acquired past information. Thus, firms with higher sales volatility might need to constantly reinvest in acquiring new information and creating knowledge.

[25] The PCI is a ranking of product's know-how diversity and complexity based on country characteristics that make them. It captures succinctly the difficulty that a firm may be facing in making a particular product, requiring an adequate firm-level capability.

handling buyers. Fourth, we simply count the product lines (at the 4-digit NACE level) that a firm produces. Handling more distinct products within a firm requires more production of information about inputs, production processes or different markets. Finally, we calculate how concentrated the product sales are: the more dispersed are the sales, the more difficult is the task to bring goods to the market.

We capture quality by computing log unit values at the 8-digit PRODCOM level. We obtain unit values by dividing sales by reported quantities.

Evidence in the cross section. We correlate these different firm-level and firm-product-level characteristics with the share of non-routine service workers, controlling for firms' size (measured by employment). We report the results in Table 7, without (top panel) and with 2-digit NACE fixed effects (lower panel; 8-digit product fixed effects for product quality) to control for industry-level determinants. We uncover that there is substantial between industry-level variation in these measures, with some industries (not shown) scoring considerably higher than others in terms of our complexity proxies. This is especially true for the PCI and product differentiation measures, where 2-digit industry effects can explain even more than 40% of overall variation. Overall, considering cross-sectional within-industry variation in particular, we find that higher shares of non-routine service workers in organizations are correlated with higher product sales volatility, higher product complexity and differentiation, a larger scope of products and a greater dispersion of sales among distinct product lines. Using within-product variation, we observe that firms with higher shares of non-routine service workers also have higher quality products.

Time series evidence. So far, we have focused on cross-sectional evidence. We now look at within-firm patterns in the time series to confirm the connection between non-routine service workers and product complexity at the core of our proposed mechanism.

One of the challenges is that many of such quantities are slow-moving; another is that increases in the share of non-routine service functions may not instantaneously lead to implementable knowledge generation. Therefore we resort to studying long-run changes. In Table 8, we show regressions of changes 2010-2015 in the above-mentioned complexity measures related to overall changes in employment and the share of non-routine workers over the period 2005-2010, controlling for employment levels and the share of non-routine workers in 2005.[26] That is, we run the following regressions:

$$\begin{aligned}\Delta(Y_{i,2015-2010}) = \alpha + \beta_1 \Delta ShNONROUT_{i,2010-2005} + \beta_2 ShNONROUT_{i,2005} + \cdots \\ \cdots + \beta_3 \Delta LN(Hours_{i,2010-2005}) + \beta_4 LN(Hours_{i,2005}) + \epsilon_i\end{aligned} \tag{1}$$

where $ShNONROUT_{i,2010-2005}$ is the change in the share of non-routine service workers between 2005 and 2010; $ShNONROUT_{i,2005}$ is the share of non-routine service workers in 2005; $LN(Hours_{i,2010-2005})$ is the change in hours between 2005 and 2010 and $LN(Hours_{i,2005})$ are the hours worked in 2005. The

[26]We do not have product-level data at the firm level from the EAP data source to be able to study these further back.

coefficient of interest is β_1.

Although the sales volatility and the Rauch differentiation (at the firm level) or product quality (at the product level; for existing products) measures do not seem to increase in the period 2010-2015 as a result of the change in the share of non-routine workers between 2005-2010, other measures do.[27] Firms that increased the share of non-routine service workers in the preceding 5 years increase their product complexity as measured by PCI measure, further their product scope and decrease their product sales concentration. This suggests that the adjustment the firms undertake may be at the extensive margin—adding new products of higher complexity—rather than an improvement in the characteristics of existing products.

4.3 Non-routine service functions and productivity

We have shown that firms with high non-routine service shares in employment generate more knowledge and have higher capabilities. We now investigate whether there is a link between functional composition (the *shares* of different functions in employment) of firms and their productivity. Our main finding is that the share of non-routine service functions is strongly correlated with standard measures of firm productivity, along management.

Approach. To address this question, we proceed in two steps. In a first step, we estimate firm productivity as the residual of the following OLS regression within each 2-digit sector:

$$LN(VA_i) = \alpha + \beta_1 LN(Capital_i) + \beta_2 LN(Hours_i) + \beta_3 LN(PredictedAverageWage_i) + \epsilon_i \quad (2)$$

where *LN(VA)* is the logarithm of value added; *LN(Capital)* is the logarithm of the value of property, plant and equipment (ppe) capital of the firm in 2015; *LN(Hours)* is the logarithm of hours worked and *LN(Predicted Average Wage)* is the logarithm of the ratio of the predicted wage bill and hours worked[28][29] and ϵ is an error term. We use the predicted instead of the actual wage bill to account for worker skill but at the same time avoid the problem of the correlation of a regressor with the error term: as is known in the literature, more productive firms may pay their workers more due to different rent-sharing practices. The first-stage inclusion of the projected wage bill in TFP estimation accounts indirectly for projected employed worker skill.

[27] There are limitations in how we are able to calculate sales volatility at the product level. We need to use 2005-2015 data from the Eurostat with many times the number of years limited especially to the 2010-2015 period. Most French manufacturing firms in sample have Rauch's index equal to the maximal level (=1) which renders the time evolution of this measure less relevant.

[28] To obtain the projected wage bill, we first regress individual remuneration within firms in 2015 on hours worked, age, age squared, sex, 2-digit occupation X function and industry fixed effects in the manufacturing sector. Then we calculate for each firm the predicted wage bill given the characteristics of its employees.

[29] We tried different specifications with different definitions of capital or including the wage bill directly without a material difference in the results.

In a second step, to investigate the correlation of function shares on TFP we then regress obtained productivity estimates on the shares of management and service functions and industry fixed effects. Given that some firms do not have all functions, we cannot use logarithms of hours worked directly and need to use shares. Since shares necessarily sum to 1, we choose "Production" as the base function of a manufacturing company. As a result, the interpretation of the presented results involves how much a share of a function is correlated with TFP relative to the share attributed to Production. We estimate the following regression by OLS:

$$LN(TFP_i) = \alpha + \sum \gamma_j share_j + \sum \zeta_k \theta_k + \varepsilon_i \tag{3}$$

where $share_i$ denotes the employment shares of functions considered (apart from production) and θ_i are 2-digit industry effects, with γ being the coefficients of interest.

The share of non-routine functions increases with firm's productivity. We report the results in Table 9. Our preferred specification is the one in column 2, with the sample trimmed in terms of TFP at 0.5% from above and below to exclude outliers.[30]

Irrespective of the specification, the share of non-routine service functions is higher in more productive firms. In particular a 1 percentage point higher share of non-routine functions is correlated with a 0.21% higher TFP. For the median firm in our sample (126 employees), a shift in hours from production to non-routine service functions by approximately 12.6 jobs is on average equivalent to an increase of productivity of 2.08%.

Consistent with the empirical result of Bender et al. (2018) on the importance of management on TFP,[31], a higher share of the management function in employment is correlated with higher TFP. An increase in the share of management in employment by 1 percentage point at the expense of production is correlated with an increase of productivity of 0.57%. In columns 7-9 version of Table 9 we explicitly include the share of managers (calculated in the way described in Section 3.1) in all functions bar for those overseeing production. We find that our correlations of different function shares with productivity are qualitatively the same, and the share of managers in the workforce is significantly correlated with TFP. Both functional division of the workforce and hierarchy (manager hours) are related to productivity.[32]

We do not find statistically significant results for the effect of changes in non-routine service shares and our TFP measure in the long-differenced regressions as in the product complexity (or profitability, below) regressions.

[30]Unconditional relationships between the major function categories' employment shares and TFP are also shown graphically in Figure C.1.

[31]These authors particularly focus on management quality. We do not have similar measures to theirs for our administrative data. However, in separate sets of regressions shown in Table C.6 we find that higher management hours shares, especially among CEOs and "cadres"—top managers strongly correlate with productivity measures.

[32]Further investigation is reported in Section C.1 of the Appendix.

Is R&D driving the results? To complete the picture, we find that no single non-routine function is driving the correlation with productivity and that this connection does not simply result from the number of distinct functions firms have.

First, we show that no particular non-routine function is driving this correlation. To this end, we work directly with the main functions as determined by INSEE, and display the results in regressions (4), (5) and (6) in Table 9. B-2-B, R&D and maintenance (routine function in our classification, but much less than production, management or transport and logistics) are statistically significantly correlated with productivity and Intellectual Services are so in the sample with 95% of observations (last column). An increase of 1 percentage point in the employment shares in any of these functions at the firm level correlates with a higher TFP at least of 0.1%. The only exception is Transport and Logistics that is not correlated with productivity in any of our specifications.

Second, we obtain that the sheer count of distinct functions in the firm is not correlated with productivity (Figure 8, lower panel) in contrast with a strong correlation with size (Figure 8, upper panel). This, along with the evidence presented above, shows that it is not the number of functions present within the firm (the extent of firm's labor force "diversification") per se but the employment share *structure* that is correlated with productivity.

4.4 Non-routine service functions and profitability

The ultimate success of a firm is the ability to generate profits. We measure profitability[33] by dividing total revenue by total costs from firms' income statements. The correlation of the share of non-routine service functions with profitability are positive and statistical in the cross-section in 2015 (columns 1 and 2 of Table 10; column 4 for the logarithm of profitability), as is the 5-year change in the share of non-routine share between 2005-2010 and the change in profitability over the 2010-2015 period (column 3 of Table 10). We conclude that higher shares of non-routine service workers are associated with higher profitability.

5 A simple model

In this section, we rationalize the documented facts with a simple model of heterogenous firms that decide on the type of goods that they produce as well as whether they dedicate labor to knowledge generation.

In our model, firms can produce differentiated goods or an homogenous good. The differentiated

[33] In the literature, profitability is shown to be positively correlated to firm markups. See De Loecker et al. (2020) among others.

good is complex to supply in the sense that its supply depends on a choice made the firm—this choice can be either to find the exact good that meets consumers' preferences, the set of inputs required to produce or the sequence of production. In this context, knowledge helps firms to become more efficient producers of the complex good.

We start by assuming that knowledge has to be generated within the firm using specialized labor. We then extend our insights to the endogenous choice of whether the firm should either generate knowledge within or outside the borders of the firm. Overall, we obtain, as in the data, that more productive firms are producing more complex goods, and have higher shares of specialized labor dedicated to knowledge generation within the firm—i.e., non-routine service functions.

5.1 The environment.

Let us consider two firms $i \in \{1, 2\}$. Both firms can produce a good that can either be simple or complex and we denote by $q \in \{simple, complex\}$ the corresponding complexity.

Production. To produce a good of complexity q, we assume that firms have to use labor but they also have to make decisions. We denote by l_i the amount of labor that firm i uses. Taking a decision amounts here to select $k \in \Omega$ where $\Omega = \{1, 2, ...N\}$ and $N = Card(\Omega)$. In the end, the number of goods of complexity q that the firm produces is in the end $A_i G(q, k, l)$, where A_i is firm i's productivity. We assume that G is an increasing, concave and differentiable function with respect to labor l. Finally, we denote by w the wage rate paid on labor.

The production of the complex good is more sensitive to the firm's decision than the production of the simple good. Formally, to make things simple, we assume an extreme form of sensitivity for the production of the complex good: there exists $k^\star \in \Omega$ such that $G(complex, k^\star, l) > 0$ and $G(complex, k, l) = 0$ for any $k \neq k^\star$. We refer to $k^\star$ as the right decision. In contrast, we assume that the production of the simple good does not depend on the choice of inputs: $G(simple, k, l)$ is constant for any $k \in \Omega$. Finally, we assume the following function form for G:

$$G(complex, k, l) = 1_{k=k^\star} l^\theta \text{ and } G(simple, k, l) = l^\theta,$$

with $\theta \in (0, 1)$.

Knowledge generation. Firms cannot freely observe which decision is the right one, $k^\star$, but can decide on their information set $\mathcal{I}_i$. This requires, however, to produce knowledge. Formally, firms can decide on their information set $\mathcal{I}_i$ but this requires to generate knowledge, that we denote by K. Depending on this stock of knowledge, firms can select its information set under the following constraint:

$$H(\mathcal{I}_i|\mathcal{I}_0) \leq K_i, \tag{4}$$

where $H(\mathcal{I}_i|\mathcal{I}_0)$ is the relative entropy—also known as the Kullback-Leibler divergence—that, following Sims (2003), measures the informational content of $\mathcal{I}_i$ relative to $\mathcal{I}_0$, i.e. the initial information set. Constraint (4) is the only one on selecting the information set and we do not assume any constraints on the signals that the firm can receive to select this information set.

We assume that, absent knowledge generation, $\mathcal{I}_0$ is such that the firm has uninformative priors on which decision is the right one,$k^\star$, and, for each $k \in \Omega$, they assign a prior probability $1/N$ that k is $k^\star$.

To generate knowledge, firms can use labor specific labor in-house—we label $l_{in,i}$ such in-house knowledge-generating labor—also paid at the wage rate w. We assume that knowledge is accumulated according to a linear production technology $K_i = l_{in,i}$.

Demand for goods. Finally, we assume, for expositional simplicity, that a representative consumer values the consumption of complex and simple goods with the following linear preferences:

$$U(c_{simple}, c^1_{complex}, c^2_{complex}) = c_{simple} + \alpha\left(c^1_{complex} + c^2_{complex}\right)$$

with $\alpha > 1$. This sets the price of the simple good at $\pi_{simple} = 1$ and the prices of complex goods at $\pi^1_{complex} = \pi^2_{complex} = \alpha > \pi_{simple} = 1$.[34]

Perfect information. As a benchmark, let us clarify that knowledge is not useful if the function $G(q, k, l)$ is perfectly observable at no cost. Firms are then perfectly able to select the right input $k^\star \in \Omega$. In addition, no labor is required to generate knowledge: $l_K = 0$.

Interpreting the model. Several comments are in order about complex goods and the role of knowledge in our model.

Mapping with empirics. Let us first connect the objects of the models to Figure 4 that organizes the concepts that we use for the empirical part. Non-routine service functions correspond to the specific labor that firms employ in-house $l_{in,i}$ to generate knowledge K_i. The capabilities of the firm can be summarized by the set of goods that the firm can produce (either the simple or the complex one). Finally, we will measure the performance of the firm by its profits and its apparent productivity—which we define more formally later.

The decision. The decision needed to produce the complex goods may have different interpretations, corresponding

[34]Notice that our results hold more generally with downward-sloping demand (see Section 5.4.

Interpretation 1 (Sourcing of inputs). Ω is a set of potential inputs and the firm has to decide which input $k \in \Omega$ to select. In this case, the function G is of the kind

$$G(q,k,l) = \sum_{k' \in \Omega} a_{k',q} 1_{k'=k} l^{\theta}$$

with $\theta \in (0,1)$. The $a_{k',q}$ is how productive it is to use the input k to produce the good of complexity q. Our assumption of extreme complexity is that $a_{k,'simple} = 1$ for all $k' \in \Omega$ and there exists $k^{\star} \in \Omega$ such that $a_{k^{\star},complex} = 1$ and $a_{k,complex} = 0$ for any $k \neq k^{\star}$.

Interpretation 2 (R&D). Let us consider, for example, a set of inputs $\mathcal{J} \subset \mathcal{N}$ with $M = Card(M) < \infty$. To produce a good, the firm needs to find the right order for assembling these inputs through R&D. In this case, Ω is the set of permutations over $\mathcal{J}$ and $N = Card(\Omega) = M!$. For example, an order of production is a mapping σ from $\mathcal{J}$ to $\mathcal{J}$ and the right order is $\sigma^{\star}$. Producing a set of inputs requires labor so that the production of the complex good is:

$$\Pi_{j=1}^{\mathcal{J}} \Pi_{m=i}^{j} 1_{\sigma(m)=\sigma^{*}(m)} l^{\theta}.$$

Indexing mappings from $\mathcal{J}$ to $\mathcal{J}$ by $k \in \Omega$ with $k^{\star}$ the index for $\sigma^{\star}$, this production can be rewritten: $1_{k=k^{\star}} l^{\theta}$ as in our model.

Interpretation 3 (Marketing). Ω is the set of characteristics of the complex good that the firm can supply and, to meet demand, the firm needs to select the right characteristics. For example, the price of a complex good with characteristics k is $\pi^{k}_{complex}$—in this case, the utility function for the representative consumer is $c_{simple} + \sum_{k \in \Omega} \alpha^{k} c^{k}_{complex}$. Under our assumption of extreme complexity, $\pi^{k^{\star}}_{complex} > 1$ and $\pi^{k}_{complex} = 0$ for any $k \neq k^{\star}$.

Complexity. In this setup, we use a cardinality-based notion of complexity similar to that in the literature on complexity in games as in, e.g., Rubinstein (1986). Here, there are more payoff-relevant states to consider in the production of the complex good.

The role of knowledge. Our assumptions capture two aspects of knowledge and the boundaries of the firm consistent with theories like the resource-view one. On the one hand, knowledge produced within the firm can be used as an input for other (physical) production within the firm. On the other hand, to what extent this knowledge allows the firm to benefit from high markups depend on whether the other firm also produces a complex good (Section 5.4). Internally-generated knowledge becomes the critical resource of the firm and the source of its comparative advantage in producing the complex good.

5.2 Optimal firm structure

In this subsection, we derive the optimal firm structure in the case where productivities $a_{k,q}$ cannot be freely observed by the firm. Given productivity A and initial information set $\mathcal{I}_0$, the problem faced by firms 1 and 2 is to maximize profits by selecting a level of complexity q, a choice k, the amount of production labor l and knowledge-generating labor l_{in} as well as an information set $\mathcal{I}$. More formally:

$$\max_{q\in Q,l,k,l_K,\mathcal{I}} A\mathbb{E}\left[\pi_q 1_{k=k'} l^\theta|\mathcal{I}\right] - w\left(l + l_{in}\right)$$
$$\text{s.t. } H(\mathcal{I}|\mathcal{I}_0) \leq l_{int}$$

where $\mathbb{E}(.|\mathcal{I})$ is the expectation operator conditional on the information set $\mathcal{I}$ and π_q is the price of the good of quality q. We drop the index i to ease reading.

Optimal knowledge generation. Let us first discuss the incentives by the firm to generate knowledge given a choice of good complexity.

First, let us note that (4) is always binding as, otherwise, the firm would reduce its amount of labor dedicated to knowledge generation l_{in} and strictly increase its profits. Second, a firm deciding to produce the simple good is better off not generating knowledge. It would imply a positive cost of $wl_{in} > 0$ and there is no gain to generate knowledge for the simple good as any choice k leads to the same production.

In contrast, not generating knowledge to produce the complex good may not be optimal. In this case, when taking a decision $k' \in \Omega$, the firm expects a production

$$A\mathbb{E}\left[\pi_{complex} 1_{k=k'} l^\theta|\mathcal{I}_0\right] = \frac{\pi_{complex} A}{N} l^\theta.$$

A larger set of choices Ω—and so, a larger $N = Card(\Omega)$ leads to a lower expected productivity of producing the complex good in the absence of knowledge generation.

On the other hand, generating knowledge is costly. Let us start with an example. Suppose that the firm would like to have perfect information to supply the complex good. Given its prior set of information, the relative entropy of the new information set is:

$$H(\mathcal{I}|\mathcal{I}_0) = \sum_{k\in\Omega} P(k) \log\left(\frac{P(k)}{Q(k)}\right).$$

where $Q(k)$ is the probability that k is the right choice based on $\mathcal{I}_0$ and $P(k)$ is this probability based on $\mathcal{I}$. By assumption, $Q(k) = 1/N$ and under perfect information $P(k') = 1$ for the right choice k' and equals 0 otherwise. As $\lim_{x\to 0} x\log x = 0$, the relative entropy is, in this case: $H(\mathcal{I}|\mathcal{I}_0) = \log N > 0$. There is then a need to generate knowledge due to (4) and select a strictly positive l_{in}.

To make the problem tractable, we consider the information structure $\mathcal{I}(p)$ so that the posterior distribution is as follows. A given $k \in \Omega$ is the right choice for supplying the complex good with probability $p \geq 1/N$ and all the other $k' \neq k$ are the right choices with probability $1/N-(p-1/N)/(N-1)$. Thus, we obtain a continuum between $p = 1/N$ and $p = 1$. In particular, we have that $p = \mathbb{E}\left[1_{k=k'}|\mathcal{I}\right]$.

The following Lemma describes how the firm generates knowledge as a function of the level of complexity of their production:

Lemma 1 (Knowledge generation)**.** *The firm producing the complex good generates knowledge:* $l_{in} > 0$*. This contrasts with a firm producing the simple good, which does not generate knowledge:* $\mathcal{I} = \mathcal{I}_0$ *and* $l_{in} = 0$*.*

Furthermore, there is no loss of generality to focus on the information sets $\{\mathcal{I}(p)\}_{p\in[1/N,1]}$ *and:*

(i) *Knowledge-generation* l_{in} *and the probability* p *are increasing with productivity* A*.*

(ii) *The probability* p *is an increasing and concave function of knowledge-generating labor* l_{in}*.*

Proof. See Appendix A.1. □

First of all, as we noted, we do not make any assumptions on the set of signals that the firm can receive and so, without loss of generality, we can directly focus on posterior beliefs.

In the left panel of Figure 9, we illustrate Lemma 1's result by plotting the optimal probability p as a function of firm's productivity A, conditional on supplying the complex good. Our calibration is such that $N = 2$ and so, the probability p takes value from $1/N = .5$ to 1. The right panel of Figure 9 illustrates the "production function" of knowledge: by increasing the amount of labor dedicated to knowledge generation, a firm increases its probability to make the right choice to supply the complex good. This production function features decreasing returns stemming from the convexity in p of conditional entropy.

Good selection and firm's structure. In the end, knowledge generation is tied to the production of the complex good so that a firm's decision to produce the complex good boils down to compare:

$$Ap\pi_{complex}l^{\theta}_{complex} - wl_{complex} - wl_{in} \geq \pi_{simple}Al^{\alpha}_{simple} - wl_{simple},$$

with $l_{complex} = (Ap\pi_{complex}/w)^{\frac{1}{1-\alpha}}$ the optimal amount of labor in production when producing the complex good and $l_{simple} = (A/w)^{\frac{1}{1-\alpha}}$ the optimal amount of labor in production when producing the simple good—in this case, the supply of the good does not depend on a decision that the firm has to take.

When $\pi_{complex} > \pi_{simple} = 1$, the marginal value of production is potentially larger for the complex good but, at the same time, knowledge generation leads to a larger cost for producing this good. This comparison of higher marginal value with larger cost leads to the following proposition:

Proposition 2 (Optimal firm structure and good production)**.** *There exists $\bar{A}$ so that a firm with $A \geq \bar{A}$ generates knowledge and produces the complex good. Otherwise, a firm with productivity $A < \bar{A}$ does not generate knowledge and produces the simple good.*

Proof. See Appendix A.2. □

Only when sufficiently productive, a firm engages in knowledge generation and produces the complex good. In contrast, when being less productive, firms specialize in the simple good and, accordingly, do not generate knowledge.

Per se, the result that more productive firms engage in higher value activities is not new.[35] However, this self-selection of higher-productivity firms into the production of the complex good does not stem from a *fixed cost* of producing the complex good but from the comparison of the marginal gain to produce such good: due to endogenous knowledge generation, a low-productivity firm is relatively more productive for the simple good than for the complex good. This pattern is reversed for high-productivity firms that generate knowledge and benefit from a higher probability to make the right decsion $k = k^{\star}$. In a way, the need for knowledge generation leads to a form of an adjustment cost to allow the organization to produce the complex good.

To illustrate this finding, we plot in Figure 10 profits from producing the simple and the complex goods as a function of productivity. As it can be observed, both goods yield 0 profits when productivity equals 0. However, profits when producing the complex good are more convex than when producing the simple good. This difference in convexity happens despite the production function is the same for the two goods but only due to the *endogenous* knowledge generation choice on the probability p—the stronger increase in the slope results from p increasing with productivity (left panel of Figure 9).

A consequence of this result is that the existence of productivity thresholds and adjustment costs of employing labor specialized in knowledge generation may cause a staggered reaction of firms to productivity shocks as found by Pozzi and Schivardi (2016) or no reaction at all if firms may not be productive enough as in Atkin et al. (2017).

Productivity advantage to produce the complex good. That a more productive firm produces more complex goods as stated by Proposition 2 is not only due to self-selection but also results into a productivity gain for such a firm. In this paragraph, we make formal this point.

[35] See Melitz (2003) among others and the general conditions for this to happen in Mrázová and Neary (2019).

For a firm producing the complex good, we can define the productivity level that it would have when producing the simple good and making the same level of profits. Formally, for knowledge-generating labor l_{in}, production labor l, $A_{apparent}$ is such that:

$$\pi_{complex} p A l^{\theta} - w(l + l_{in}) = \pi_{simple} A_{apparent} (l + l_{in})^{\theta} - w(l + l_{in})$$

From Proposition 2, we can state the following:

Corollary 3. *When $A \geq \bar{A}$, $A_{apparent} \geq A$, with strict inequality when $A > \bar{A}$.*

Proof. The case where $A = \bar{A}$ is trivial as the firm is indifferent between the complex and the simple good. Suppose that $A > \bar{A}$. Suppose that $A_{equivalent} \leq A$. Then the firm is strictly better off not producing the complex good to produce the simple good, which would yield a profit of at least $\pi_{simple} A(l + l_{in})^{\theta} - w(l + l_{in})$. □

5.3 The role of management

Let us now extend our model to think about the role of management. To this purpose, we consider management labor l_M. For simplicity, we assume that knowledge generation takes place inside firms. Management is arguably complement to production but it is also complement to knowledge generation as management, consistently with the literature, plays an important role in gathering, disseminating, processing (Radner, 1993) or coordinating knowledge produced within the firm. From this perspective, adding service functions may contribute to increasing the need for coordinating different tasks: marketing, R&D and production for example.

To model these complementarities, we focus on the following modified problem for the firm:

$$\max_{q \in Q, l, l_M l_K, \mathcal{I}} A\pi(q) E\left[1_{k=k^\star \text{ or } q=simple} l^{\theta} | \mathcal{I}\right] l_M^{\beta} - w\left(l + l_{in} + l_M\right), \tag{5}$$

$$\text{s.t. } H(\mathcal{I}|\mathcal{I}_0) \leq l_{in}^{\gamma} l_M^{\delta} \tag{6}$$

with θ, β, γ and δ positive coefficients such that $\theta + \beta < 1$ and $\gamma + \delta \leq 1$.[36]

The first order condition with respect to management writes:

$$E\left[A\pi(q) 1_{k=k^\star \text{ or } q=simple} l^{\theta} | \mathcal{I}\right] \beta l_M^{\beta-1} + \lambda (l_{in})^{\gamma} \delta l_M^{\delta-1} = w$$

with λ the Lagrange multiplier associated with (6). Inspecting this condition, one can observe that management labor l_M increases with production labor l and firm's productivity A_i but l_M also increases

[36]This formulation encompasses many approaches in the literature. For example, if $\gamma = \delta$, $\theta = \beta$ the wage bill for management is equivalent to a coordination cost as in Becker and Murphy (1992).

with knowledge-generating labor l_{in} and the shadow value of knowledge generation as measured by the Lagrange multiplier λ. The rest of the analysis is not modified with respect to the previous paragraph.

In such a setting, higher productivity leads to more management labor all the more when it also leads to more labor dedicated to knowledge generation:

Corollary 4. *When management is complement to knowledge generation, firms generating more knowledge hire relatively more management labor.*

We illustrate this finding in Figure 11 where we plot knowledge-generating and management labor as a function of productivity for a calibrated version of the model. Consistently with our results, only the most productive firms are producing the complex good and then also generate knowledge and hire specific labor—this can be observed in the Figure as knowledge-generating labor is plotted by the red dashed line. Such presence of knowledge labor leads to a higher demand for management labor as in Corollary 4, which can be observed in the Figure by the larger elasticity of management labor with respect to productivity for firms generating knowledge—management labor is plotted by the black plain line.

Remark. In this paragraph, we have left unmodeled the precise motive for the complementarity between knowledge-generating and management labor. In models of the organization of the firm, analyzing such complementarity in more details would allow to investigate how knowledge labor should be organized: should it be centralized at the level of the firm, e.g., to benefit from increasing returns, or decentralized to adapt to local conditions. We leave the related questions to future research.

5.4 Strategic knowledge generation within the firm

Why firms are producing knowledge in-house? In this section, we enrich the model along two dimensions to answer this question. First, we allow firms to generate knowledge either inside—as in the benchmark model—but also outside the firm, through expert firms. Second, to allow for strategic interactions between the two firms, we assume that the complex goods supplied by the two firms are substitutes.

The key motive to keep knowledge generation within the firm in this extension is that knowledge generation outside the firm can exacerbate competition. We obtain this by assuming that expert firms cannot commit not to sell the knowledge that they have produced to other firms (see, among others Veldkamp, 2006, for a similar assumption). As a result of this incomplete market friction, purchases of knowledge outside the firm also make knowledge more available to other firms making them more likely to also produce competitive complex goods that can substitute for the firm's own products.

Knowledge generation inside and outside the firm. To generate knowledge, we now assume that firms can either use labor specific labor inside—as previously, we label $l_{in,i}$ such knowledge-generating labor inside the firm—or the firm can purchase knowledge outside the firm—we denote by $l_{out,i}$ such purchases. We assume that knowledge is then accumulated according to $K_i = l_{in,i} + l_{out,i}$. Purchases of outside knowledge take place at the endogenous price P_{out}.

Outside knowledge is produced by competitive expert firms. Expert firms sell knowledge at the price P_{out} and they use labor L_{out} at a wage rate w to generate knowledge. Importantly, when an expert firm produces knowledge for one firm, we assume that the consulting firm can sell this knowledge also to the other firm. In particular, we assume that it is not possible to write contracts to prevent this. The only constraint is that the consulting firm can sell only what she has generated in terms of knowledge. In the end, expert firms maximize:

$$P_{out}\left(l_{out,1} + l_{out,2}\right) - wL_{out} \text{ with } L_{out} \geq \max_{i\in\{1,2\}} l_{out,i}.$$

The free entry condition in the expert firm sector implies average cost pricing:

$$P_{out} = \frac{w \max_{i\in\{1,2\}} l_{out,i}}{l_{out,1} + l_{out,2}}. \tag{7}$$

Demand for goods. As before, we assume that the price of the simple good is $\pi_{simple} = 1$. The prices of complex goods are as follows:

$$\pi^1_{complex} = \alpha - \gamma c^1_{complex} - \eta c^2_{complex},$$
$$\pi^2_{complex} = \alpha - \gamma c^2_{complex} - \eta c^1_{complex},$$

with α, γ and η positive parameters and $\alpha > 1$. The important element here is that the price of complex good of one firm is lower when the other firm also produces a complex good. Notice that these demand functions can be obtained from assuming a linear-quadratic utility function.

Strategic internal knowledge generation. Let us investigate the decision to keep inside knowledge generation or to purchase knowledge outside. We focus on results about the split of knowledge generation K between l_{out} and l_{in}.

The problem to solve for a firm is:

$$\max_{l,l_{in},l_{out},p} Ap\pi_{complex}(l_{out})l^{\theta} - w\left(l + l_{in}\right) - P_{out}l_{out}.$$

We focus on firm 1 that, as it is more productive, also produce more knowledge. Two elements are important. On the one hand, generating knowledge outside leads to a lower price $\pi_{complex}(l_{out})$. The reason is that generating knowledge outside leads to a lower price P_{out} for the other firm that can then produce more of its complex good. On the other hand, shifting from knowledge generation outside to inside the firm is costly. Of course, this cost is an increasing function of the targeted level of knowledge, but when productivity is sufficiently large, the targeted level of knowledge levels off. As a result, the comparison between the gains to keep knowledge generation in-house and the costs is standard and leads to:

Proposition 5. *When A^1 is large enough, firm 1 generates knowledge internally. The threshold on A^1 is a decreasing function of η, the degree of substitutability between the complex goods produced by firms 1 and 2.*

Proof. See Appendix A.3. □

Several comments are in order. First, more productive firms have an incentive to generate knowledge internally: by doing so, they enjoy higher markups on their production but at a higher cost for knowledge generation—generating knowledge inside is, at least weakly, costlier than generating knowledge outside.

Naturally, this incentive to generate knowledge inside is a function of degree of competition with firm 2: more substitutability between the two complex goods that the two firms produce reduces the market power of firm 1. Also, if there are other gains to generate knowledge within the firm—e.g., due to communication costs outside the firm or there is learning by doing—, firm 1 is also more prone to keep knowledge generation within the firm.

Remark. It is important to note that Proposition 5 is not inconsistent with more productive firms outsourcing more. In particular, they may have a comparative advantage compared with less productive firms to outsource tasks that are useless to accumulate knowledge and to guarantee market power.

5.5 Empirical implications and further discussion

In this final subsection, we first connect the results from this section to our empirical findings in the previous sections. Then, we discuss the type of workers that are needed for knowledge generation and how this fits with what we observe for non-routine service functions. Finally, we discuss the relation of our model to various theories of the firm.

Empirical implications. We derive a set of empirical implications of the model. To this purpose, let us consider two firms with two different productivities levels $A_{low} < A_{high}$.

Using our previous results, we obtain the following proposition that makes predictions on the difference between these two firms.

Proposition 6. *Suppose that $A_{low} < \bar{A} < A_{high}$. Then:*

(i) *The high-productivity firm produces a more complex good than the low-productivity one.*

(ii) *Its share of knowledge-generating labor is larger.*

(iii) *Its share of management labor is higher.*

(iv) *Its knowledge generation is higher.*

(v) *Its markup is higher.*

More productive firms engage in the production of more complex goods, hire more knowledge-generating labor and, concomitantly, more management—under the assumption that knowledge generation is complement to management.

What kind of workers are employed in knowledge generation? To link our results to data, it is also useful to make more precise what we mean by labor engaged in knowledge generation. In our benchmark model, labor responsible for generating knowledge reduces the uncertainty regarding the decision critical in the supply of the complex good. Two functions within firms seem relevant for this role: R&D and B-2-B (purchases). The former is about designing the right product, i.e. identifying the right set of inputs that one needs and the way to assemble them in order to produce a good, and the latter is about making sure to have the actual sourcing of (high quality) inputs.

As we discussed, our model can also encompass goods that are complex to sell and so, knowledge generation may be important in finding the ideal variety desired by consumers or informing customers about firm products. In our data, jobs in advertising, marketing, sales or economic consultants that may critical for assessing the precise demand are gathered in the intellectual services or B-2-B (sales) functions. R&D workers are key in producing useful knowledge about technologies and production while lawyers on the legal challenges (e.g., driverless cars, intellectual property protection) on the feasibility of producing different products. All in all, our model rationalizes that workers in B-2-B, R&D and intellectual services generate knowledge that raises the efficiency of complex good production and leads to higher productivity and profitability of the firm.

Our model and theories of the firm. How does our model fit within existing theories of the firm? One alternative for explaining why non-routine service functions are within firms would be the property-rights theory.

Indeed, service functions with low routineness scores that generate knowledge involve by definition tasks that are less standardizable, requiring more customized actions than other functions like production. This means that for employees performing them, by the nature of the task or their skill, their effort and final output may not be easily measurable and therefore monitored. The way these functions are performed may require firm-specific investments on the side of workers but and/or yield firm-specific output. That naturally gives rise to contractual incompleteness (see, e.g., Costinot et al., 2011) in the provision of non-routine labor services.

These functions, however, also require high human capital and may not require many physical assets for production. To the contrary, they may need more intangible capital. The understanding of why such functions are kept within firms is thus largely outside of the scope of the property-rights theory (see Rajan and Zingales, 1998): one of the contractual parties cannot *own* employees, and the mere fact of employing them within the organization as opposed to sourcing their services as outside contractors may not solve any hold-up problems.[37]

One answer to this can be that of the critical resource theories of the firm going back to Wernerfelt (1984)—that employment of non-routine service workers within the firm provides access to some critical resource. In our case, this would be the knowledge generated by these workers that would become a source of comparative advantage of those firms. While Rajan and Zingales (1998) provide an explanation why employees with high human capital may be retained within the borders of organizations, it does not explain why larger or more productive firms would acquire a more *diverse* workforce, where workers have multiple functions; furthermore, they consider the "critical resource" to be exogenously given. In our framework, we abstain from modeling the contractual relations between owners and workers, but focus on the creation of the critical resource in the form of generated knowledge.

6 Conclusion

In this paper, we provide new facts on the structure of firms and on service functions that constitute an important fraction of employment in manufacturing firms. These facts, and the important role non-routine service functions play in firms cannot be easily rationalized by existing vertical or hierarchical theories of the firm. We rationalize these facts by a model in the spirit of critical resource theories of the firm where firms can generate knowledge resolving uncertainty to produce higher-value complex goods using specific labor. Model predictions are consistent with the data; higher shares of non-routine service functions and management are correlated with higher productivity and profitability; and measures of knowledge generation and product complexity. Our results suggest that functional composition of the workforce is another important driver of TFP.

Our description of service functions is mainly in the cross-section. A natural question is about the time evolution of service functions. First, the nature of some of those functions may have changed as well; for example, some IT services became standardized, codifiable and therefore could be outsourced as they cease to be a part of important comparative advantage of (manufacturing) firms. The development of tradable business services in many countries is a witness to that. Moreover, the relative importance

[37]See also comments by Hart (2017), p. 1735.

of these functions may have changed over time, either for some or all firms, as a result of the need to generate more knowledge to produce ever more complex goods. From this perspective, an interesting question is whether the resulting evolution of the ability of firms to engage in more complex production has allowed only some firms to increase their market power, consistently with the rise in markups explored by De Loecker et al. (2020). Perhaps engaging in employment structure transformation has also permitted firms to cope with increased (international) competition similarly to increased innovation found in Bloom et al. (2016). An increased share of locally provided, within-firm services in firms makes their goods less *tradable* and open to outside competitive forces. Causal inference of the role of functions in firms would be important as well. We leave these crucial questions for future research.

We identify a tension in firm organization. The need to generate knowledge to produce complex goods may increase the need for coordination (Becker and Murphy, 1992), and be complementary with more intensive management, while successful experimentation by non-routine service workers may require autonomy that in turn invites decentralization and adaptation (Dessein and Santos, 2006). This introduces interesting tradeoffs that should be addressed.

Bibliography

ANTRÀS, P. (2003): "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, 118, 1375–1418.

ATALAY, E., A. HORTAÇSU, AND C. SYVERSON (2014): "Vertical Integration and Input Flows," *American Economic Review*, 104, 1120–1148.

ATKIN, D., A. CHAUDHRY, S. CHAUDHRY, A. K. KHANDELWAL, AND E. VERHOOGEN (2015): "Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan," *American Economic Review*, 105, 537–544.

ATKIN, D., A. CHAUDHRY, S. CHAUDRY, A. K. KHANDELWAL, AND E. VERHOOGEN (2017): "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," *The Quarterly Journal of Economics*, 132, 1101–1164.

ATKIN, D., A. COSTINOT, AND M. FUKUI (2021): "Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom?" NBER Working Papers 29500, National Bureau of Economic Research, Inc.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The skill content of recent technological change: An empirical exploration," *Quarterly journal of economics*, 118, 1279–1333.

BANDIERA, O., A. PRAT, S. HANSEN, AND R. SADUN (2020): "CEO Behavior and Firm Performance," *Journal of Political Economy*, 128, 1325–1369.

BASTOS, P., J. SILVA, AND E. VERHOOGEN (2018): "Export Destinations and Input Prices," *American Economic Review*, 108, 353–92.

BECKER, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70, 9–49.

BECKER, G. S. AND K. M. MURPHY (1992): "The Division of Labor, Coordination Costs, and Knowledge," *Quarterly Journal of Economics*, 107, 1137–1160.

BENDER, S., N. BLOOM, D. CARD, J. V. REENEN, AND S. WOLTER (2018): "Management Practices, Workforce Selection, and Productivity," *Journal of Labor Economics*, 36, 371–409.

BERGEAUD, A., C. MAZET-SONILHAC, C. MALGOUYRES, AND S. SIGNORELLI (2021): "Technological Change and Domestic Outsourcing," IZA Discussion Papers 14603, Institute of Labor Economics (IZA).

BERNARD, A. B. AND T. C. FORT (2015): "Factoryless Goods Producing Firms," *American Economic Review*, 105, 518–23.

BLOOM, N., M. DRACA, AND J. V. REENEN (2016): "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity," *Review of Economic Studies*, 83, 87–117.

BLOOM, N., L. GARICANO, R. SADUN, AND J. V. REENEN (2014): "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science*, 60, 2859–2885.

CALIENDO, L., G. MION, L. D. OPROMOLLA, AND E. ROSSI-HANSBERG (2020): "Productivity and Organization in Portuguese Firms," *Journal of Political Economy*, 128, 4211–4257.

CALIENDO, L., F. MONTE, AND E. ROSSI-HANSBERG (2015): "The anatomy of French production hierarchies," *Journal of Political Economy*, 123, 809–852.

CHANDLER, A. D. (1962): *Strategy and Structure*, MIT Press.

COSTINOT, A., L. OLDENSKI, AND J. RAUCH (2011): "Adaptation and the Boundary of Multinational Firms," *The Review of Economics and Statistics*, 93, 298–308.

DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 135, 561–644.

DEFEVER, F. (2006): "Functional fragmentation and the location of multinational firms in the enlarged Europe," *Regional Science and Urban Economics*, 36, 658–677.

DESSEIN, W., A. GALEOTTI, AND T. SANTOS (2016): "Rational Inattention and Organizational Focus," *American Economic Review*, 106, 1522–36.

DESSEIN, W. AND A. PRAT (2022): "Organizational Capital, Corporate Leadership, and Firm Dynamics," *Journal of Political Economy*, forthcoming.

DESSEIN, W. AND T. SANTOS (2006): "Adaptive Organizations," *Journal of Political Economy*, 114, 956–985.

FORT, T. C. (2016): "Technology and Production Fragmentation: Domestic versus Foreign Sourcing," *The Review of Economic Studies*, 84, 650–687.

GARICANO, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108, 874–904.

GARICANO, L., C. LELARGE, AND J. V. REENEN (2016): "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 106, 3439–3479.

GARICANO, L. AND T. VAN ZANDT (2013): "Hierarchies and the Division of Labor," in *The Handbook of Organizational Economics*, ed. by R. Gibbons and J. Roberts, Princeton: Princeton University Press, chap. 15, 604–654.

GARICANO, L. AND Y. WU (2012): "Knowledge, Communication, and Organizational Capabilities," *Organization Science*, 23, 1382–1397.

GIORCELLI, M. (2019): "The Long-Term Effects of Management and Technology Transfers," *American Economic Review*, 109, 121–52.

GOOS, M., A. MANNING, AND A. SALOMONS (2014): "Explaining job polarization: Routine-biased technological change and offshoring," *American Economic Review*, 104, 2509–26.

GRANT, R. M. (1996): "Toward a knowledge-based theory of the firm," *Strategic management journal*, 17, 109–122.

GROSSMAN, S. J. AND O. D. HART (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.

HARRIGAN, J., A. RESHEF, AND F. TOUBAL (2021): "The March of the Techies: Job Polarization Within and Between Firms," *Research Policy*, 50.

HART, O. (2017): "Incomplete Contracts and Control," *American Economic Review*, 107, 1731–52.

HART, O. AND J. MOORE (1990): "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119–1158.

HELPMAN, E. (1984): "A Simple Theory of International Trade with Multinational Corporations," *Journal of Political Economy*, 92, 451–471.

JOVANOVIC, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50, 649–670.

KOGUT, B. AND U. ZANDER (1992): "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology," *Organization Science*, 3, 383–397.

LAZEAR, E. P. (2009): "Firm-Specific Human Capital: A Skill-Weights Approach," *Journal of Political Economy*, 117, 914–940.

LOCKETT, A., S. THOMPSON, AND U. MORGENSTERN (2009): "The development of the resource-based view of the firm: A critical appraisal." *International Journal of Management Reviews*, 11, 9 – 28.

MACKOWIAK, B. A., F. MATEJKA, AND M. WIEDERHOLT (2020): "Rational Inattention: A Review," CEPR Discussion Papers 15408, C.E.P.R.

MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

MRÁZOVÁ, M. AND J. P. NEARY (2019): "Selection Effects with Heterogeneous Firms," *Journal of the European Economic Association*, 17, 1294–1334.

PATAULT, B. AND C. LENOIR (2021): "How valuable are business networks? Evidence from sales managers in international markets," Mimeo, University of Amsterdam and Insee.

PORTER, M. E. (1985): *Competitive advantage: creating and sustaining superior performance*, New York: FreePress.

POZZI, A. AND F. SCHIVARDI (2016): "Demand or productivity: what determines firm growth?" *RAND Journal of Economics*, 47, 608–630.

RADNER, R. (1993): "The Organization of Decentralized Information Processing," *Econometrica*, 61, 1109–1146.

RAJAN, R. G. AND L. ZINGALES (1998): "Power in a Theory of the Firm," *Quarterly Journal of Economics*, 113, 387–432.

——— (2001): "The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms*," *The Quarterly Journal of Economics*, 116, 805–851.

RAUCH, J. E. (1999): "Networks versus markets in international trade," *Journal of International Economics*, 48, 7–35.

RUBINSTEIN, A. (1986): "Finite automata play the repeated prisoner's dilemma," *Journal of Economic Theory*, 39, 83–96.

SIMS, C. A. (2003): "Implications of rational inattention," *Journal of monetary Economics*, 50, 665–690.

SUTTON, J. (2012): *Competing in Capabilities*, Oxford University Press.

SUTTON, J. AND D. TREFLER (2016): "Capabilities, Wealth, and Trade," *Journal of Political Economy*, 124, 826–878.

TAMBE, P., L. HITT, D. ROCK, AND E. BRYNJOLFSSON (2020): "Digital Capital and Superstar Firms," NBER Working Papers 28285, National Bureau of Economic Research, Inc.

TEECE, D. J. (1982): "Towards an economic theory of the multiproduct firm," *Journal of Economic Behavior & Organization*, 3, 39–63.

TEECE, D. J., G. PISANO, AND A. SHUEN (1997): "Dynamic capabilities and strategic management," *Strategic Management Journal*, 18, 509–533.

TOWNSEND, R. M. (1979): "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 21, 265–293.

VELDKAMP, L. L. (2006): "Media Frenzies in Markets for Financial Information," *American Economic Review*, 96, 577–601.

WERNERFELT, B. (1984): "A resource-based view of the firm," *Strategic Management Journal*, 5, 171–180.

ZEIRA, J. (1987): "Investment as a Process of Search," *Journal of Political Economy*, 95, 204–210.

Tables and Figures

Table 1: Further description of functions using INSEE documentation

Code	Function	Further description	Example of occupations
CONREC	R&D	Jobs in conception, research and innovation.	Engineers and technicians in R&D.
PREINT	Intellectual services	Jobs providing specific knowledge for consulting, expertise, etc.	Lawyers, advertising, communication, IT, architects, etc.
AGRICU	Agriculture and fishing	Jobs in agriculture, fishing, lumbering	Farmer, Farm hands, etc.
BTP	Construction and public works	–	Engineers, technicians in construction, builder, carpenters, etc.
FABRIC	Production	Jobs connected to any process involved in the production of tangible goods and energy.	Engineers, technicians and workers in production.
COMINT	B-to-B	wholesale and business-to-business trade, both sales and purchases.	Buyers, salespersons, sales executives, etc.
GESTIO	Management	CEOs, management and administrative staff.	–
LOGIST	Transport and logistics	Both passenger and good transport	Engineers in logistics, drivers, handlers, dockers, etc.
ENTREP	Maintenance	Jobs for maintenance and repair (excluding construction)	Repair mechanics, cleaners, gardeners, etc.
DISTRI	Retail	–	Cashiers, butchers, salespersons, etc.
SERPRO	Local services	Daily life services (excluding transport, retail, education and health).	Hairdressers, cooks, real estate agents, etc.
EDUFOR	Education and formation	Jobs in primary, secondary and upper education and professional training.	Teachers, education trainers, etc.
SANSOC	Health and social work	–	Medical doctors, pharmacists, nurses, childcare, social worker, etc.
CULLO	Culture and leisure	–	Librarians, journalists, artists, sports instructors, etc.
ADMPUB	Public administration	All jobs related to public administration (excluding health and education but including security and justice).	–

Note: See https://www.insee.fr/fr/statistiques/1893116 for further documentation.

Table 2: Sample statistics

Variable	Source	unit	Mean	Std. Dev	Min.	Max.
property, plant and equipment	FARE	1000 euros	47120.87	262488.6	1.319	1.42e+07
intangible capital	FARE	1000 euros	7659.016	56802.75	-.162	2352462
capital	FARE	1000 euros	54779.89	290688.9	1.465	1.45E+07
value added	FARE	1000 euros	22856.92	86871.67	34.705	2895507
sales	FARE	1000 euros	88459.83	423847.3	1109.17	2.01E+07
output	FARE	1000 euros	77682.19	390757.2	12.79	2.00E+07
total wages paid	DADS	euros	1.07E+07	3.41E+07	1065385	1.07E+09
projected wages	DADS	euros	9663794	2.88E+07	1090598	9.27E+08
total hours worked	DADS	hours	475824.2	1214838	88166	3.75E+07
Distinct product lines (8 digit)	EAP		2.9106	4.292801	1	119
Hhi of sales of products	EAP		0.756131	0.271374	0.054292	1
Weighted volatility of sales growth	PRODCOM		0.172628	0.102032	0.024402	2.117789
Raw markup	FARE		0.037564	0.100719	-0.66804	2.067379

Table 3: Distribution of basic functions across firms in terms of hours worked

Variable	Mean	Std. Dev.	Min	Max	Median	Share of firms with function	Share in total hours worked	Share in total wage bill
Public administration	0.0%	0.4%	0.0%	12.5%	0.0%	2.3%	0.0%	0.0%
Agriculture and fishing	0.1%	1.2%	0.0%	54.2%	0.0%	7.3%	0.1%	0.1%
Construction and public works	1.1%	4.3%	0.0%	79.2%	0.0%	38.6%	0.9%	0.6%
B-2-B	6.4%	7.1%	0.0%	84.6%	4.1%	90.0%	6.3%	9.5%
R&D	5.7%	7.7%	0.0%	75.8%	3.1%	82.0%	9.3%	12.3%
Culture and leisure	0.2%	1.2%	0.0%	44.0%	0.0%	20.9%	0.2%	0.1%
Retail	1.5%	7.1%	0.0%	97.0%	0.0%	37.0%	1.4%	1.0%
Education and training	0.0%	0.8%	0.0%	58.6%	0.0%	6.9%	0.1%	0.1%
Maintenance	7.0%	8.5%	0.0%	96.5%	4.9%	94.4%	7.1%	6.7%
Production	54.9%	20.0%	0.0%	100.0%	57.5%	99.8%	51.3%	44.5%
Management	11.6%	7.3%	0.0%	94.6%	10.2%	99.3%	11.5%	13.9%
Transport and logistics	9.2%	8.5%	0.0%	87.2%	7.1%	96.8%	8.6%	7.2%
Intellectual services	1.7%	3.2%	0.0%	75.1%	0.9%	65.5%	2.4%	3.2%
Health and social work	0.2%	1.0%	0.0%	40.2%	0.0%	26.7%	0.3%	0.4%
Local services	0.3%	1.7%	0.0%	64.0%	0.0%	25.0%	0.4%	0.3%
Non-routine service	13.8%	12.2%	0.0%	84.8%	10.5%	96.5%	18.0%	25.0%
Routine service	16.3%	11.8%	0.0%	97.6%	13.8%	99.1%	15.7%	13.9%
Other functions	3.5%	9.0%	0.0%	98.2%	1.1%	86.2%	3.4%	2.7%
Non-routine service without R&D	8.1%	8.3%	0.0%	84.8%	5.7%	93.1%	8.7%	12.7%
Shares without manager positions								
Non-routine service	11.0%	10.3%	0.0%	82.0%	8.2%	96.5%	14.9%	
Routine service	14.8%	11.2%	0.0%	92.4%	12.4%	98.7%	14.2%	
Other functions	3.1%	8.5%	0.0%	97.1%	0.9%	84.9%	3.1%	
Non-routine service without R&D	5.8%	6.6%	0.0%	76.5%	4.0%	93.1%	6.5%	

The first five columns reports statistics on the within-firm shares of functions. The share is defined by the number of hours worked in a given function divided by the total number of hours. The sixth column reports the share of firms that has a given function in-house. The last column gives the share of the hours worked in the function in total hours worked for the entire sample.

Table 4: Summary of routinness measures for different functions

Variable	Observations	Mean	Std. Dev	Min	Max
Construction and public works	2589	-0.10	0.29	-1.50	0.46
B-2-B	6042	-0.67	0.11	-0.82	0.05
R&D	5436	-0.59	0.15	-0.82	-0.40
Culture and leisure	1303	-0.19	0.30	-0.73	1.24
Retail	2483	0.25	0.83	-1.52	1.41
Maintenance	6330	0.07	0.32	-1.00	1.59
Production	6699	0.34	0.32	-0.82	2.24
Management	6668	0.68	0.66	-0.75	2.24
Transport and logistics	6502	0.24	0.92	-1.50	2.24
Intellectual services	4396	-0.58	0.16	-1.00	-0.33
Health and social work	1725	-0.49	0.22	-1.00	-0.33
Local services	1681	-0.35	0.28	-0.75	0.03

Table 5: Non-routine service functions in 2015 and firm innovation 2016-2018

	Innovation index	Product innovation	Intellectual property development	Process innovation	Innovation index	Product innovation	Intellectual property development	Process innovation
Non-routine lateral	7.433***	1.700***	2.575***	3.159***				
	(0.647)	(0.148)	(0.334)	(0.420)				
R&D					6.969***	1.355***	2.895***	2.719***
					(1.418)	(0.221)	(0.698)	(0.697)
Other non-routine lateral					7.782***	1.959***	2.333***	3.489***
					(1.026)	(0.289)	(0.458)	(0.726)
Ln of hours worked	1.103***	0.194***	0.383***	0.526***	1.108***	0.198***	0.380***	0.531***
	(0.121)	(0.017)	(0.047)	(0.075)	(0.117)	(0.017)	(0.046)	(0.072)
Constant	-11.630***	-1.872***	-4.462***	-5.297***	-11.713***	-1.934***	-4.405***	-5.375***
	(1.589)	(0.230)	(0.620)	(0.976)	(1.532)	(0.219)	(0.607)	(0.929)
N	1571	1571	1571	1571	1571	1571	1571	1571
N_clust	24	24	24	24	24	24	24	24
R^2	0.2555	0.2022	0.2183	0.1335	0.2557	0.2037	0.2187	0.1339
industry FE	Y	Y	Y	Y	Y	Y	Y	Y

Innovation measures from the 2018 of the French version of the Community Innovation Survey (CIS). We grouped yes(=1)/no answers to 15 questions into an overall index (overall sum) and distinct groups. *Product innovation* is the sum of indicator variables whether the firm conducted product or service innovation (maximum value = 2). *Intellectual property development* is the sum of indicator variables on different aspects of intellectual property – patents, trade secrets etc. (maximum value = 6). *Process innovation* is the sum of indicator variables whether the firm had different types of process innovation (maximum value = 7).

Sample trimmed at 0.5% at each tail of estimated TFP for all firms. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 6: Intangibles and functions

	intangible capital / hour worked	intangible capital / hour worked	intangible capital / hour worked	intangible capital / (ppe + intangible capital)	intangible capital / total capital
Management	4.492***	4.423***	4.761***	3.754***	3.168***
	(0.510)	(0.478)	(0.528)	(0.246)	(0.252)
Non-routine service	3.851***		3.859***		
	(0.363)		(0.316)		
Routine service	-0.103	-0.114	-0.116	-0.848**	-0.822**
	(0.474)	(0.472)	(0.474)	(0.327)	(0.304)
R&D		3.629***		2.722***	2.436***
		(0.384)		(0.450)	(0.453)
Other non-routine service		4.011***		4.158***	3.831***
		(0.452)		(0.386)	(0.391)
Other functions	1.979***	1.967***	1.929***	2.674***	2.621***
	(0.492)	(0.494)	(0.481)	(0.321)	(0.325)
Ln of hours worked	0.264***	0.267***	0.258***	-0.004	-0.045
	(0.026)	(0.025)	(0.025)	(0.038)	(0.040)
CONSTANT	-10.367***	-10.395***	-10.329***	-4.009***	-3.493***
	(0.387)	(0.385)	(0.375)	(0.514)	(0.538)
N	6565	6565	6305	6565	6565
clusters	24	24	24	24	24
R^2	0.2254	0.2255	0.2310	0.2451	0.2148
trim	1%	1%	5%	1%	1%

In this table, we regress different measures of intensity of intangible capital on the shares of functions within the firm. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 7: Non-routine service functions, product complexity and product quality

	Sales growth volatility	Product complexity (PCI)	PCI*(PCI >0)	Rauch product differentiation	Rauch product differentiation <1	Log of number of products	Product sales concentration	Product quality
Non-routine lateral	0.098**	1.761***	1.341***	0.344**	0.533**	0.761***	-0.241***	
	(0.036)	(0.427)	(0.346)	(0.152)	(0.209)	(0.176)	(0.039)	
Ln of hours worked	0.005	-0.027	-0.023	-0.021**	-0.029	0.103***	-0.027***	
	(0.003)	(0.030)	(0.021)	(0.009)	(0.026)	(0.015)	(0.005)	
CONSTANT	0.083**	0.468	0.629**	1.120***	1.029**	-1.017***	1.248***	
	(0.037)	(0.379)	(0.252)	(0.085)	(0.413)	(0.182)	(0.056)	
N	6233	5790	5790	5758	1957	6235	6233	
clusters	23	23	23	23	23	23	23	
R^2	0.0177	0.0830	0.0969	0.0415	0.0374	0.0736	0.0421	
industry FE	N	N	N	N	N	N	N	
trim	1%	1%	1%	1%	1%	1%	1%	
Non-routine lateral	0.033**	0.102	0.170*	0.066*	0.218***	0.532***	-0.170***	2.150***
	(0.015)	(0.203)	(0.087)	(0.035)	(0.066)	(0.145)	(0.043)	(0.195)
Ln of hours worked	0.005*	0.009	0.003	-0.006**	0.009*	0.104***	-0.030***	-0.140***
	(0.003)	(0.012)	(0.009)	(0.002)	(0.005)	(0.014)	(0.004)	(0.043)
CONSTANT	0.053	-0.480***	-0.003	0.726***	0.388***	-0.998***	1.292***	-0.248
	(0.034)	(0.146)	(0.112)	(0.028)	(0.062)	(0.172)	(0.052)	(0.529)
N	6233	5790	5790	5758	1957	6235	6233	9627
clusters	23	23	23	23	23	23	23	22
R^2	0.1436	0.5425	0.5005	0.4143	0.5501	0.1058	0.0726	0.8697
industry FE	Y	Y	Y	Y	Y	Y	Y	product FE
trim	1%	1%	1%	1%	1%	1%	1%	1%

Top panel: correlations without 2-digit industry level effects. Lower panel: with 2-digit industry level effects except for product quality where 8-digit product effects are included.
Sales growth volatility is calculated as the standard deviation of (log)growth of 8-digit products sales in the EU 27 (EU 28 without France) over the period 2005-2014 weighted by firm sales. *Product complexity (PCI)* is the firm-level sales-weighted (at the 8-digit product level) 2015 measure of product complexity from the Harvard Atlas of Economic Complexity. *PCI*(PCI >0)'* is the PCI index for PCI>0 and zero otherwise. *Rauch product differentiation* is the sales-weighted (at the 8-digit product level) measure of product differentiation from Rauch (1999) where homogenous goods are coded as "0", reference priced products are coded as "0.5", and differentiated as "1". We adopt the "conservative" classification with results similar while using the "liberal" one. *Rauch product differentiation < 1* retains firms that produce some non-differentiated products. *Log of number of products* measures the number of distinct 4-digit product lines that a firm produces. *Product sales concentration* is calculated as the HHI of sales in different 4-digit products. *Product quality* is calculated as the logarithm of the unit value (sales / quantity) at the 8-digit PRODCOM product level with firms using VF1 mode of production (the firm produces by itself and commercializes the product). Observations are thus at the firm-product level and product-level fixed effects are included in the specification.
Data for the calculation of firm-level measures of product count and concentration and product-level unit values from EAP 2015. Eurostat PRODCOM database of sold production for volatility of sales growth at the 8-digit NACE Rev. 2 code. Data for two sectors – NACE 10 and 11 from EAP survey "PRODCOM: Production commercialisée des industries agricoles alimentaires". Product-level volatility included for products that had at least 5 years of data in the period 2010-2015, but calculated over entire range 2005-2014. For 8-digit lines with less data, the volatility of sold production at the 2-digit NACE assigned.
Sample trimmed at 0.5% at each tail of estimated TFP. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 8: Changes in non-routine service functions, product complexity and quality

Change in:	Sales growth volatility	Product complexity (PCI)	PCI*(PCI > 0 in 2005)	Rauch product differentiation	Rauch product differentiation <1 in 2005	Log of number of products	Product sales concentration	Quality
Change in non-routine lateral share 2005-2010	-0.000	0.038**	0.040**	-0.007	-0.021	0.107***	-0.040**	-0.213
	(0.007)	(0.017)	(0.018)	(0.005)	(0.016)	(0.025)	(0.015)	(0.235)
Non-routine lateral share in 2005	0.001	0.048**	0.048**	-0.004	-0.011	0.110**	-0.038	-0.111
	(0.008)	(0.021)	(0.019)	(0.003)	(0.013)	(0.051)	(0.023)	(0.166)
Change of ln of hours worked 2005-2010	-0.001	0.000	-0.003	-0.000	0.004	0.020	-0.009*	0.033
	(0.001)	(0.004)	(0.004)	(0.001)	(0.003)	(0.012)	(0.004)	(0.024)
Ln of hours worked in 2005	0.000	-0.002	-0.002	-0.001***	-0.004***	-0.001	0.003*	0.019
	(0.000)	(0.002)	(0.002)	(0.000)	(0.001)	(0.004)	(0.002)	(0.015)
CONSTANT	-0.012**	0.015	0.036*	0.016***	0.062***	0.062	-0.044**	0.078
	(0.005)	(0.022)	(0.018)	(0.004)	(0.013)	(0.044)	(0.018)	(0.174)
N	12283	10232	10232	10980	3063	12284	12283	9913
clusters	21	21	21	21	21	21	21	21
R^2	0.0081	0.0095	0.0074	0.0129	0.0345	0.0068	0.0069	0.7294
trim	1%	1%	1%	1%	1%	1%	1%	1%

Correlations with 2-digit industry level effects.
Firms over 10 employees in 2005, 2010 and 2015.
Sales growth volatility is the standard deviation of (log)growth of 8-digit products sales in the EU 27 (EU 28 without France) over the period 2005-2014 weighted by firm sales. *Product complexity (PCI)* is the firm-level sales-weighted (at the 8-digit product level) measure of product complexity from the Harvard Atlas of Economic Complexity from 2010 and 2015. Firms trimmed at 2.5% of PCI values at each tail to counter the problem of mean reversion. *PCI*(PCI >0)* is the PCI index for PCI>0 and zero otherwise. *Rauch product differentiation* is the sales-weighted (at the 8-digit product level) measure of product differentiation from Rauch (1999) where homogenous goods are coded as "0", reference priced products are coded as "0.5", and differentiated as "1". We adopt the "conservative" classification with results similar while using the "liberal" one. *Rauch product differentiation < 1* retains firms that produce some non-differentiated products. *Log of number of products* measures the number of distinct 4-digit product lines that a firm produces. *Product sales concentration* is calculated as the HHI of sales in different 4-digit products. *Raw markup* is total revenue/total cost at the firm level.
Data for the calculation of firm-level measures of product count and concentration from EAP 2010 and 2015 and Eurostat PRODCOM database of sold production (for volatility of sales growth at the 8-digit NACE Rev. 2 code). Data for two sectors – NACE 10 and 11 unavailable for 2010 and hence not included. Product-level volatility included for products that had at least 5 years of data in the period 2010-2015, but calculated over entire range 2005-2014. For 8-digit lines with less data, the volatility of sold production at the 2-digit NACE assigned.
Sample trimmed at 0.5% at each tail of estimated TFP. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 9: Productivity and functions

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Management		0.524***	0.570***	0.437***	0.536***	0.586***	0.446***	0.529***	0.559***	0.407***
		(0.058)	(0.069)	(0.067)	(0.063)	(0.070)	(0.066)	(0.083)	(0.087)	(0.079)
Non-routine lateral		0.155***	0.208***	0.220***				0.148***	0.202***	0.203***
		(0.036)	(0.051)	(0.038)				(0.040)	(0.055)	(0.046)
	B-2-B				0.118	0.177**	0.188***			
					(0.078)	(0.079)	(0.054)			
	R&D				0.167***	0.236**	0.215***			
					(0.060)	(0.089)	(0.062)			
	Intellectual services				0.209	0.262	0.403***			
					(0.177)	(0.193)	(0.142)			
Routine lateral		0.031	0.046	0.047				0.027	0.042	0.036
		(0.022)	(0.036)	(0.031)				(0.025)	(0.036)	(0.031)
	Maintenance				0.050	0.120**	0.131***			
					(0.039)	(0.044)	(0.039)			
	Transport and logistics				0.014	-0.009	-0.014			
					(0.034)	(0.041)	(0.039)			
Managers (share)								0.225**	0.300***	0.332***
								(0.106)	(0.093)	(0.099)
Other functions		0.094**	0.101**	0.112**	0.095**	0.104**	0.112**	0.095**	0.099*	0.111**
		(0.041)	(0.048)	(0.047)	(0.039)	(0.046)	(0.043)	(0.043)	(0.050)	(0.047)
CONSTANT		-0.084***	-0.084***	-0.065***	-0.085***	-0.083***	-0.064***	-0.082***	-0.084***	-0.065***
		(0.010)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)
N		6648	6648	6380	6648	6648	6380	6648	6648	6380
N_clust		24	24	24	24	24	24	24	24	24
R^2		0.0183	0.0223	0.0267	0.0185	0.0228	0.0278	0.0176	0.0215	0.0264
industry FE		N	Y	Y	N	Y	Y	N	Y	Y
trim		1%	1%	5%	1%	1%	5%	1%	1%	5%

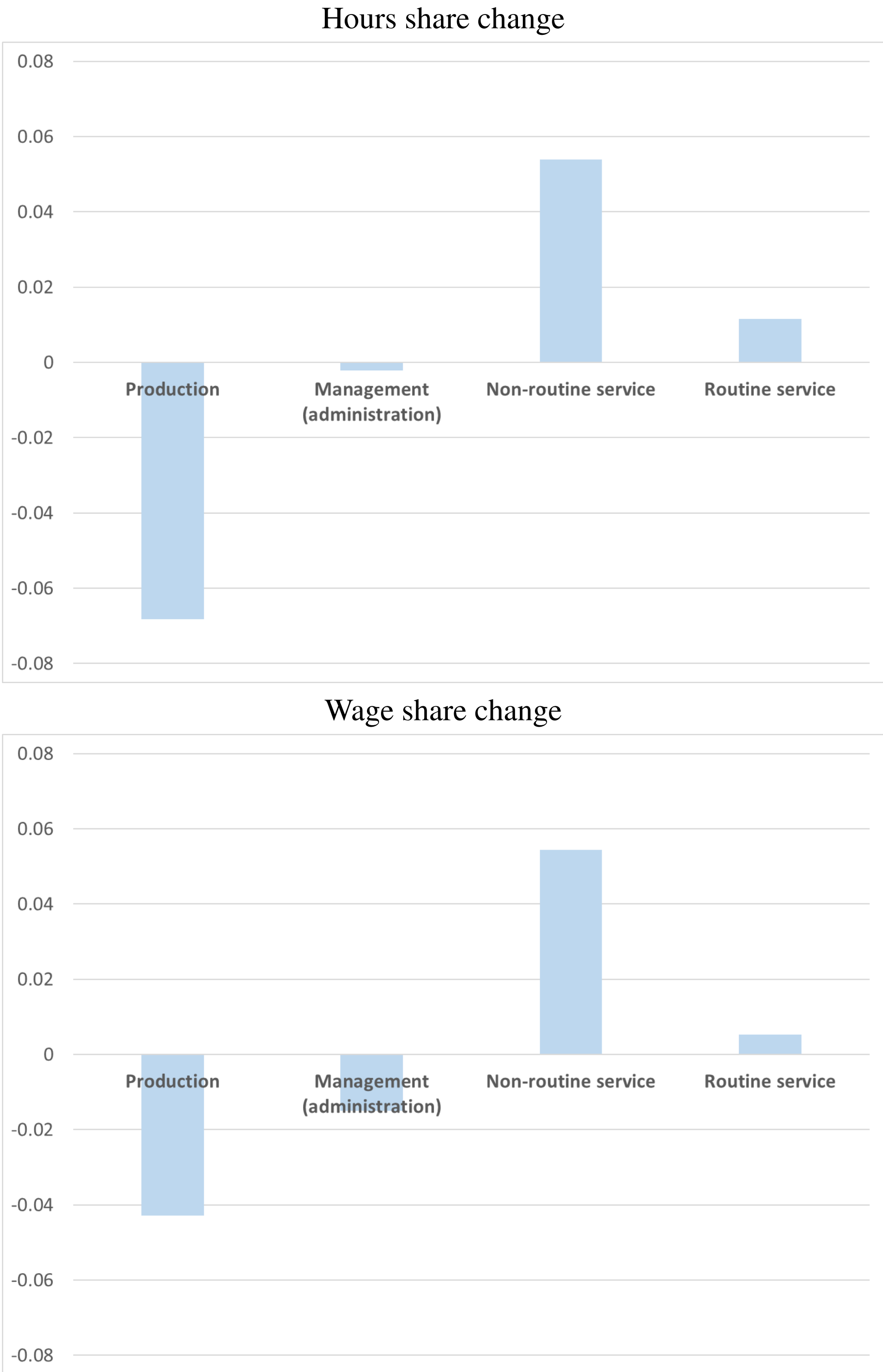
In this table, we regress productivity on the shares of functions within the firm. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 10: Non-routine service functions and profitability

	Profitability 2015	Profitability 2015	Profitability 2015	Change in profitability 2010-2015	Log of profitability 2015
Non-routine lateral	0.059**	0.048**	0.045***		0.645***
	(0.023)	(0.020)	(0.016)		(0.164)
Ln of hours worked	0.003	0.002	0.002		0.023
	(0.002)	(0.002)	(0.002)		(0.020)
Change in non-routine lateral share 2005-2010				0.018*	
				(0.009)	
Non-routine lateral share in 2005				0.016	
				(0.010)	
Change of ln of hours worked 2005-2010				-0.012***	
				(0.002)	
Ln of hours worked in 2005				-0.001	
				(0.001)	
Constant	-0.014	0.011	0.010	0.000	-3.732***
	(0.027)	(0.021)	(0.021)	(0.010)	(0.246)
N	6648	6648	6413	13637	4952
clusters	24	24	24	21	24
R^2	0.0082	0.0463	0.0455	0.0058	0.0430
industry FE	N	Y	Y	Y	Y
trim	1%	1%	5%	1%	1%

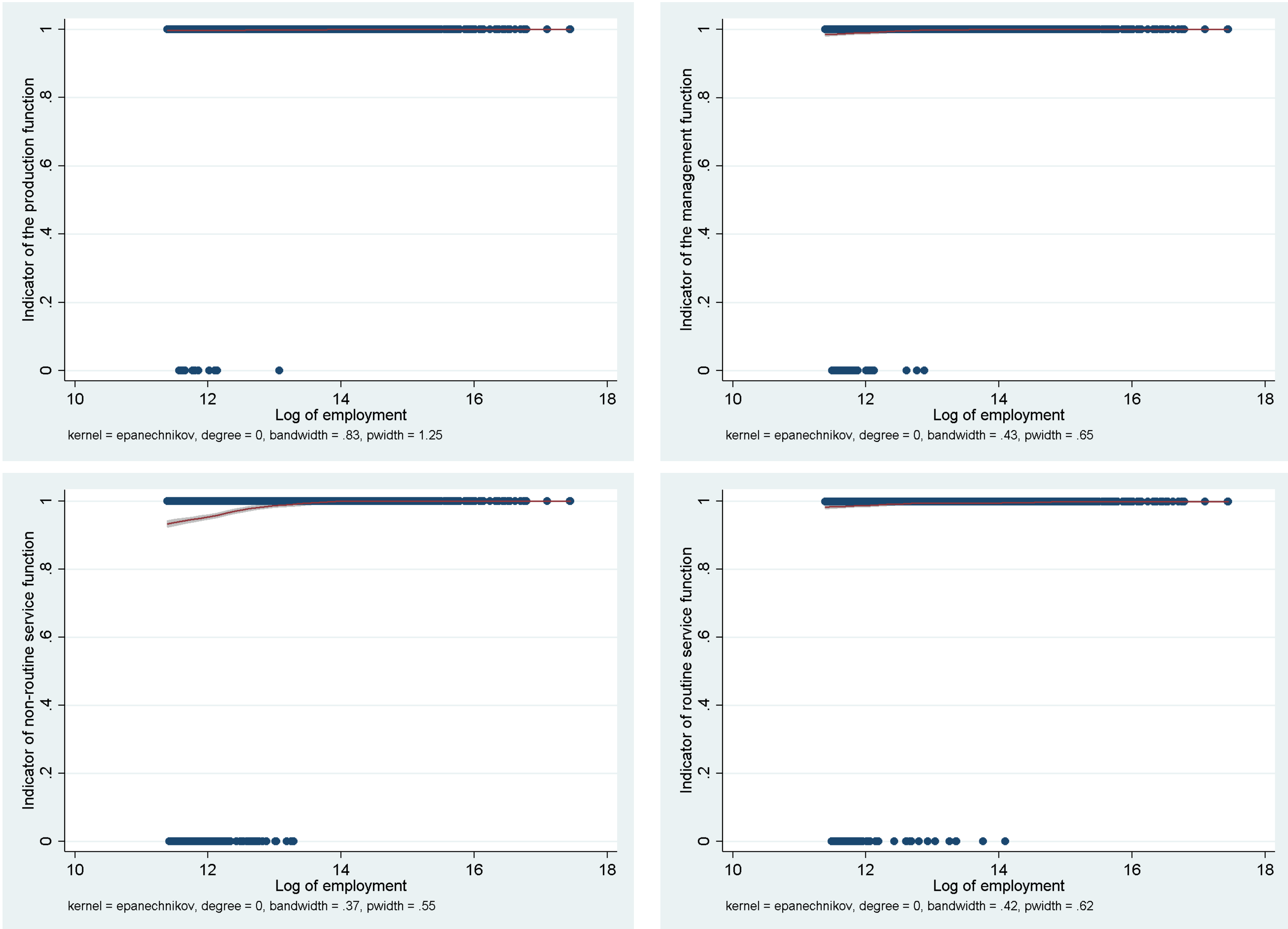
Profitability is total revenue/total cost at the firm level. Sample trimmed at 0.5% at each tail of estimated TFP. For the long-differenced regression, firms over 10 employees in 2005, 2010 and 2015. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Figure 5: Changes in shares of hours or wages in different functions 1999-2015.



The upper panel shows the change in the aggregate share of hours worked while the lower in wage bill shares in production, management ("gestion" or administration in French), non-routine and routine service functions. Sample: firms in manufacturing with employment >50 workers in 1999 or 2015.

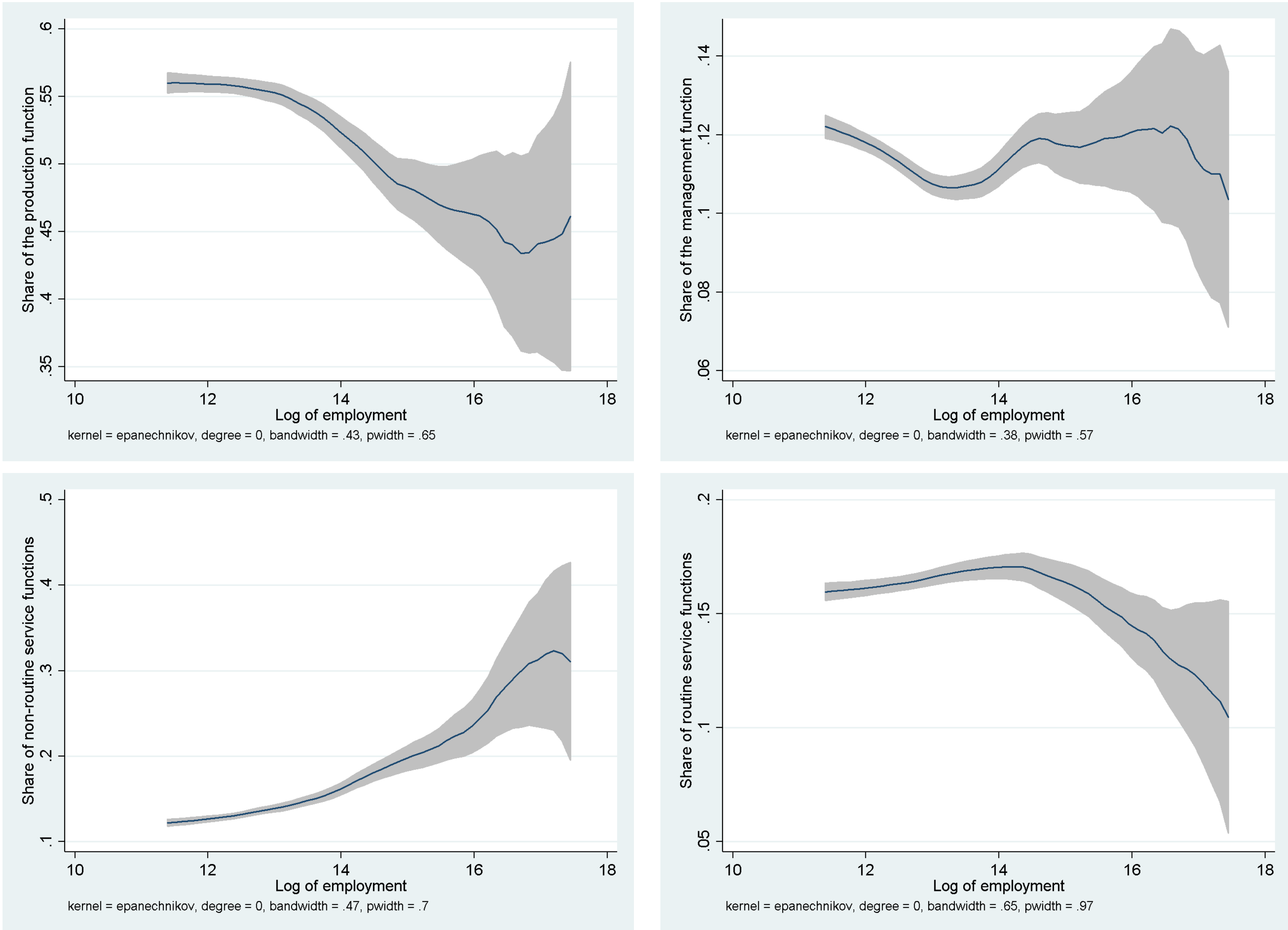
Figure 6: Firm employment and the presence of different functions within firms.



Figures show a kernel estimate of the relationship between firm employment (measured in full position equivalent hours) and an indicator whether a function is represented within a firm. 95% confidence intervals around the estimate are shown. Sample trimmed at the top and bottom 2.5% estimate of TFP.

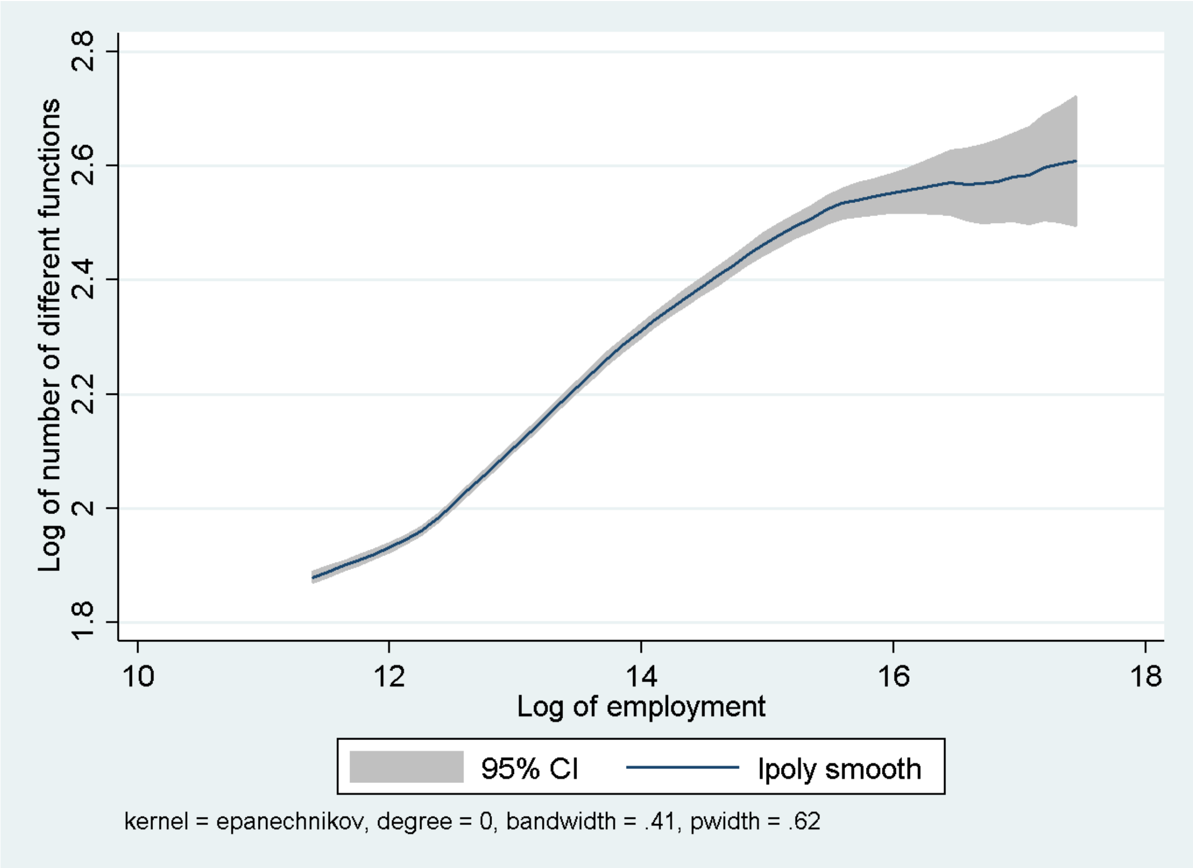
Larger firms have more functions present within firms and size is an important predictor of whether a function is within a firm or not. See also the top panel of Figure 8.

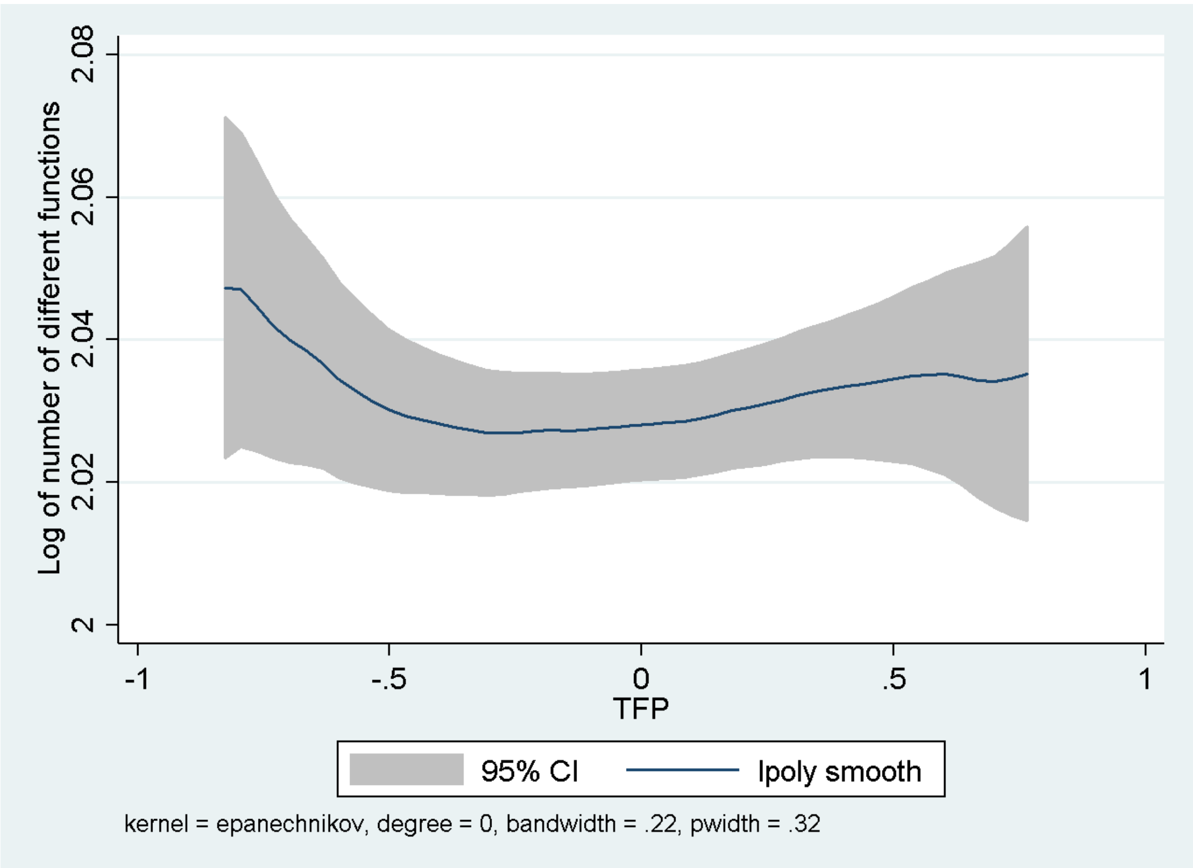
Figure 7: Firm employment and share of different functions.



Figures show a kernel estimate of the relationship between firm employment (measured in full position equivalent hours) and the employment share within different functions. The values may not sum to one as "other functions" category is not depicted. 95% confidence intervals around the estimate are shown. Sample trimmed at the top and bottom 2.5% estimate of TFP.

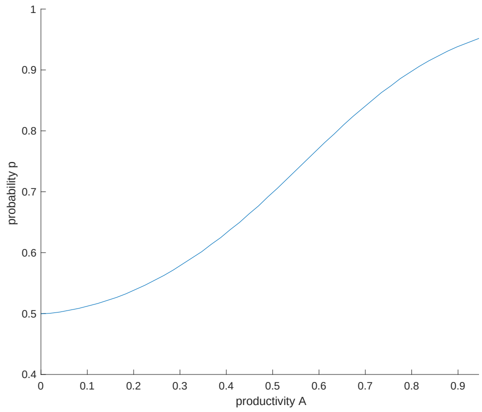
Figure 8: Number of functions vs. employment and TFP.

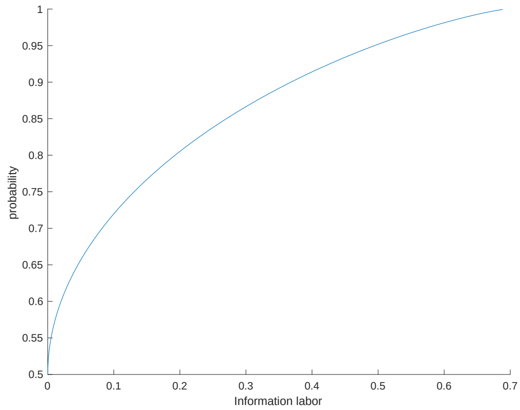




Figures show a kernel estimate of the relationship between respectively firm employment or TFP and the logarithm of the number of different functions. 95% confidence intervals around the estimate are shown. Sample trimmed at the top and bottom 2.5% estimate of TFP.

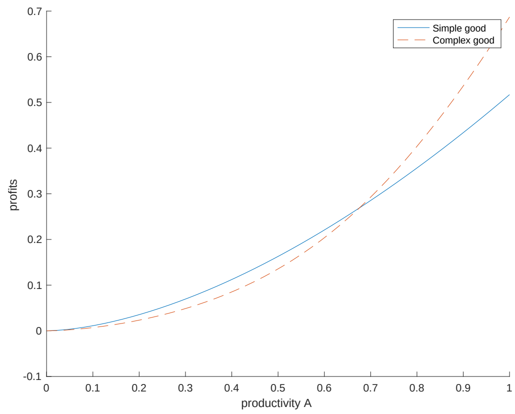
Figure 9: Probability p as a function of productivity and knowledge-generating labor





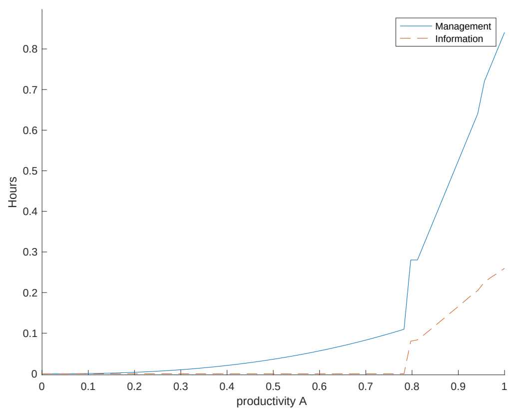
In this Figure, we plot the optimal probability p as a function of productivity (left panel) and as function of knowledge-generating labor l_K (right panel). The calibrated parameters are $N = 2$, $\alpha = .4$, $\pi(complex) = 1.5$, $\pi(simple) = 1$ and $w = .5$.

Figure 10: Optimal profits as a function of productivity.



In this Figure, we plot profits as a function of productivity. The calibrated parameters are $N = 2$, $\alpha = .4$, $\pi(complex) = 1.5$, $\pi(simple) = 1$ and $w = .5$.

Figure 11: Management and knowledge-generating labor as a function of productivity.



In this Figure, we plot labor in management and knowledge generation as a function of productivity. The calibrated parameters are $N = 2$, $\alpha = .4$, $\beta = .2$, $\gamma = \delta = .5$, $\pi(complex) = 1.5$, $\pi(simple) = 1$ and $w = .5$.

A Proofs

A.1 Proof of Lemma 1.

Let us consider a probability distribution such that k' has now probability $1/N+\epsilon$ and all the other inputs have probability $1/N-\epsilon/(N-1)$, with ϵ arbitrarily small. The amount of information of such distribution is such that:

$$H(\mathcal{I}|\mathcal{I}_0) = (1/N+\epsilon)\log\frac{1/N+\epsilon}{1/N} + (N-1)(1/N-\epsilon/(N-1))\log\frac{1/N-\epsilon/(N-1)}{1/N}$$

When ϵ is small enough:

$$H(\mathcal{I}|\mathcal{I}_0) = N\epsilon^2/2 + N\epsilon^2/(2(N-1)) = \Gamma(N)\epsilon^2$$

The problem solved locally by the firm is then:

$$\max_{\epsilon}\max_{l}\left\{\pi_{complex}A\left(\frac{1}{N}+\epsilon\right)(l)^{\alpha} - wl_2 - w1/2\Gamma(N)\epsilon^2\right\}$$

We have l satisfying:

$$\alpha\pi_{complex}A\left(\frac{1}{N}+\epsilon\right)l^{\alpha-1} = w.$$

The optimal ϵ then satisfies:

$$\begin{aligned}\pi_{complex}A\left(l^{\alpha} + \left(\frac{1}{N}+\epsilon\right)\alpha l^{\alpha-1}\frac{\partial l}{\partial\epsilon}\right) &= w\Gamma(N)\epsilon\\ \pi_{complex}Al^{\alpha} + w\frac{\partial l}{\partial\epsilon} &= w\Gamma(N)\epsilon\end{aligned}$$

As the left hand term is strictly positive, we have $\epsilon > 0$.

Let us turn to the second part of the Lemma. First, let us consider a probability distribution such that the probability of k' is p. The probability distribution for all the other $i\in\Omega$ that minimizes

$$p\log p - \log N + \sum_{i\neq k'} p\log p_i$$

is the uniform distribution. As the distribution on all other $i\in\Omega$ is payoff irrelevant, it is then without loss of generality to consider the $\mathcal{I}(p)$.

Second, let us note that any distribution such that the probability of k' is $p\geq 1/N$ and other probabilities are $f_i(p)$, with $\sum_i f_i(p) = 1-p$ and $f_i(p)$ decreasing, we have that $H(\mathcal{I}|\mathcal{I}_0)$ is increasing and convex in p. Indeed, the derivative of $H(\mathcal{I}|\mathcal{I}_0)$ with such distribution is:

$$\log p + \sum_i \frac{\partial f_i}{\partial p}\log f_i(p) > 0,$$

and the second order derivative is:

$$\frac{1}{p} + \sum_i\left(\frac{\partial f_i}{\partial p}\right)^2\frac{1}{f_i(p)} > 0$$

The problem solved by firms is then:

$$\max_p \max_l \{\pi_{complex} A p l^{\alpha} - wl - wC(p)\}$$

with $H(\mathcal{I}|\mathcal{I}_0) = C(p)$. Clearly, p is increasing in A and, as $C(p)$ is increasing and convex in p, we have $l_{I,i}$ increasing in A.

Finally, as $H(\mathcal{I}(p)|\mathcal{I}_0)$ is increasing in p and the information constraint binds, we have p is increasing in l_K. As $H(\mathcal{I}(p)|\mathcal{I}_0)$ is strictly increasing and convex, we also have that the inverse function, that is p as a function of l_K is concave, which proves (ii).

A.2 Proof of Proposition 2.

First of all, note that ultimately $l_{I,i} = \log N$ and the firm operates under perfect information. As $\pi(complex) > \pi(simple)$, this implies that when $A \geq \bar{A}$, the production of the complex good dominates the production of the simple good.

Using the envelope theorem, the derivative of profits with respect to A are $\pi_{complex} p l^{\theta}_{complex}$ for the profits associated with producing the complex good and $\pi_{simple} l^{\theta}_{simple}$ for the profits associated with producing the simple good. Using the first order conditions for labor, that are:

$$A\pi_{complex} p \theta l^{\theta}_{complex} = wl_{complex} \text{ and } A\pi_{simple} \theta l^{\theta}_{simple} = wl_{simple},$$

we have $l_{complex} = (Ap\pi_{complex}/w)^{1/(1-\theta)}$ and $l_{simple} = (A\pi_{simple}/w)^{1/(1-\theta)}$. As a result, the difference between these two derivatives are:

$$\pi_{complex} p l^{\theta}_{complex} - \pi_{simple} l^{\theta}_{simple} = \frac{w}{A}(l_{complex} - l_{simple}) = \left(\frac{A}{w}\right)^{\frac{\theta}{1-\theta}} \left([\pi_{complex} p]^{\frac{1}{1-\theta}} - [\pi_{simple}]^{\frac{1}{1-\theta}}\right).$$

As p is continuously increasing from $1/N$ to 1 when A increases from 0 to ∞, the difference is either always positive when $\pi_{complex}/N > \pi_{simple}$ or negative and then positive when $\pi_{complex}/N < \pi_{simple}$. In the end, as with $A = 0$, there is no production of both goods, in the first case, $\bar{A} = 0$ and, in the second case, $\bar{A} > 0$.

A.3 Proof of Proposition 5.

Let us consider firm 1 with a productivity A_1 so that p_1 is arbitrarily close to 1. In this case, knowledge generation inside is equivalent to paying a fixed cost $wl_{in} - wl_{out}/(l_{out} + l^2_{out})l_{out} - \kappa$. On the other hand, generating knowledge inside leads to a higher price. So there exists a sufficiently high productivity A^1 so that firm 1 prefers to enjoy a higher price and pay the fixed cost.

B Data sources for the variables used

Data sources:

- **Function classification.** 4-digit occupations based on the French PCS classification are classified into 15 different functions by the French statistical institute INSEE. The current methodology and classifications are accessible at https://www.insee.fr/fr/statistiques/1893116.

- **Individual hours worked (by occupation) and wages** at the firm (an entity with a "SIREN" identifier) level: French administrative social security data "DADS-Postes" for 1999, 2005, 2010 and 2015. Data accessible after administrative permissions through the secure data hub modems via CASD (https://www.casd.eu/en/).

- **Balance sheet data** (tangible and intangible capital, sales, materials purchased, markups etc.): FICUS-FARE 2010 and 2015. Data accessible after administrative permissions through the secure data hub modems via CASD (https://www.casd.eu/en/).

- **Routineness measures.** Taken from Goos et al. (2014) (originally from Autor et al. (2003)) translated into two-digit PCS and then into individual functions. To translate these indices and obtain exposure to automation we merge the exposure classifications of Goos et al. (2014) (that include RTI in their dataset) based on 2-digit ISCO occupation classification into the 2015 4-digit PCS classification. Function assignment is available at the 4-digit CS level, so we obtain routineness measures for functions by weighting hours worked in occupations within the function different RTI indexes.

- **Data at the firm-product level.** 8-digit product level data (sales) at the SIREN level to calculate product scope, concentration and product quality (unit values) come from two datasets: (i) the EAP "PRODCOM: Production commercialisée des industries agricoles alimentaires" for food and beverage manufacturing (NACE 10 and 11) from 2015 (this survey program started in 2011) and (ii) Enquête Annuelle de Production (EAP) of 2010 and 2015 for the remainder of manufacturing industries. Data accessible after administrative permissions through the secure data hub modems via CASD (https://www.casd.eu/en/). Both data sets are used by the INSEE to process French data for the Eurostat PRODCOM data sets.

- **8-digit product level** data from pan-EU (excluding France) PRODCOM data from EUROSTAT for years 2005-2015 to calculate sales growth variability at the product level. Data accessible from https://ec.europa.eu/eurostat/web/prodcom.

- **Product complexity indexes** from the Atlas of Economic Complexity at Harvard University accessible from https://atlas.cid.harvard.edu/rankings/product. Firm-level (SIREN) product complexity calculated matching firm-product data from the INSEE's EAP surveys (PRODCOM data, described above) with HS4 product code complexity indexes.

- **Rauch product differentiation classification** from Rauch (1999). Accessible at https://econweb.ucsd.edu/~jrauch/rauch_classification.html.

- **Innovation indices.** Community Innovation Survey ran by the INSEE, 2010-2018 for product innovation measures at the firm (SIREN) level. Data accessible after administrative permissions through the secure data hub modems via CASD (https://www.casd.eu/en/).

C Additional figures and tables

Table C.1: Productivity and detailed functions

	2-digit-NACE FE	Multi-plant firms	2-digit-NACE FE
Management: CEOs	1.353***	1.635**	1.560***
	(0.420)	(0.708)	(0.382)
Management: cadres	0.590**	0.533	0.501**
	(0.269)	(0.423)	(0.241)
Managers: mid-level	0.470**	0.544*	0.312*
	(0.208)	(0.286)	(0.175)
Managers: office workers	0.491***	0.498***	0.322***
	(0.123)	(0.125)	(0.111)
B-2-B: purchases	0.895**	0.804*	1.014***
	(0.336)	(0.392)	(0.270)
B-2-B: sales	0.093	0.237*	0.105*
	(0.088)	(0.124)	(0.056)
R&D	0.145*	0.154	0.108**
	(0.079)	(0.105)	(0.050)
Intellectual services: IT	0.277	0.458	0.289
	(0.191)	(0.315)	(0.180)
Intellectual services: legal services	5.105	5.892	6.241**
	(4.454)	(5.436)	(2.824)
Intellectual services: marketing	4.336***	4.689**	3.534***
	(1.034)	(1.671)	(1.246)
Intellectual services: economic consulting	-0.879	1.152	0.747
	(0.920)	(0.690)	(0.732)
Intellectual services: other	0.008	-0.335	0.049
	(0.293)	(0.528)	(0.196)
Maintenance: cadres	0.663***	0.770**	0.629***
	(0.147)	(0.326)	(0.145)
Maintenance: technicians	0.057	0.052	0.060
	(0.053)	(0.103)	(0.052)
Maintenance: lowest level	0.103	0.121*	0.113*
	(0.061)	(0.061)	(0.055)
Transport and logistics: cadres	0.286	0.783	0.809*
	(0.580)	(0.862)	(0.447)
Transport and logistics: mid-level	0.205	0.397	0.070
	(0.301)	(0.340)	(0.274)
Transport and logistics: lowest-level	-0.011	-0.042	-0.018
	(0.042)	(0.069)	(0.039)
Production: cadres	0.428***	0.370	0.388***
	(0.125)	(0.225)	(0.105)
Production: technicians and foremen	-0.041	-0.098	-0.012
	(0.047)	(0.093)	(0.035)
Other functions	0.103**	0.145***	0.116**
	(0.047)	(0.030)	(0.044)
CONSTANT	-0.093***	-0.138***	-0.075***
	(0.014)	(0.019)	(0.014)
N	6648	2917	6380
clusters	24	24	24
R^2	0.0306	0.0494	0.0385
trim	1%	1%	5%

In this table, we regress productivity on the share of functions within the firm, when splitting functions depending on the level of skills of workers. [?] Sample trimmed at 0.5% at each tail of estimated TFP. ***, **, and * denote statistical significance at the $1\ \%$, $5\ \%$, and $10\ \%$ levels.

Table C.2: Summary statistics on detailed functions

Variable	Mean	Std. Dev.	Min	Max	Median	Share of firms with function	Share in total hours worked
Public administration	0.0%	0.4%	0.0%	12.5%	0.0%	2.3%	0.0%
Agriculture and fishing	0.1%	1.2%	0.0%	54.2%	0.0%	7.3%	0.1%
Construction and public works	1.1%	4.3%	0.0%	79.2%	0.0%	38.6%	0.9%
B-2-B	6.4%	7.1%	0.0%	84.6%	4.1%	90.0%	6.3%
R&D	5.7%	7.7%	0.0%	75.8%	3.1%	82.0%	9.3%
Culture and leisure	0.2%	1.2%	0.0%	44.0%	0.0%	20.9%	0.2%
Retail	1.5%	7.1%	0.0%	97.0%	0.0%	37.0%	1.4%
Education and training	0.0%	0.8%	0.0%	58.6%	0.0%	6.9%	0.1%
Maintenance	7.0%	8.5%	0.0%	96.5%	4.9%	94.4%	7.1%
Production	54.9%	20.0%	0.0%	100.0%	57.5%	99.8%	51.3%
Management	11.6%	7.3%	0.0%	94.6%	10.2%	99.3%	11.5%
Transport and logistics	9.2%	8.5%	0.0%	87.2%	7.1%	96.8%	8.6%
Intellectual services	1.7%	3.2%	0.0%	75.1%	0.9%	65.5%	2.4%
Health and social work	0.2%	1.0%	0.0%	40.2%	0.0%	26.7%	0.3%
Local services	0.3%	1.7%	0.0%	64.0%	0.0%	25.0%	0.4%
B-2-B: sales	5.3%	6.8%	0.0%	84.6%	3.0%	87.1%	5.1%
B-2-B: purchases	1.1%	1.5%	0.0%	46.6%	0.7%	62.6%	1.2%
Maintenance: cadres	0.5%	1.5%	0.0%	82.2%	0.0%	43.4%	0.7%
Maintenance: technicians	2.6%	5.3%	0.0%	83.3%	1.4%	71.7%	3.0%
Maintenance: low-skilled	3.9%	6.0%	0.0%	86.3%	2.1%	81.5%	3.4%
Management: CEOs	0.6%	0.9%	0.0%	11.5%	0.0%	45.8%	0.3%
Management: cadres	2.3%	2.6%	0.0%	47.6%	1.7%	83.3%	3.2%
Management: middle managers	1.7%	2.2%	0.0%	25.2%	1.1%	69.2%	2.1%
Management: office workers	7.1%	5.7%	0.0%	92.2%	5.8%	97.6%	5.9%
Intellectual services: IT high skill	0.4%	1.6%	0.0%	55.3%	0.0%	31.8%	0.6%
Intellectual services: IT medium skill	0.4%	1.2%	0.0%	33.8%	0.0%	34.9%	0.5%
Intellectual services: legal services	0.0%	0.2%	0.0%	5.1%	0.0%	8.1%	0.1%
Intellectual services: economic consulting	0.2%	0.8%	0.0%	34.2%	0.0%	18.2%	0.3%
Intellectual services: marketing and communication	0.1%	0.4%	0.0%	7.7%	0.0%	15.7%	0.2%
Intellectual services: other	0.6%	2.0%	0.0%	74.2%	0.0%	48.6%	0.7%

Table C.3: Productivity and functions: TFP estimated with intangible capital

	2-digit- NACE FE	Multi-plant firms	2-digit- NACE FE
Management	0.514***	0.518***	0.364***
	(0.071)	(0.149)	(0.062)
Non-routine service	0.203***	0.218***	0.210***
	(0.049)	(0.071)	(0.039)
Routine service	0.060	0.110*	0.051
	(0.039)	(0.062)	(0.036)
Other functions	0.040	0.149	0.067
	(0.043)	(0.147)	(0.044)
CONSTANT	-0.078***	-0.058**	-0.050***
	(0.014)	(0.022)	(0.014)
N	6564	3668	6300
clusters	24	24	24
R^2	0.0189	0.0218	0.0212
trim	1%	1%	5%

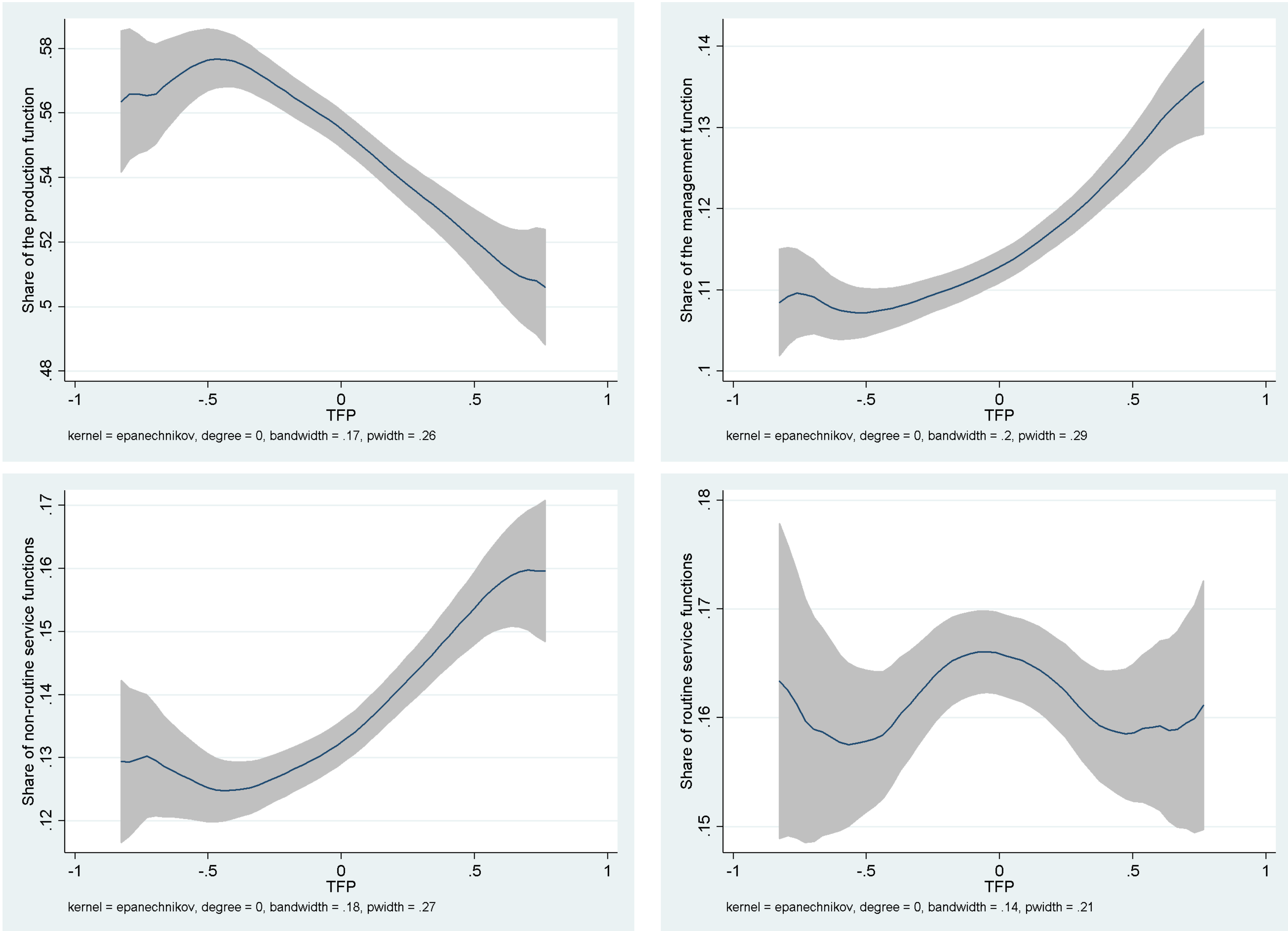
In this table, we regress productivity on the share of functions within the firm, when productivity is estimated also using intangible capital. [?] Sample trimmed at 0.5% at each tail of estimated TFP. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table C.4: Non-routine service functions in 2015 and process innovation between 2016-2018

	production and R&D	logistics	Information and communication	accounting, mgmt & administration	procedures and external links	organization, HR, decision making	marketing
Non-routine lateral	0.267*	0.328***	0.620***	0.263***	0.512***	0.405***	0.761***
	(0.132)	(0.113)	(0.084)	(0.093)	(0.121)	(0.099)	(0.094)
Ln of hours worked	0.121***	0.080***	0.076***	0.056***	0.063***	0.079***	0.050***
	(0.012)	(0.016)	(0.017)	(0.013)	(0.013)	(0.012)	(0.016)
Constant	-1.102***	-0.903***	-0.805***	-0.526***	-0.712***	-0.790***	-0.459**
	(0.159)	(0.205)	(0.217)	(0.165)	(0.173)	(0.158)	(0.204)
N	1571	1571	1571	1571	1571	1571	1571
N_clust	24	24	24	24	24	24	24
$R2$	0.1008	0.0766	0.0740	0.0474	0.0653	0.0740	0.0812
industry FE	Y	Y	Y	Y	Y	Y	Y
R&D	0.365	0.033	0.601***	0.274*	0.623***	0.440*	0.384*
	(0.314)	(0.100)	(0.120)	(0.148)	(0.137)	(0.216)	(0.218)
Other non-routine lateral	0.194	0.551***	0.635***	0.256**	0.429**	0.379***	1.046***
	(0.172)	(0.167)	(0.131)	(0.108)	(0.203)	(0.131)	(0.145)
Ln of hours worked	0.120***	0.083***	0.076***	0.056***	0.062***	0.079***	0.054***
	(0.012)	(0.015)	(0.016)	(0.012)	(0.013)	(0.012)	(0.016)
Constant	-1.085***	-0.956***	-0.808***	-0.524***	-0.692***	-0.783***	-0.527**
	(0.157)	(0.192)	(0.210)	(0.159)	(0.174)	(0.160)	(0.203)
N	1571	1571	1571	1571	1571	1571	1571
N_clust	24	24	24	24	24	24	24
$R2$	0.1011	0.0807	0.0740	0.0475	0.0658	0.0741	0.0869
industry FE	Y	Y	Y	Y	Y	Y	Y

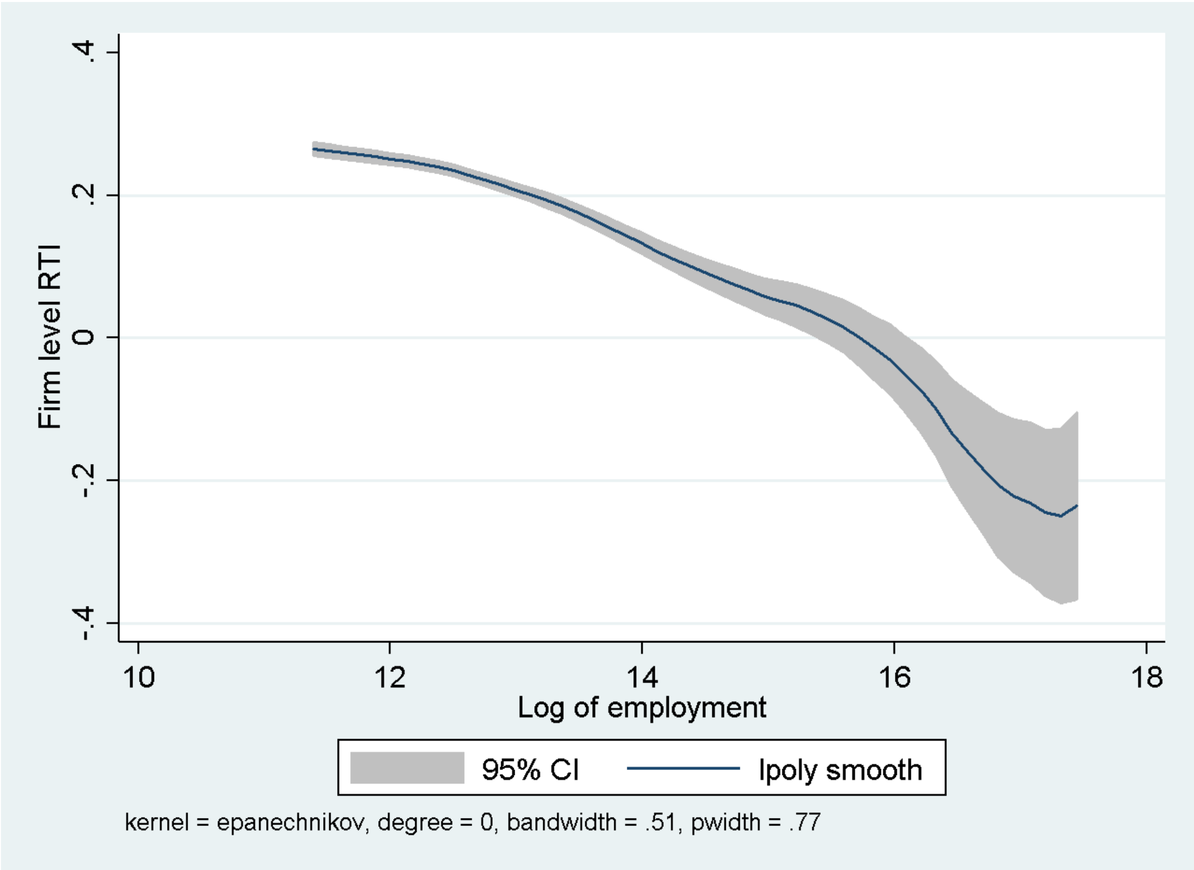
Process innovation measures from the 2018 of the French version of the Community Innovation Survey (CIS).
Top panel: correlations with overall share of non-routine service functions. Lower panel: with a distinction between R&D and other non-routine services. 2-digit industry level effects in all specifications.
Process innovation measures from the 2018 of the French version of the Community Innovation Survey (CIS). Dummy variables yes(=1)/no answers to 7 distinct questions about process innovation about different areas of firm activity.
Sample trimmed at 0.5% at each tail of estimated TFP for all firms. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

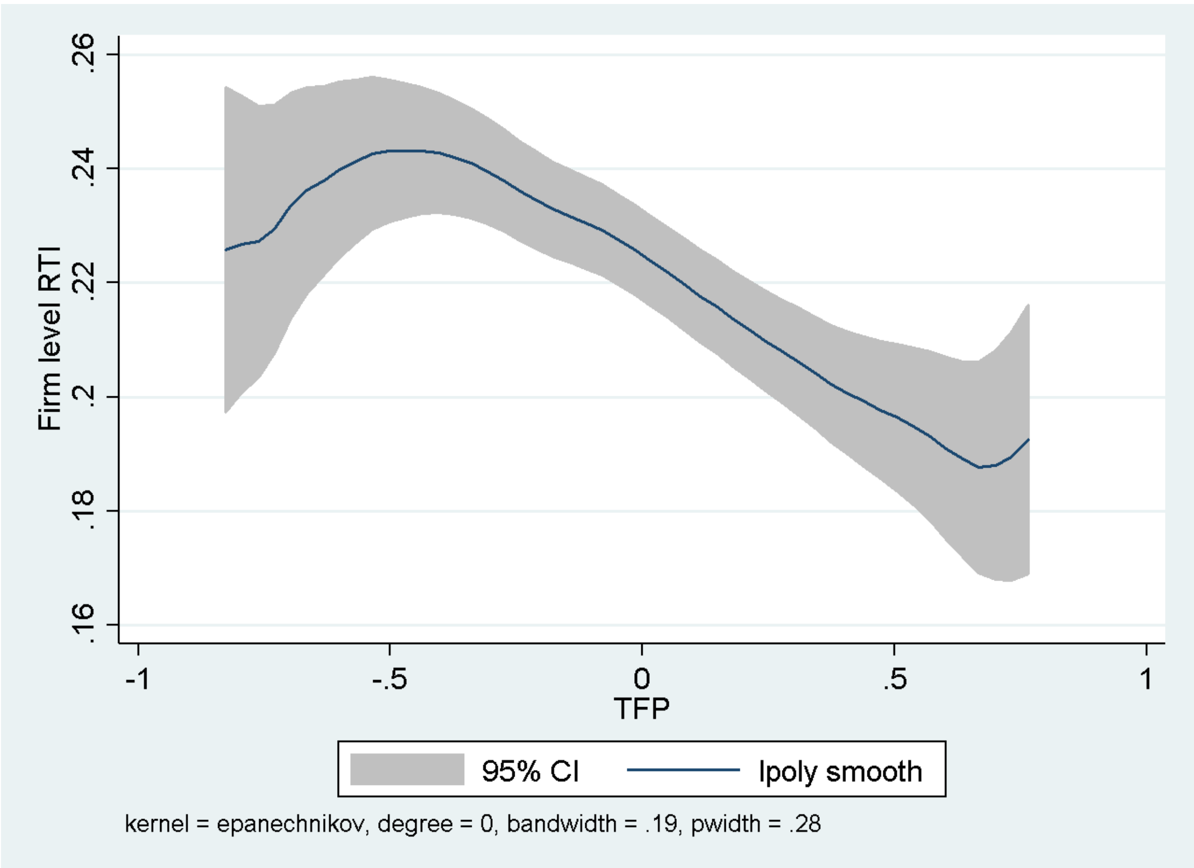
Figure C.1: Firm TFP and share of different functions.



Figures show a kernel estimate of the relationship between firm TFP and the employment share within different functions. The values may not sum to one as "other functions" category is not depicted. 95% confidence intervals around the estimate are shown. Sample trimmed at the top and bottom 2.5% estimate of TFP.

Figure C.2: Firm-level routinness vs. employment and TFP.





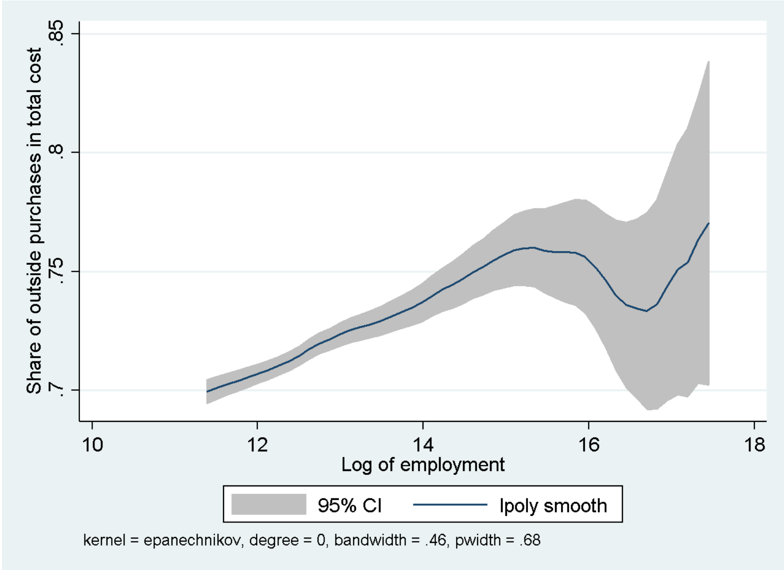
Figures show a kernel estimate of the relationship between respectively firm employment or TFP and the routineness measure at the firm level. 95% confidence intervals around the estimate are shown. Sample trimmed at the top and bottom 2.5% estimate of TFP.

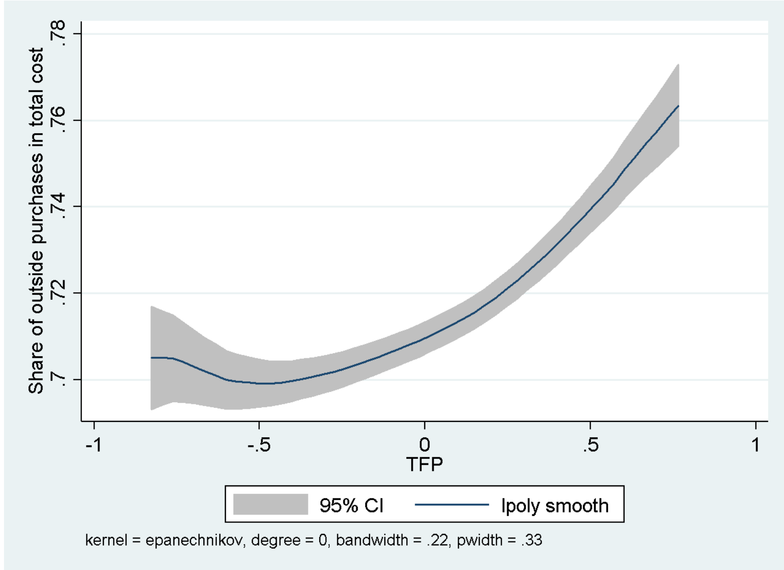
Table C.5: Function shares and outsourcing intensity

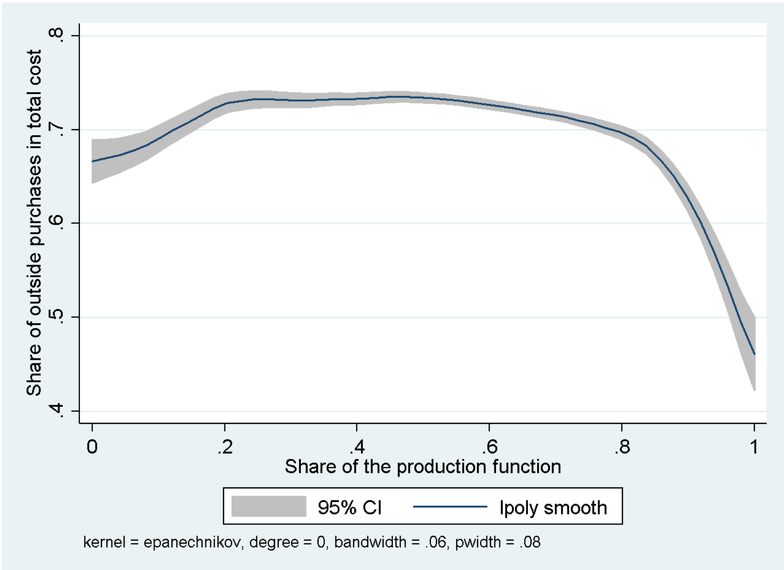
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ln of hours worked	0.015***		0.017***	0.015***	0.014***	0.012***	0.013***	0.012***	0.015***
	(0.003)		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TFP_narrow_v2		0.050***							
		(0.008)							
Management			0.340***						
			(0.064)						
Routine service				0.120*					
				(0.064)					
Production workers					-0.110***	0.355***			
					(0.032)	(0.112)			
$(Production workers)^2$						-0.468***			
						(0.104)			
Non-routine service							0.116**	0.428***	
							(0.046)	(0.098)	
$(Non - routine service)^2$								-0.615***	
								(0.136)	
B-2-B: purchases									0.922***
									(0.286)
Constant	0.597***	0.790***	0.543***	0.579***	0.683***	0.615***	0.616***	0.617***	0.598***
	(0.039)	(0.000)	(0.044)	(0.039)	(0.042)	(0.047)	(0.039)	(0.039)	(0.039)
N	6648	6648	6648	6648	6648	6648	6648	6648	6648
clusters	24	24	24	24	24	24	24	24	24
r2	0.1748	0.1824	0.2051	0.1838	0.1962	0.2229	0.1824	0.1957	0.1831
industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

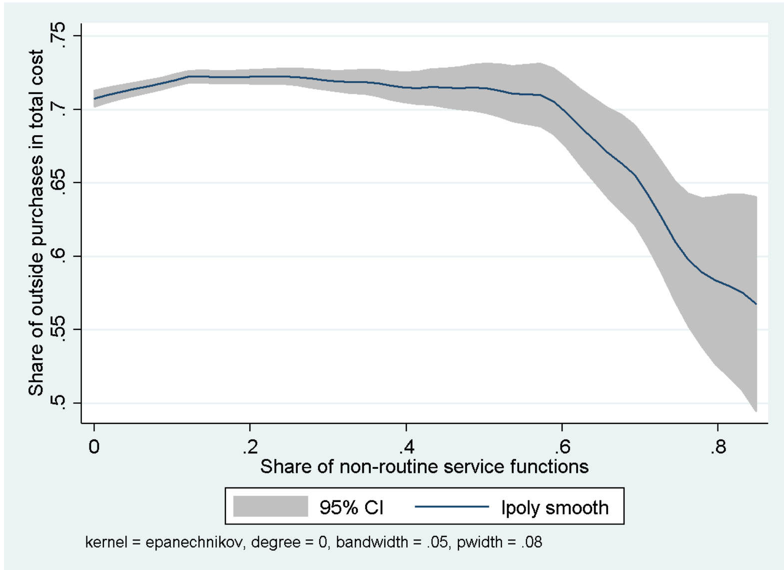
This table shows correlations of outsourcing intensity (outside purchases of materials, goods and services including rent) in total operating cost with hours worked, TFP and shares of different functions. Sample trimmed at 0.5% at each tail of estimated TFP. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Figure C.3: Share of outside purchases in total cost vs. employment and TFP.









Figures show a kernel estimate of the relationship between respectively firm employment, TFP, the share of production labor or non-routine service workers and the share in total cost of materials and services purchased outside of the firm. 95% confidence intervals around the estimate are shown. Sample trimmed at the top and bottom 2.5% estimate of TFP.

C.1 More on service functions and hierarchies

In this section, we provide evidence that the role of non-routine service functions differ from the hierarchical knowledge production discussed in the literature (see Garicano, 2000; Caliendo et al., 2015, among others) and provide a complementary view of firm organization. Non-routine service functions typically do not have workers in the lowest skill layers (CS 5 and 6), but most e.g., at the hierarchical level of "cadres" performing them are not managers. Moreover, if, given the skill content of some of those functions we would assign them a place within a hierarchy, hierarchies of many firms would appear to be bell-shaped instead of pyramid-like as required by management theories. This is important, because our approach points how empirical study of organizations could be refined.

We have already shown using the Appendix Table C.1 that both functional and hierarchical composition of labor are important correlates of productivity.

To investigate this issue further, we split further the service functions into distinct subfunctions imposing the hierarchical structure used by Caliendo et al. (2015)[38] and investigate e.g., the correlation of different subfunctions shares with firm productivity.[39] If the hierarchical-vertical view would account purely for productivity differentials between firms, the obtained coefficients for all functions on the same hierarchical layer should be of the same sign and magnitude—and the nature of the tasks performed at each layer would not matter.

We superimpose functions with hierarchical layers. Results are shown in Table C.6, with employment shares of lowest-ranking employees and workers (CS 5 and 6 as the base category). We focus our attention on the first column. Interestingly, only few of the functions X hierarchy levels shares' come out as statistically significant. It is not necessarily true that shares of subfunctions with higher skilled workers (as witnessed by the PCS classification) are always correlated with higher firm productivity. Higher CEO (a management function) share in hours worked is strongly correlated with productivity. Higher shares of cadres (the second-highest level layer after the CEO) in management, B-2-B, maintenance and production are correlated with TFP, while those in all other are not.[40] Among functions in middle- and lowest hierarchical levels we find that the shares of hours worked in total employment of lowest level (CS 5 or 6) B-2-B and management (e.g., office clerks) workers are positively related with TFP. We reject the Wald tests of equality of coefficients for the "cadres" (CS3) layer at 1% and for the mid-level (CS4) at 2% level. These results suggest that—as R&D or intellectual services functions' shares come as statistically significant overall as shown in Table 9 and discussed in Section 4.3—it may be the team output of a function that matters and not the hierarchical structure. We also observe that the "management" function at all levels is an important correlate of productivity, confirming the notions advanced e.g., in Bender et al. (2018). We conclude that although hierarchy structure is correlated with productivity overall as argued in Caliendo et al. (2020), functional composition of the workforce may matter as well.

C.2 Productivity - Levisohn and Petrin (2003)

[38] These authors use the 1 digit PCS classification of occupations and group jobs into 4 hierarchy layers: CEOs (CS code "2"); senior staff or management positions (CS "3"); employees at a supervisor level (CS "4"); other qualified or non-qualified white and blue-collar workers (heterogenous group of CS "5" and "6" codes)

[39] This is possible by using the 4-digit PCS classification.

[40] Statistical significance of the correlation of these functions is unrelated to the size of particular subfunction in overall workforce (Table C.2).

Table C.6: Productivity and detailed functions

	2-digit-NACE FE	Multi-plant firms	2-digit-NACE FE
Management: CEOs	1.375***	1.676**	1.582***
	(0.405)	(0.699)	(0.370)
Management: cadres	0.727**	0.652	0.610**
	(0.262)	(0.410)	(0.239)
Managers: mid-level	0.582**	0.706**	0.454**
	(0.211)	(0.318)	(0.174)
Managers: office workers	0.488***	0.493***	0.321***
	(0.126)	(0.130)	(0.113)
B-2-B: cadres	0.317***	0.512***	0.294***
	(0.106)	(0.149)	(0.078)
B-2-B: mid-level	0.006	0.017	0.059
	(0.122)	(0.150)	(0.097)
B-2-B: lowest level	1.650**	1.873**	1.015
	(0.769)	(0.872)	(0.692)
R&D: cadres	0.202	0.176	0.171
	(0.121)	(0.160)	(0.107)
R&D: technicians	0.110	0.171	0.095
	(0.114)	(0.181)	(0.108)
Intellectual services: cadres	0.062	0.726**	0.420
	(0.382)	(0.324)	(0.341)
Intellectual services: mid-level workers	0.305	-0.087	0.262
	(0.269)	(0.407)	(0.203)
Maintenance: cadres	0.683***	0.763**	0.659***
	(0.151)	(0.324)	(0.148)
Maintenance: technicians	0.041	0.040	0.048
	(0.053)	(0.102)	(0.050)
Maintenance: lowest level	0.104	0.117**	0.113*
	(0.064)	(0.053)	(0.058)
Transport and logistics: cadres	0.313	0.780	0.846*
	(0.593)	(0.840)	(0.480)
Transport and logistics: mid-level	0.232	0.447	0.099
	(0.301)	(0.367)	(0.275)
Transport and logistics: lowest-level	-0.017	-0.046	-0.022
	(0.042)	(0.068)	(0.037)
Production: cadres	0.429***	0.334	0.398***
	(0.136)	(0.233)	(0.108)
Production: technicians and foremen	-0.041	-0.094	-0.011
	(0.047)	(0.091)	(0.035)
Other functions	0.112**	0.154***	0.124***
	(0.044)	(0.026)	(0.043)
CONSTANT	-0.095***	-0.141***	-0.077***
	(0.014)	(0.017)	(0.014)
N	6648	2917	6380
clusters	24	24	24
R^2	0.0273	0.0460	0.0346
trim	1%	1%	5%

In this table, we regress productivity on the share of functions within the firm, when TBA. [?] Sample trimmed at 0.5% at each tail of estimated TFP. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table C.7: Productivity (Levisohn and Petrin, 2003) and functions

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Management		0.505***	0.550***	0.477***	0.481***	0.542***	0.455***	0.582***	0.612***	0.481***
		(0.072)	(0.076)	(0.066)	(0.076)	(0.079)	(0.066)	(0.092)	(0.098)	(0.082)
Non-routine lateral		0.319***	0.410***	0.363***				0.353***	0.455***	0.381***
		(0.057)	(0.077)	(0.053)				(0.071)	(0.096)	(0.066)
	textitB-2-B				0.385***	0.470***	0.432***			
					(0.106)	(0.107)	(0.066)			
	textitR&D				0.252***	0.366***	0.284***			
					(0.077)	(0.093)	(0.071)			
	textitIntellectual services				0.352**	0.428**	0.456***			
					(0.165)	(0.176)	(0.150)			
Routine lateral		0.047	0.065	0.074				0.056	0.071	0.076
		(0.046)	(0.079)	(0.045)				(0.044)	(0.073)	(0.045)
	textitMaintenance				0.088	0.194	0.172**			
					(0.060)	(0.116)	(0.066)			
	textitTransport and logistics				-0.006	-0.042	-0.007			
					(0.033)	(0.036)	(0.037)			
Managers (share)								0.133	0.219**	0.295***
								(0.097)	(0.094)	(0.095)
Other functions		0.174***	0.191**	0.186***	0.168***	0.191***	0.184***	0.177**	0.186**	0.191***
		(0.059)	(0.071)	(0.061)	(0.057)	(0.067)	(0.057)	(0.063)	(0.074)	(0.064)
CONSTANT		-0.112***	-0.112***	-0.104***	-0.108***	-0.108***	-0.100***	-0.113***	-0.113***	-0.105***
		(0.015)	(0.016)	(0.011)	(0.015)	(0.013)	(0.010)	(0.016)	(0.015)	(0.011)
N		6312	6312	6056	6312	6312	6056	6312	6312	6056
N_clust		24	24	24	24	24	24	24	24	24
R^2		0.0282	0.0348	0.0415	0.0287	0.0361	0.0430	0.0280	0.0346	0.0405
industry FE		N	Y	Y	N	Y	Y	N	Y	Y
trim		1%	1%	5%	1%	1%	5%	1%	1%	5%

In this table, we regress productivity estimated via the Levisohn and Petrin (2003) method on the shares of functions within the firm. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels. Sample: Firms with more than 50 employees both in 2014 and 2015.

Table C.8: Routineness and productivity

Item	col 1	col 2	col 3	col 4	col 5	col 6	col 7	col 8
Firm-level	-0.056***							
	(0.017)							
B-2-B		-0.085**						
		(0.031)						
R&D			-0.068***					
			(0.024)					
Maintenance				-0.026*				
				(0.013)				
Production					-0.057***			
					(0.014)			
Management						-0.000		
						(0.011)		
Transport and logistics							0.004	
							(0.005)	
Intellectual services								-0.066**
								(0.030)
CONSTANT	0.029***	-0.040*	-0.024	0.009***	0.037***	0.010	0.008***	-0.025
	(0.005)	(0.020)	(0.015)	(0.001)	(0.006)	(0.007)	(0.000)	(0.018)
N	6580	5918	5327	6202	6564	6535	6371	4293
clusters	24	24	24	24	24	24	24	24
R^2	0.0043	0.0036	0.0045	0.0039	0.0044	0.0026	0.0031	0.0045

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.