Journal of Economics & Management Strategy

ORIGINAL ARTICLE OPEN ACCESS

Firm Hierarchy and the Market for Knowledge

Fabio Pieri[1,2] | Massimiliano Vatiero[1,3]

[1]Department of Economics and Management, University of Trento, Trento, Italy | [2]Centro Studi Luca d'Agliano, Milan, Italy | [3]Law Institute, Università della Svizzera italiana, Lugano, Switzerland

Correspondence: Fabio Pieri (fabio.pieri@unitn.it)

Received: 4 November 2022 | **Revised:** 6 June 2024 | **Accepted:** 14 September 2024

Funding: Massimiliano Vatiero gratefully acknowledges financial support from the "Brenno Galli" Fondo (Fondo Ricerca e Sviluppo, USI).

Keywords: firm hierarchy | market for knowledge | number of layers | span of control

ABSTRACT

This paper sheds light on the role of the market for knowledge in shaping a firm's hierarchy—that is, the span of control and the number of layers. We predict that the larger the extent of the market for knowledge, the larger the span of control and the fewer the layers. We test our predictions using a rich database representing industrial firms in Italy over the period 2004–2017. Our identification strategy employs a difference-in-difference approach that exploits the cross-regional variability in the extent of the local market for knowledge-intensive business services and the cross-industry heterogeneity in the level of technological exposure of industrial firms to such a market. We find that a thicker regional market for knowledge is associated with relatively flatter firms in industries that use knowledge-intensive business services more intensively. The results are confirmed when we use instrumental variables for the extent of the market for knowledge, test the sensitivity of the estimates to omitted variable bias, and perform a series of robustness checks.

JEL Classification: D21, D22, D23, L22, L23

1 | Introduction

Economists, management consultants, international organizations, and the popular press recognize the pivotal role of knowledge in creating the competitive advantage of modern firms. So, how do firms manage knowledge? This work examines how firms determine their hierarchical structure to manage the knowledge required in the production processes. In particular, we hypothesize and test that the shape of a firm's hierarchy depends, *ceteris paribus*, on the availability of knowledge outside of the firm's boundaries—that is, on the extent of the market for knowledge.

We exploit a powerful idea of Garicano (2000), who has shown that the hierarchical structure of organizations increases the utilization rate of knowledge by shielding specialized supervisors or managers from solving routine problems and allowing them to focus on solving exceptional problems instead.[1] Thus, the greater the level of predictability within a firm's production process, the greater the likelihood that the problems the firm faces are relatively simple and solvable with the knowledge available at lower layers of the firm. In that case, adding upper layers of specialized supervisors is unnecessary. Otherwise, whenever the match between problems and solutions is costly at a firm's lower layers, the firm adds an upper layer of problem-solvers to leverage their knowledge. In that way, the hierarchy reduces the learning costs inherent in solving problems but increases the costs of communicating solutions between layers.

But when exceptional problems are frequent, what are the alternatives to adding layers of supervisors? In such cases, a firm may resort to the market for knowledge, typically by buying the solution from providers of specialized business services, instead of relaying the problems to supervisors or managers. To the best of our knowledge, the theoretical and empirical literature on knowledge-based hierarchies has not paid enough attention to that alternative.[2]

Our work fills this gap by investigating how the market for knowledge affects firm hierarchy. In particular, we seek to answer the following question: how do firms shape their hierarchies if they can acquire the knowledge that they need from the market? To answer this question, the theoretical part of the paper adds the market for knowledge to the basic model of Garicano (2000) to investigate how it affects the fundamental dimensions of firm hierarchy, that is, the span of control and the number of layers. We predict that the lower the cost of buying solutions in the market for knowledge, the larger the span of control and the fewer the layers, ceteris paribus.

In the empirical analysis, we assume that the greater the extent of the local market for knowledge (i.e., the thicker it is), the greater the availability of solutions for industrial firms located there and, in turn, the lower the cost of buying solutions in the market. We use an interaction model that exploits the cross-regional variability in the extent of the market for knowledge-intensive business services (KIBS) and the cross-industry heterogeneity in the level of technological exposure of industrial firms to the use of KIBS. The market for KIBS is one in which solutions for production problems are bought and sold by firms.[3] The empirical model is inspired by the difference-in-difference (DiD) approach pioneered by Rajan and Zingales (1998b).[4] We consider that if the thickness of the market for knowledge has a direct effect on firm hierarchy, this effect is likely to be larger in industries that are relatively more exposed to the local market for knowledge. For this reason, industries that use KIBS more intensively as inputs are naturally (due to idiosyncratic demand and technological factors) more exposed to the extent of the local market for knowledge and are the "treatment" group. Conversely, industries that use KIBS less intensively are the "control" group.

We confirm our theoretical predictions using representative data on the hierarchies of industrial firms in Italy, which we complement with information on the extent of the market for KIBS in the region where the firm is headquartered. The extent of the regional market for KIBS is associated with a higher average span of control (i.e., the number of subordinates per supervisor) and a lower number of layers in industries that use business services more intensively. Since the extent of the market for knowledge may be endogenous to the firm hierarchy, we adopt an instrumental variable (IV) approach by using information on the share of regional firms with access to broadband and the share of regional employment in cooperative firms as IVs for the extent of the regional market for KIBS. The IV estimates confirm that a thicker regional market for KIBS is associated with relatively flatter firms (i.e., larger span of control and lower number of layers) in industries that use KIBS intangibles more intensively. We complement this evidence with a test of the sensitivity of the estimates to omitted variable bias and a series of robustness checks.

1.1 | Related Literature

Our paper contributes to several streams of research. The first one is the mentioned literature on knowledge-based hierarchies (Garicano 2000, 2010; Garicano and Rossi-Hansberg 2004, 2006, 2012, 2015; Garicano and Wu 2012), wherein three structural factors that shape firm hierarchy have received the most attention: technology, product market competition, and geographic frictions. Regarding technology, Garicano and Rossi-Hansberg (2004) have examined how reducing the costs of communicating knowledge through the use of new communication technology (CT) enables hierarchies with a larger span of control. Garicano and Rossi-Hansberg (2006, 2012), who included both CT and information technologies (ITs) as factors that shape a firm's hierarchy, showed that CT increases both the span of control and the number of layers in a firm, while the adoption of new IT that reduces learning costs leads to a decrease in the number of layers. Bloom et al. (2014) built proxies for both IT and CT in a large sample of firms in Europe and the United States and showed the heterogeneous effect of these types of technology on firm hierarchy. Delmastro (2002) obtained results similar to Bloom et al. (2014) but in a sample of firms active in Italy's metalworking industry. Product market competition affects firm hierarchy, as well. Guadalupe and Wulf (2010) found that firms in the United States reduced their number of layers and increased the span of control of their top managers due to increased competitive pressures following the liberalization of trade between Canada and the United States in 1989. Caliendo and Rossi-Hansberg (2012) have additionally shown that trade liberalization reduces the number of layers of nonexporters but increases it for exporters. More recently, Barba Navaretti et al. (2024) found that introducing technical barriers to trade in importing countries prompted firms to increase their share of managers at the top hierarchical layer. Cooke, Fernandes and Ferreira (2021) examined a comprehensive deregulation event that occurred across municipalities in Portugal from 2005 to 2009 and found that pro-competitive deregulation in the product market implied a reduction in the number of layers and an increase in the managers' span of control. Finally, in a recent paper, Gumpert, Steimer and Antoni (2022) have studied how geographic frictions (the distance between a firm's headquarters and its plants) affect hierarchy. They show both theoretically and empirically (by using data on German firms) that geographical frictions hamper knowledge flows within and between the plants of a firm and, therefore, increase its number of layers. Caliendo, Monte, and Rossi-Hansberg (2015) and Caliendo et al. (2020) have shifted their focus away from these structural factors and instead examined how demand shocks affect the number of layers in French and Portuguese firms, respectively.

We contribute to that stream of literature by investigating how the extent of the market for knowledge affects firm hierarchy. First, we construct a model in which firms can pay external specialized problem-solvers instead of bearing higher costs for learning and communicating knowledge. In our setting, using the market for knowledge is an alternative to maintaining deeper hierarchies. Second, we conduct, to the best of our knowledge, the first empirical test on the role of the market for knowledge in firm hierarchies, by employing granular data about the layers within organizations in a representative sample of industrial firms in Italy.

Second, and more broadly, our work relates to the literature that addresses hierarchy as a tool for supervising subordinates (Williamson 1985).[5] Rajan and Zingales (2001) introduced a scenario in which an entrepreneur gives their subordinates access to the core of the firm's technology—the so-called "critical resource" (Rajan and Zingales 1998a)—to specialize and enhance their knowledge. Although proximity provides better access to the core technology and thus enhances learning, subordinates may expropriate that core, establish their own businesses, and, in turn, compete with the entrepreneur. To diminish that hold-up risk, entrepreneurs should reduce the number of layers between themselves and subordinates and increase their span of control. In that case, a puzzle emerges: while theories on knowledge-based hierarchies (Garicano 2000, 2010; Garicano and Rossi-Hansberg 2004, 2006, 2012, 2015; Garicano and Wu 2012) predict positive returns from deeper hierarchies in terms of utilization rate of knowledge, Rajan and Zingales (2001) and the hold-up theory seem to suggest negative returns from deeper hierarchies in terms of looser control over subordinates and higher risk of expropriation. However, if knowledge can be bought into the market, the hold-up theory would predict firms with a higher number of layers and a lower span of control because proximity to the critical resource and shorter hierarchies to control subordinates become less necessary. Instead, our primary theoretical and empirical results are supportive of the theory on knowledge-based hierarchies, as they show that the larger the extent of the market for knowledge, the lower the number of layers and the larger the span of control. These findings make it clear that shorter hierarchies may well depend on a larger market for knowledge, the development of which may reduce the need for proximity to the critical resource and the expropriation costs borne by entrepreneurs.

Third, our paper contributes to the knowledge-based theory of the firm, according to which a firm is a coordinated nexus of workers' knowledge (e.g., Kogut and Zander 1992; Conner and Prahalad 1996; Grant 1996). According to this theory, a firm's organization is not primarily designed to allocate tangible resources, but is a tool for improving the use of intangible assets across the firm's units, including managers' specialized knowledge and oversight (Atalay, Hortaçsu, and Syverson 2014; Garicano and Hubbard 2016; Halac and Prat 2016). Our paper relates to this literature in two ways. First, we empirically unpack a firm's knowledge into related "blocks" proxied by workers' level of occupations (i.e., white collars, blue collars, middle managers, managers, and the entrepreneur). Second, we offer a more comprehensive theoretical framework in which using the market for knowledge is an alternative to maintaining firm hierarchy. In our framework, knowledge can be provided both within the firm through upper layers of supervisors and by specialized problem-solvers in the market for knowledge. In that sense, the potential knowledge set of a firm is larger than the firm's boundaries (Brusoni, Prencipe, and Pavitt 2001) and increases as the market for knowledge expands. Thus, we employ a broad conceptualization of the market for knowledge that embraces transactions whose object is an intangible service (i.e., a solution) that supports firms' production processes (Arora, Fosfuri, and Gambardella 2001).

Last, our paper contributes to the literature on the role of KIBS in developing clients' industries (O'Farrell and Moffat 1995) and, by extension, economies, and territories. Several works have provided sound evidence that KIBS are associated with structural changes within territories (Duranton and Puga 2005; Meliciani and Savona 2015), industrial firms' innovation (Muller and Zenker 2001), and productivity. In this respect, our work suggests that the extent of the market of knowledge, specifically the relevance of KIBS in a region, shapes the hierarchical structure of the industrial firms located there.

The rest of the paper is organized as follows. In Section 2, we propose a theoretical framework that identifies how, ceteris paribus, the market for knowledge affects the span of control and the number of layers. From that theoretical framework, we derive our primary predictions. Section 3 describes the data and the variables used in this work to test our predictions. In Sections 4 and 5, we present the empirical analysis and results, respectively. Section 6 provides some concluding remarks.

2 | A Theoretical Setting

2.1 | Our Starting Point: Garicano (2000)

According to Garicano (2000), a firm's hierarchy is a means of managing knowledge: when workers in lower layers of the firm encounter problems that they cannot solve, they turn to supervisors in the next-higher layer in the hierarchy.

Let us assume that all agents involved in a production process are endowed with one unit of time that they spend in the firm. Let $\Omega \subset R^{+}$ be the set of all possible production problems that a firm may face and $A_i \subset \Omega$ be the set of problems that an agent, i, is able to solve—that is, problems within their knowledge set. Production requires that problem $z \in \Omega$ be solved, which happens whenever $z \in A$. If the problem z can be solved, then the potential for production becomes Q units of output. Instead, if an agent cannot solve the problem, then they will ask their supervisor in the next-highest layer within the firm's hierarchy for a solution.

Each agent can learn the knowledge to solve the problem at the cost of $c < 1$, in which c is the learning cost per unit of time. Thus, learning how to solve all problems in the interval $[0, z]$ has a cost of cz. If the supervisor has the solution, then they will spend a fraction of a unit of time, $h < 1$—that is the communication cost per unit of time—on communicating the solution to the agent in the lower layer. Otherwise, and if the conditional probability of solving the problem is enough to bear the associated costs, the supervisor will ask their own supervisor at the next-highest layer in the firm, and so on, until the problem reaches the highest layer, occupied solely by the entrepreneur. Thus, fractions c and h represent two kinds of costs that a firm incurs to *make* solutions to production problems internally. The learning cost, c, negatively depends on several factors, including the range of expertize acquired by members of each layer (e.g., via education, on-the-job training, and learning by doing). The communication cost, h, depends on communications technology, including investments in CT and process innovation.

To add tractability, problems are ordered from the most common to the most exceptional and drawn from an exponential distribution. Thus, a firm solves the fractions of problems $F(z) = 1 - e^{-\lambda z}$. The unique parameter of the exponential distribution, $\lambda > 0$, thus characterizes the predictability of the firm's production process. The greater the predictability, λ, the greater the likelihood that production problems are relatively simple and therefore solvable with the knowledge available within the lower layers of the firm.

Consider the case of an entrepreneur who hires a number of workers, n, and forms a hierarchy with two layers, the top layer of which will be occupied by the entrepreneur. Let us use subscript E to indicate the entrepreneur and subscript w to indicate workers. Each worker has a knowledge set, A_w, whereas the entrepreneur's knowledge set is A_E. The firm solves problems in the interval $\left[0, z_w^2 + z_E^2\right]$, in which the superscript "2" indicates that the firm has two layers. Each worker solves a fraction of problems $\left(1 - e^{-\lambda z_w^2}\right)$, whereas the problems that workers cannot solve are relayed to the entrepreneur. The entrepreneur spends h_E unit of time on communicating the solution if the problem falls within their knowledge set. In that way, the entrepreneur can deal, at most, with $\frac{1}{h_E}$ problems. The entrepreneur's time constraint is $hne^{-\lambda z_w^2} \leq 1$, meaning that they can hire $n = \frac{e^{\lambda z_w^2}}{h}$ workers at most.

The learning cost is lower at each successively higher layer within the firm, namely $c_w > c_E$. That condition is the primary reason for maintaining a hierarchy as a means to increase the utilization rate of knowledge. Because the communication of solutions flows downward from supervisors to subordinates, adding a layer increases the cost of such communication, however. Thus, while taking account of that trade off, the firm has to decide how many layers to have, what each agent on each layer needs to learn, and whom each agent should ask when confronted with an exceptional problem. The optimal organization has a pyramidal shape, with production workers at the base and fewer supervisors or managers who learn knowledge to solve exceptional problems (Figure 1A).

2.2 | Our Contribution to Theory: Firm Hierarchy and the Market for Knowledge

In our study, we add a market for knowledge to Garicano (2000), basic model and examine the potential *make-or-buy* dilemma that a firm may face for increasing the utilization rate of knowledge. In that dilemma, instead of relaying an unsolved exceptional problem to the next-highest layer in the hierarchy, the firm could pay someone outside the firm, typically a business service provider, to solve it. Thus, according to our model, the hierarchy's shape depends on the costs of learning and communicating solutions internally—that is, "making" the solutions—versus the cost of "buying" those solutions in the market for knowledge (Figure 1B).

Let us indicate with p (where $p < 1$) the cost of buying a solution in the market. Similar to c and h, p is normalized in terms of units of time; each agent can save time—that is, find solutions in the market at the cost of p—instead of spending time learning how to solve the problem (c) and/or communicating the solution (h). We assume that the greater the extent of the market for knowledge, the lower the cost of buying solutions in that market.

In a previous extension of Garicano (2000), model resembling our work with the market for knowledge, Garicano and Rossi-Hansberg (2006) conceived that each agent can buy solutions from providers of consultant services at the cost of a consulting fee. In their theoretical framework, the authors predicted the consequences of that market for the wage inequality between workers and managers within the firm. Moreover, they derived, under wage equalization, an equivalence between knowledge-based hierarchies and (knowledge) transactions in the market for consultant services. Although Garicano and Rossi-Hansberg's (2006) framework and ours share certain features,

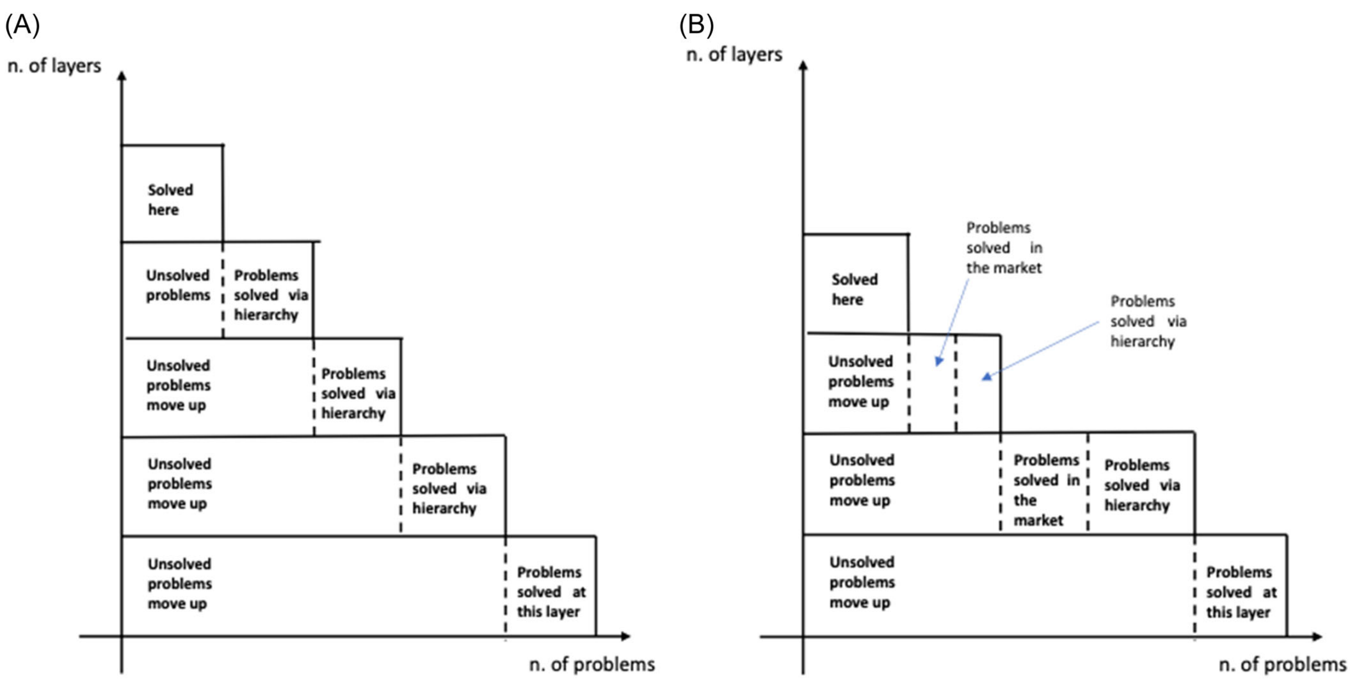

FIGURE 1 | Knowledge-based hierarchy and the market for knowledge. (A) Garicano's setting and (B) our setting. [Color figure can be viewed at wileyonlinelibrary.com]

there are two major differences. First, Garicano and Rossi-Hansberg (2006) investigated labor market equilibrium and focused on the wages of workers and managers when firms can buy knowledge in the market for consultant services, and their framework reflects that intention. Instead, we are interested in the consequences of the extent of the market for knowledge for the hierarchical structures of firms. Second, whereas Garicano and Rossi-Hansberg (2006) stressed the equivalence between hierarchy and the market for knowledge, in our framework firms manage knowledge using a combination of both internal hierarchy and the market for knowledge, the latter of which becomes more relevant—for the firm—as it becomes thicker.

2.3 | The Model and Testable Predictions

The design of a firm's hierarchy involves determining the number of layers or depth, L, which represents the vertical dimension of the hierarchy, and the span of control or breadth, B, which represents the horizontal dimension of the hierarchy. Garicano (2000) has identified the first key trade-off faced by the firm as being between learning (c) and communication (h) costs. Indeed, by adding layers of problem-solvers, the firm increases the utilization rate of knowledge, thereby economizing on learning but increasing communication costs. We add a second trade-off, one between the knowledge managed through the hierarchy (which involves learning and communication costs, c and h, respectively) and the knowledge bought in the market.

Given the existence of a market for knowledge, the supervisor or manager may now deal with a maximum of $\frac{1}{h} + \frac{1}{p}$ problems. The entrepreneur's time constraint becomes $\frac{\text{hpne}^{-\lambda z_w^2}}{p+h} \leq 1$, and thus they can hire at maximum of $n = \frac{e^{\lambda z_w^2}(p+h)}{ph} = \frac{e^{\lambda z_w^2}}{h} + \frac{e^{\lambda z_w^2}}{p}$ subordinates. The net expected output is:

$$y_M^2 = Q\left(1 - e^{-\lambda\left(z_w^2 + z_E^2\right)}\right) - c_E z_E^2 - n c_w z_w^2 - p e^{-\lambda\left(z_w^2 + z_E^2\right)} \quad (1)$$

in which y_M^L indicates the output of a firm with L layers that can use the market (subscript M) for knowledge. Equation (1) can be rewritten as follows:

$$y_M^2 = Q\left(1 - e^{-\lambda\left(z_w^2 + z_E^2\right)}\right) - c_E z_E^2 - \left(\frac{e^{\lambda z_w^2}}{h} + \frac{e^{\lambda z_w^2}}{p}\right) c_w z_w^2 - p e^{-\lambda\left(z_w^2 + z_E^2\right)}.$$

In that configuration, $L = 2$ and $B = \frac{e^{\lambda z_w^2}}{h} + \frac{e^{\lambda z_w^2}}{p}$. Thus, the entrepreneur may buy solutions instead of making them within the firm. The span of control is negatively related to p: *ceteris paribus*, a reduction in p allows the entrepreneur to save time and devote it to controlling a greater number of subordinates.

That setting can be readily generalized to a firm with $L > 2$ layers in which subordinates, placed in layer $l - 1$, ask their supervisors placed in the layer l to solve production problems. Alternatively, the organization can buy knowledge in the market. If the firm's knowledge set is $A = z_E^L + \sum_{l=1}^{L-1} z_l^L$, then the expected output of the firm is:

$$y_M^L = Q(1 - e^{-\lambda A}) - c_E z_E^L - \sum_{l=1}^{L-1} n_l^L c_l^L z_l^L - p e^{-\lambda A} \quad (2)$$

subject to a time constraint for each manager at the layer l, presented as:

$$\frac{\text{hpn}_{l-1}^L e^{-\lambda z_{l-1}^L}}{p+h} \leq 1. \quad (3)$$

From that setting, we first derive two testable predictions, as follows.

Prediction 1. *The span of control at each layer* l (B_l) *negatively depends on the cost of buying solutions in the market* (p).

Proof. Given the time constraint (Equation 3), the span of control of one manager at the layer l is

$$B_l = \frac{e^{\lambda z_{l-1}^L}}{h} + \frac{e^{\lambda z_{l-1}^L}}{p}. \quad (4)$$

According to Equation (4), the higher the cost of buying solutions in the market, the narrower the span of control. By the same token, a reduction in p incentivizes the firm to buy solutions in the market for knowledge instead of "making" them internally. As a result, the manager's average time available to control subordinates increases, as does their span of control. □

Prediction 2. *The number of layers* (L) *positively depends on the cost of buying solutions in the market* (p).

Proof. Consider two firms, the first of which has L layers, whereas the second has $(L - k)$ layers. The first firm thus has more depth than the second one and exploits its greater hierarchy to solve exceptional problems. As such, without the market for knowledge, the first firm would produce a higher output than the second one. Now assume that the second firm, with $(L - k)$ layers, can access and use the market for knowledge, while the first, with L layers, cannot. As a consequence, the second firm, with $(L - k)$, may acquire the needed knowledge from the market and thereby obtain a higher expected output despite having fewer layers than the first firm. Using Equation (2), we have that $y_M^{L-k} > y^L$ if the cost of buying solutions in the market is less than the cost of adding the Lth layer. Thus, a firm characterized by a short hierarchy may have a higher expected output.

For instance, consider two firms, one with four layers and the other with three layers. The firm with three layers could obtain a higher expected output than the firm with four layers, if:

$$Q\left(1 - e^{-\lambda A^3}\right) - c_3 z_3^3 - \sum_{l=1}^{2} n_l^3 c_l^3 z_l^3 - p e^{-\lambda A^3} > Q\left(1 - e^{-\lambda A^4}\right) - c_4 z_4^4 - \sum_{l=1}^{3} n_l^4 c_l^4 z_l^4.$$

TABLE 1 | Our predictions.

Hierarchical dimensions	Determinants (expectations)	
Span of control (B)	Cost of buying solutions in the market p (−)	Prediction 1
	Predictability λ (+)	Prediction 3
	Learning costs c (−)	
	Communication costs h (−)	
Number of layers (L)	Cost of buying solutions in the market p (+)	Prediction 2
	Predictability λ (−)	Prediction 4
	Learning costs c (+)	
	Communication costs h (ambiguous)	

That condition may hold when the firm with three layers has access to the market for knowledge, whereas the firm with four layers does not.

The market for knowledge affects the number of layers. While a higher p incentives to add upper layers, a reduction in p makes solutions of productive problems cheaper and encourages upper layers to be substituted with knowledge transactions in the market. Thus, the firm reduces the depth of its hierarchy. □

With reference to Garicano (2000, 889–893 and his predictions 5, 6, and 7), the model also allows making two additional predictions, as follows.

Prediction 3. *The span of control (B_l) is positively related to the predictability of the process and negatively related to learning costs and communication costs (proof: Supporting Information S1: Section A.1).*

Prediction 4. *A reduction in communication costs* (h) *may increase or decrease the number of layers; a reduction in the learning costs* (c) *does not increase and may decrease the number of layers; and a reduction in the predictability* (λ) *of the process does not reduce but may increase the number of layers (proof: Cf. Supporting Information S1: Section A.1).*

Table 1 summarizes our predictions.

3 | Data and Descriptive Analysis

3.1 | Data Sources

We use an original database compiled by recovering information from various sources. The data on firms come from the *Rilevazione Longitudinale Imprese e Lavoro* (RIL), a mandatory survey conducted in Italy in five waves (i.e., 2005, 2007, 2010, 2015, and 2018) by the *Istituto Nazionale per l'Analisi delle Politiche Pubbliche* (INAPP). The RIL data cover a representative sample of Italian partnerships and limited liability companies of all size classes operating in the private, nonagricultural sectors. The information contained in RIL refers to the year before the publication of the wave.[6]

Various features of the RIL survey made it appropriate for our work. First, because the survey's aim was to collect accurate information from firms in Italy about the characteristics of their labor demand, including the number of employees per occupational category, we are able to develop proxies for the span of control and the number of layers within firms. Second, RIL includes a wide range of questions that allow us to control for a large set of firms' observable characteristics, particularly ones related to the primary parameters of the model: predictability (λ), learning costs (c), and communication costs (h). Third, the RIL database contains information regarding the industry in which each firm is active and the geographical region where it is located. That information is critical to assessing the role of the extent of the market for knowledge in the hierarchies of firms. Indeed, as detailed in Section 3.3, as a proxy for the market for knowledge, we use the information about the extent of the market for KIBS in the (NUTS 2)[7] region where each firm's headquarters are located. We also take advantage of the fact that, because some industries are more intensive than others in the use of KIBS as inputs in their production processes, firms operating in different industries likely show heterogeneous levels of technological exposure to the extent of the market for knowledge. In line with previous works (Colombo and Delmastro 1999; Delmastro 2002; Guadalupe and Wulf 2010; Caliendo, Monte, and Rossi-Hansberg 2015; Caliendo et al. 2020; Barba Navaretti et al. 2024), we focus on industrial firms.[8] In fact, industrial sectors are highly intensive in research and development (R&D), the most innovative, and where advances in productivity most often occur (Pilat et al. 2006; Castaldi 2009). For these reasons, the hierarchical organization of knowledge is certainly relevant in industrial firms.

Data regarding the extent of the market for KIBS in the NUTS 2 regions of Italy are obtained from the Italian National Institute of Statistics (ISTAT) database on firms (*Imprese*; *Risultati economici delle imprese—reg* and *Statistiche regionali sulla struttura delle imprese*). Data on the input–output (I–O) dependency of industrial sectors on KIBS sectors are derived from the Italian national I–O *use* tables maintained by ISTAT. Finally, the information relevant for the construction of the IVs (i.e., the diffusion of broadband connections among firms and the share of regional employment in cooperative businesses) is obtained from the ISTAT database "Territorial Indicators for Development Policies" (in particular the item "Information Society," and "Social Capital"). In Supporting Information S1: Section A.3, we provide additional information on our data sources and our procedure for constructing the proxies for the extent of the market for KIBS, the control variables and the IVs.

3.2 | Data on Firm Hierarchy

We need to operationalize in the data the two primary dimensions of a firm's hierarchy, namely the span of control and the number of layers. A layer (l) is conceived as a group of employees with comparable knowledge and who perform tasks with a similar level of authority. The RIL database allows us to track the number of employees in four occupational categories, consistently defined over the five waves of the survey: (i) managers, (ii) middle managers, (iii) clerical employees and white-collar workers, and (iv) blue-collar workers. In line with Caliendo, Monte and Rossi-Hansberg (2015) and Caliendo et al. (2020), we group and order those occupational categories into four layers:

- layer 1: workers, meaning all employees in occupational categories (iii) and (iv);[9]
- layer 2: middle-managers, meaning all employees in occupational category (ii);
- layer 3: managers, meaning all employees in occupational category (i); and
- layer 4: entrepreneur, or owner, who is, by definition, present in all firms (Caliendo and Rossi-Hansberg 2012).

After defining what constitutes a layer (l), we can construct proxies for the two dimensions of firm hierarchy. We count the number of nonempty[10] layers to measure the hierarchy's vertical dimension (Colombo and Delmastro 1999; Colombo and Delmastro 2004; Colombo and Grilli 2013; Caliendo, Monte, and Rossi-Hansberg 2015; Caliendo et al. 2020):

$$L_{ft} = \text{number of non-empty layers in firm} f \text{ in year} t. \quad (5)$$

Thus, L_{ft} is a count of the layers in firm f in year t and ranges from 1 (i.e., a one-layer firm or self-employed worker) to 4 (i.e., at least one employee in each layer).

We measure the horizontal dimension of the hierarchy of firm f in year t, B_{ft}, by calculating the average supervisors' span of control at the firm level. Using the information on employees across layers, we first calculate the span of control of supervisors at each layer l (B_{lft}) as the number of subordinates in layer $l-1$ per specialized supervisor. If the adjacent superior layer (l) is empty, then the ratio is calculated by dividing the number of employees in layer $l-1$ by the number of supervisors in the nonempty layer, $l+1$, up to the layer of the entrepreneur who, by definition, is present in all firms. After calculating B_{lft}, we take its average across the number of nonempty layers in firm f to obtain the average span of control of supervisors at the firm-level:[11]

$$B_{ft} = \frac{\sum_l B_{lft}}{L_{ft} - 1}. \quad (6)$$

By construction, firms with $L_{ft} = 1$ show $B_{ft} = 0$. We first include these firms in the empirical analysis, and later check the robustness of the results to the exclusion of them (Section 5.4.4).

To determine whether the classification of occupational categories in the RIL database is suitable for constructing proxies for the hierarchy of industrial firms in Italy, we present the distribution of firm size by the number of layers, as shown in Figure 2.

Firms with more layers tend to have more employees (Figure 2A) and higher turnover (Figure 2B). These findings are consistent with both theory (e.g., Caliendo and Rossi-Hansberg 2012) and empirical evidence from firms in France

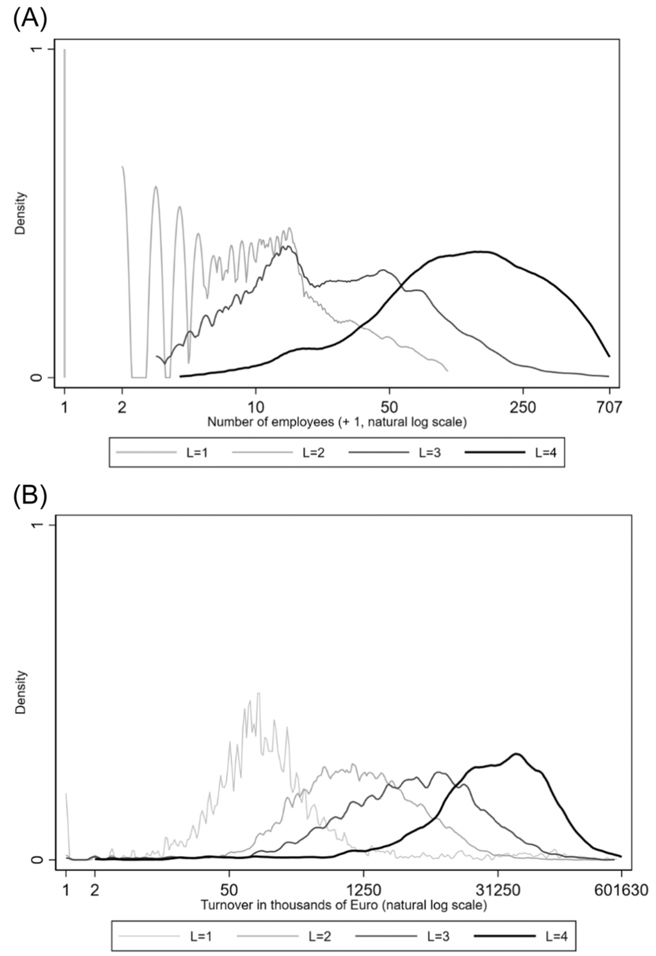

FIGURE 2 | Firm size distribution by number of layers, L_{ft}. (A) firm size is proxied by the number of employees and (B) firm size is proxied by turnover. [Color figure can be viewed at wileyonlinelibrary.com]

(Caliendo, Monte, and Rossi-Hansberg 2015) and Portugal (Caliendo et al. 2020; Cooke, Fernandes, and Ferreira 2021). In Figure 3 we graphically show the hierarchy of a representative industrial firm in Italy over the period 2004–2017.

In Figure 3, each layer is a rectangle, with a height equal to the percentage of firms in the sample with that layer. The length of layers 1, 2, and 3 is equal to the average number of employees therein. The typical firm thus takes a pyramidal shape, with lower layers having, on average, more employees, as predicted by Garicano (2000) and Caliendo and Rossi-Hansberg (2012). In our data, intermediate layers (i.e., managers and middle managers) are less frequent than the extreme ones (i.e., entrepreneur and workers). Back to the figure, the sum of the lengths of layers 1, 2, and 3 equals the average firm size (i.e., number of employees) in the sample, whereas the sum of the layers' height indicates the firm's depth. The representative hierarchy in the sample shows $L_{ft} = 2.313$ layers and an average span of control of $B_{ft} = 11.985$ subordinates per supervisor, as the arrows in Figure 3 show.

Table 2 provides some descriptive statistics regarding the number of firms, the horizontal and vertical dimensions of their hierarchies and their size in 2004, 2006, 2009, 2014, and 2017.

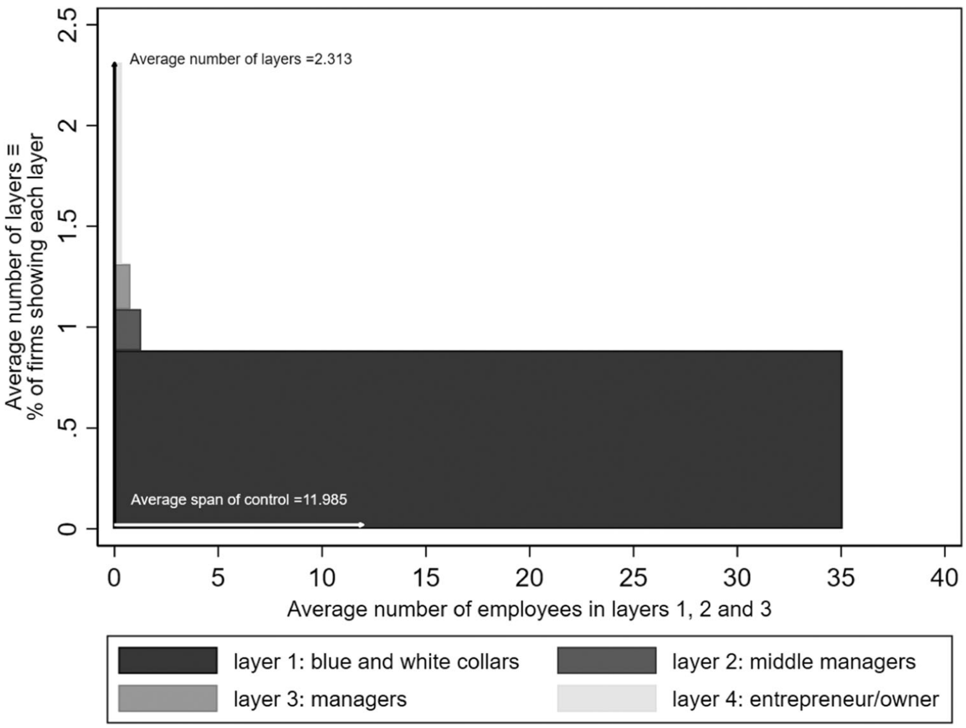

FIGURE 3 | The representative hierarchy in our sample; Italian industrial firms, 2004–2017.

TABLE 2 | Firm hierarchical dimensions and firm size; averages by wave of RIL.

RIL wave	Year	No. of firms	Span of control (B_{ft})	Number of layers (L_{ft})	Firm size (no. of employees)
2005	2004	8043	10.081	2.413	31.823
2007	2006	8897	10.305	2.138	27.657
2010	2009	8638	10.608	2.225	30.938
2015	2014	10,610	13.516	2.348	43.762
2018	2017	10,575	14.434	2.420	47.433
Total			11.985	2.313	37.106

From 2004 to 2017, the average number of layers (L_{ft}) in the industrial firms in Italy initially shrank before recovering, namely from 2.413 in 2004 to 2.225 in 2009, before rising back to 2.42 in 2017. The average span of control of supervisors, B_{ft}, increased. Those facts, combined with a general increase in firm size, indicate a flattening of industrial firms in Italy during the period analyzed.[12]

3.3 | Data on the Market for Knowledge

Although we do not directly observe the cost of the solutions (p in Section 2) that firms can buy in the market for knowledge, we use the information about the extent of the market for KIBS, in the (NUTS 2) region r where firm f is located, in terms of the share of regional employment that it represents, BS_SH_{rt}. We assume that the greater the extent of (i.e., the thicker) the market for KIBS, the greater the availability of solutions for firms located there and, in turn, the lower the cost p of buying solutions in the market for knowledge. This implies that the supervisor can save more time for controlling subordinates or can supervise a higher number of subordinates, given the time available.

KIBS are a particular group of services, which provides consulting activities to their clients and try to solve problems by leveraging on their knowledge and expertize (Muller and Zenker 2001). Solutions to clients' problems are provided by combining and applying various technologies and types of knowledge (Miles 2005). From an operational point of view, we adopt a standard definition of KIBS, commonly used by both policy makers (European Commission 2009) and scholars (Rubalcaba and Kox 2007; Schnabl and Zenker 2013; Evangelista, Lucchese, and Meliciani 2015). Based on this definition, KIBS gather the following (NACE rev 1.1) 2-digit industries: 72 (Computer and related activities), 73 (R&D) and 74 (Other business activities) for the years 2004 and 2006, and the (NACE rev 2) 2-digit industries 62 (Computer programming, consultancy and related activities), 63 (Information service activities), 69 (Legal and accounting activities), 70 (Activities of head offices; management consultancy activities), 71 (Architectural and engineering activities; technical testing and analysis), 72 (Scientific R&D), 73 (Advertising and market research), 74 (Other professional, scientific and technical activities), and 75 (Veterinary activities) for the years 2009, 2014, and 2017.[13]

To consider activities that could be bought by an industrial knowledge-based hierarchy to solve problems it faces in its production process, we perform a sensitivity analysis to the adoption of different (either stricter or broader) groupings of sectors among those listed above (see Section 5.4.3).

Because IT and CT have facilitated communication and the dissemination of knowledge, KIBS are arguably spatially neutral—that is, they may cover the entire country and not only one region. Even so, because face-to-face contact and tacit knowledge matter, business service providers most often deal with clients who require close, direct interaction to conceive solutions to their problems. Geographical proximity is thus relevant for the interactions between buyers and suppliers of knowledge, namely between industrial firms and KIBS (Koschatzky 1999; Muller and Zenker 2001; Meliciani and Savona 2015).

Figure 4 and Supporting Information S1: Table A.3 show the heterogeneity in the extent of the market for KIBS across the Italian NUTS 2 regions during the period 2004–2017.

Some northern regions (e.g., Piemonte, Lombardia) stand out among the ones with the highest shares of employees in KIBS. Lazio, the region of the capital city, is also highly intensive in KIBS. Smaller regions (e.g., Marche, Abruzzo, and Molise) are characterized by lower shares of KIBS.

Firms are heterogeneous in their technological exposure to the extent of the market for knowledge. Indeed, some firms belong to industries that use KIBS more intensively than others. We compute the coefficient w^{BS}_{j2004} from the I–O use table for Italy, which measures the share of intermediate costs of industry j corresponding to inputs provided by the KIBS sectors in 2004, the beginning of the period analyzed. The coefficient captures the dependence of industry j on business services, BS, and reflects technological (i.e., exogenous) differences across industries. Table 3 shows the cross-industry heterogeneity in the dependence of industrial firms on KIBS in 2004 (col. 1), and, on average, in the period 2004–2017 (col. 2).

Mining and quarrying is the sector most dependent on KIBS with approximately 12% of its inputs coming from KIBS sectors. Manufacture of machinery and equipment sector follows with 8.3% of inputs coming from those sectors in 2004 (and 9.8%, on average, in the period 2004–2017). By contrast, the supply and distribution of electricity, gas, and steam and water supply sector is least dependent on KIBS, with only 3.8% of its inputs coming from KIBS sectors in 2004 (and 4.4%, on average, in the period 2004–2017).

3.4 | Control Variables

Given the richness of the RIL database, we are able to control for a large set of firms' observable characteristics. The model shown in Section 2 drives the choice of those variables. Indeed, we want to identify the effect of the extent of the market for knowledge of B_{ft} and L_{ft}, in addition to the roles played by the key parameters of the model (i.e., λ, h, and c). We construct proxies for those parameters using some indicators. Both investments in R&D and the introduction of new products are expected to be associated with a less predictable production process (λ). Process innovation is expected to be related to a lower cost of communicating knowledge within the firm (h), while the share of temporary employees is expected to be positively associated with the learning cost (c) of solving problems. Firm size, in terms of the number of employees, controls for the expected positive relationship between hierarchical structure and the firm size.

We also include a vector of time-variant controls at the industry(j)−region(r) level, which is expected to be correlated to both dimensions of firm hierarchy and the extent of the market for knowledge. In particular, we control for (i) the percentage of firms that have invested in R&D activities, (ii) the median value of labor productivity (i.e., turnover per number of employees), (iii) the percentage of firms that exported part of their products and services, and (iv) the relative size (i.e., in terms of the number of employees) of industry j with respect to the regional economy.[14]

4 | Econometric Analysis

4.1 | Identification Strategy

As explained in Section 3, we do not observe the price (p) of the solutions purchased by industrial firms in the market for knowledge. For this reason, we use a proxy for the extent of the market for knowledge, namely the share of employment represented by KIBS in the NUTS 2 region (r) in which firm f is located ($\mathrm{BS\_SH}_{rt}$). We assume that the larger the extent of the market for KIBS, the greater the availability of solutions and, in turn, the lower the cost of buying them in the market.

Our ideal experiment, thus, is to compare two identical production processes, a first one that takes place in a region with a thicker market for knowledge and a second one that is located in a region with a thinner market for knowledge. Predictions 1 and 2 would make us expect, ceteris paribus, the production process located in the region with a ticker market of knowledge (higher BS_SH_{rt}) to show a larger average span of control (B_{ft}) and a lower number of layers (L_{ft}) with respect to the identical production process located in the region with a thinner market for knowledge.

Thus, a standard approach would be to estimate the following reduced form:

$$H_{ft} = \beta_0 + \beta_1 \mathrm{BS\_SH}_{rt} + V'_{rt}\theta + \alpha_r + \tau_t + \varepsilon_{ft} \tag{7}$$

in which the dependent variable $H_{ft} = \{B_{ft}, L_{ft}\}$ denotes either dimension of the hierarchy (i.e., the span of control or the number of layer) of firm f in year t, $\mathrm{BS\_SH}_{rt}$ is the proxy for the extent of the market for knowledge, V_{rt} is a vector of time-variant controls at the regional level, α_r and τ_t are, respectively region- and time-fixed effects and ε_{ft} is a standard error term.

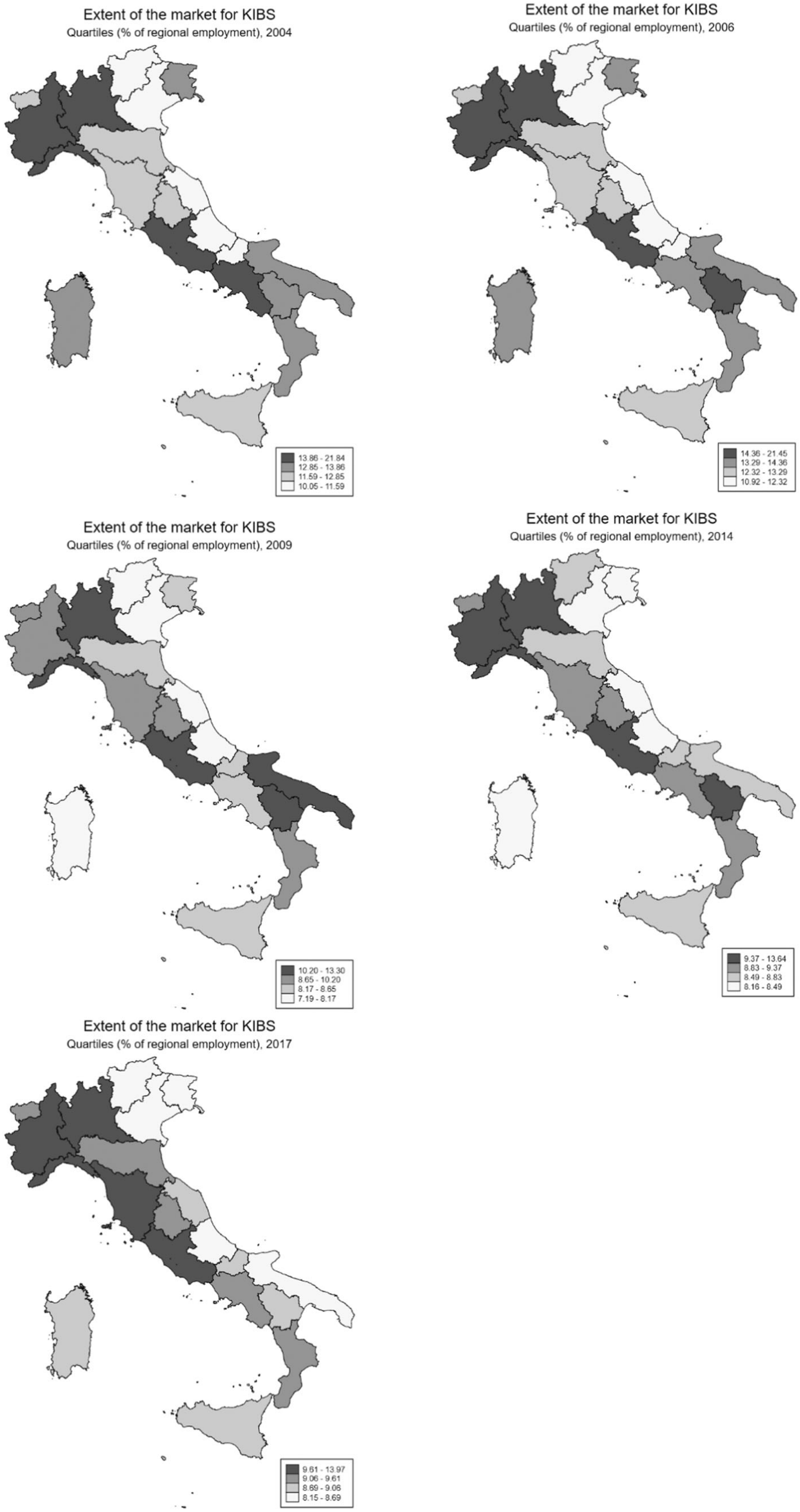

FIGURE 4 | Geographical distribution of KIBS, by NUTS 2 region, BS_SH_{rt}; Italy, selected years.

TABLE 3 | Share of intermediate costs of industry j, w_j^{BS}, that corresponds to the inputs provided by the KIBS.

Industry j	Coefficient of use of KIBS, by industry; 2004, w_{j2004}^{BS}	Coefficient of use of KIBS, by industry; average 2004–2017, $w_{\overline{j2004-2017}}^{BS}$
Mining and quarrying	0.116	0.122
Manufacture of food, beverages and tobacco	0.047	0.060
Manufacture of textile and wearing; wood; paper and reproduction	0.069	0.080
Manufacture of coke; chemicals; metals	0.067	0.059
Manufacture of machinery and equipment	0.083	0.098
Other manufacturing	0.057	0.068
Supply and distribution of electricity, gas, steam and water supply	0.038	0.044
Average across industries	0.067	

However, this analysis faces several identification challenges, which we detail in the following sections.

4.2 | Confounding Factors and Selection on Observables

First, the proxy for the extent of the market for knowledge, BS_SH$_{rt}$, varies at the regional−year level only. It is thus difficult to control for an exhaustive list of confounding factors that may affect both the development of the market for knowledge and firms' hierarchy at the region−year level: several of them may be unobserved. This may be the case of local labor market conditions, such as, the regional (mis-)match between labor demand and labor supply in terms of skills and occupations and wages' dynamics of workers with different knowledge.

The DiD identification strategy inspired by Rajan and Zingales (1998b) exploits the fact that if the thickness of the market for knowledge has a direct effect on firm hierarchy, this effect is likely to be larger in industries that are relatively more exposed to the local market for knowledge. For this reason, we assume that industries that use KIBS more intensively as inputs in their production processes are naturally (due to idiosyncratic demand and technological factors) more exposed to the extent of the local market for knowledge and are the "treatment" group. Conversely, industries that use of KIBS less intensively are "control" group. The impact of the extent of the market for knowledge of firms' hierarchy can be thus investigated by comparing differences in B_{ft} and L_{ft} between industries more exposed to the development of the local market for knowledge and their less exposed counterparts. We proxy the natural intensity in the use of KIBS as inputs with the coefficient w_{j2004}^{BS} calculated from the I–O use table. This coefficient measures the share of intermediate costs of industry j that correspond to inputs provided by the KIBS sectors at the beginning of the period analyzed (2004). w_{j2004}^{BS} is time-invariant and is obtained from I–O use table for the Italian economy (thus, it does not refer to any NUTS 2 region in particular) to reflect idiosyncratic differences across industries (Barone and Cingano 2011; Fiszbein et al. 2024).

We interact BS_SH$_{rt}$ with w_{j2004}^{BS} and insert it in our empirical model. Now that the main explanatory variable has an additional dimension (the industry, j), we can complement the year fixed effects with regional time trends to account for region-specific trends in the hierarchy. In the most demanding specification, we replace the regional time trends and the year fixed effects with a vector of region−year fixed effects to flexibly control for any unobserved time-varying confounder at the region−year level.

The coefficient of BS_SH$_{rt}$•w_{j2004}^{BS} is of our primary interest, because it captures the relative effect of BS_SH$_{rt}$ in the treatment group (industries with higher use of KIBS) with respect to the control group (industries with lower use of KIBS).[15] As suggested by Bassanini and Garnero (2013) and Bottasso, Conti, and Sulis (2017), this DiD approach only allows the identification of a differential effect of the extent of KIBS between more and less exposed industries, and not a direct one. Nonetheless, an indication on the direction of the rough average effect is provided, if one assumes further that in less exposed industries the effect is zero or of the same sign and smaller than in the more exposed industries.

A second identification challenge relates to the fact that localization into a region with a ticker or thinner market for knowledge may depend on firm characteristics (heterogeneity) that relate to the production processes conducted by industrial firms. We should control for all factors that make the two production processes comparable. Garicano (2000), model and the extension we propose in Section 2 guide our choice of observable covariates. Comparable production processes show similar degree of predictability (λ), similar cost for acquiring knowledge (c), and similar cost of communicating knowledge (h). The inclusion of the vector X_{ft} of firm-level controls, which gathers proxies for λ, h, and c serves this purpose (Angrist and Pischke 2009) and increases the precision of estimates. Moreover, to control for unobservable time-invariant characteristics of the production processes, we include a vector of fixed effects at the industry−region level, α_{jr}. In this way, we exploit the variability within industry–region pairs, where production processes should be more comparable.[16]

Third, to minimize the risk of an omitted variable bias, we control for a vector of time-variant controls (V_{jrt}) at the industry

(j)–region (r) level, which is expected to be correlated to both H_{ft} and w^{BS}_{j2004}•BS_SH$_{rt}$. In particular, we control for: (i) the share of firms that have invested in R&D activities, to control for differences in the technological intensity across industries and territories (Delmastro 2002); (ii) the median value of labor productivity, as an indirect and coarse measure of competition in a given industry–region (Cooke, Fernandes, and Ferreira 2021); (iii) the percentage of firms that export their products to control for the degree of openness (Guadalupe and Wulf 2010); (iv) the relative size of industry j with respect to the regional economy to account for market size, which affects the supply of skilled workers (Garicano and Hubbard, 2012).

We estimate using ordinary least square (OLS) the reduced form that identifies the role of the market for knowledge in the two dimensions of firm hierarchy (B_{ft} and L_{ft}):

$$H_{ft} = \beta_0 + \beta_1 \mathrm{BS\_SH}_{rt} + \beta_2 \mathrm{BS\_SH}_{rt} \bullet w^{\mathrm{BS}}_{j2004} + \beta_3 w^{\mathrm{BS}}_{j2004} + X'_{ft}\varphi + V'_{jrt}\theta + \alpha_{jr} + \tau_t + \varepsilon_{ft}, \quad (8)$$

in which the dependent variable H_{ft}, the covariate BS_SH$_{rt}$ and the error term are the same as in Equation (7), and firm-level controls, industry–region-level controls and vectors of fixed effects are defined as above. In all specifications, we cluster standard errors at the industry–region–year level—that is, the level of the "treatment" (BS_SH$_{rt}$•w^{BS}_{j2004})—as suggested by Abadie et al. (2023).[17] While β_1 can no longer be estimated when the region–year fixed effects are introduced and β_3 is absorbed by the industry–region fixed effects, the coefficient of the interaction term β_2, which is of our primary interest, can still be identified. β_2 tells us the differential effect of the extent of the market for knowledge, BS_SH$_{rt}$, in the two dimensions of firm hierarchy, between firms in industries that are more exposed to KIBS and firms in industries that are less exposed.

We expect $\beta_2 > 0 (< 0)$ when the dependent variable is the average span of control (the number of layers).

4.3 | Endogeneity: IV Approach

Given the possible correlation between the extent of the market for KIBS and unobserved shocks due to a measurement error in the explanatory variable, reverse causality, and omitted variables, all of which may generate endogeneity and biased estimates, we strengthen our empirical analysis with an IV approach. Due to classical measurement error (e.g., Wooldridge 2010), if BS_SH$_{rt}$ is measured with error, the OLS estimates will be biased toward zero (downward). As for reverse causality, firms with fewer layers and a larger span of control should, all else being equal, exhibit more predictable production processes (i.e., a higher λ, as indicated by Predictions 3 and 4 in Section 2).[18] Greater predictability implies a higher likelihood of encountering relatively simpler production problems that neither require a deeper hierarchy nor reliance on the market for knowledge to be solved, and can be addressed by the lower layers in the firm. Consequently, these firms would demand lower quantities of KIBS inputs, thereby leading to a lower number of transactions between industrial firms and KIBS providers and, consequently, a lower extent of the regional market for KIBS. Finally, regarding omitted variables that correlate with both the treatment (the extent of the market for knowledge) and the outcome (the hierarchy), we detailed at length in Section 4.2 our attempt to minimize such a problem through a set of industry–region time-variant characteristics that the theory proposed, V_{jrt}, and vectors of fixed effects. In addition, in Section 4.4, we use the framework developed by Altonji, Elder and Taber (2005a) and Oster (2019) to test the sensitivity of the estimated effect, $\hat{\beta}_2$, to unobservables and show that these are not a significant source of bias in $\hat{\beta}_2$ in our analysis.

Overall, due to measurement error and reverse causality in the context of this paper, we expect OLS estimates to be characterized by a downward bias. Following the previous arguments, both BS_SH$_{rt}$ and the interaction BS_SH$_{rt}$•w^{BS}_{j2004} are endogenous in Equation (8). The IV approach adds to Equation (8) the first-stage regression of BS_SH$_{rt}$•w^{BS}_{j2004}:

$$\mathrm{BS\_SH}_{rt} \bullet w^{\mathrm{BS}}_{j2004} = \alpha_0 + Z'_{rt-1} \bullet w^{\mathrm{BS}}_{j2004}\gamma + X'_{ft}\varphi + V'_{jrt}\theta + \alpha_{jr} + \tau_t + \omega_{rjt}. \quad (9)$$

Valid instruments (Z'_{rt-1}) need to correlate well with BS_SH$_{rt}$ (relevance), to be uncorrelated with disturbances in both Equations (8) and (9) (orthogonality) and to correlate with the two dimensions of H_{ft} only through BS_SH$_{rt}$ (exclusion restriction). We use information at the NUTS 2 level on the percentage of firms with more than 10 employees that have a broadband connection and on the share of regional employment in cooperative enterprises to obtain two IVs for BS_SH$_{rt}$.[19]

The diffusion of broadband connections among firms goes hand in hand with the development of a market for KIBS (Mack and Rey 2014; Tranos and Mack 2015). It is unlikely that firms' decisions about their hierarchies have had a direct impact on the diffusion of broadband connections within NUTS 2 regions. In fact, the trend in broadband access during the period under analysis was driven by European (supranational) targets (European Court of Auditors 2018).[20] Nonetheless, the exclusion restriction holds if the pervasiveness of the broadband connection has no direct effect on firm hierarchy independent of its relationship with the extent of the market for KIBS. Possibilities arise as to why the development of a broadband connection may affect—not through BS_SH$_{rt}$—one or both dimensions of H_{ft}. First, the broadband connection can be expected to expedite and lower the cost of interfirm exchange of knowledge (OECD 2008) and the cost of communicating knowledge (h) within the firm. For this reason, we control for the introduction of process innovations, which are expected to be related to a lower cost of communicating knowledge within the firm. Second, broadband connection may affect firm hierarchy through a booster for innovation and competition in the local product market (Akerman, Gaarder, and Mogstad 2015; Xu, Watts, and Reed 2019). As technology and competition have been shown to affect firms' choices about their hierarchy (Delmastro 2002; Guadalupe and Wulf 2010; Cooke, Fernandes, and Ferreira 2021; Belloc 2022), we control for measures of productivity and investments in R&D in both Equations (8) and (9). Meanwhile, the share of regional workers employed in the cooperative businesses is expected to be a valid instrument because it can be a proxy for the level of trust in the market relationships between firms in a given territory (Casadesus-Masanell and Khanna 2003; Sabatini, Modena, and Tortia 2014). At the same time, the relevance of cooperatives in a territory may be related to long-term factors that are hardly influenced by the organizational decisions of most of the firms in that area.[21]

4.4 | Sensitivity to Omitted Variables

Despite our efforts to include a relevant vector of time-varying controls, omitted variable bias would persist if the observed controls were an incomplete proxy for the true omitted variables. In this case, $\hat{\beta}_2$ is a biased estimate of the true effect, β_2. Therefore, we test the sensitivity of the estimated effect of the extent of the market for knowledge on firm hierarchy to unobservables. We adopt the method developed by Oster (2019), which is now common practice in empirical economics and builds on previous work by Altonji, Elder, and Taber (2005a). This method is based on the assumption that the relationship between the regressor of interest ($BS\_SH_{rt} \bullet w^{BS}_{j2004}$) and the omitted variables (W_{2t}) can be recovered from the relationship between the regressor of interest and the observables (X'_{ft}, V'_{jrt} in Equation (8) which we denote W_{1t} for simplicity of notation). A first key parameter in this approach is the proportional degree of selection, δ, which is the ratio of the magnitude of selection on unobservables to the magnitude of selection on observables.[22] The other key parameter is R^2_{long}, which is the R^2 from a hypothetical regression (*long regression*) of the dependent variable on the main regressor and both the observed (W_{1t}) and unobserved controls (W_{2t}). As a rule of thumb, Oster (2019) suggested choosing a value of $R^2_{\text{long}} = 1.3 \bullet R^2_{\text{medium}}$, where R^2_{medium} is the R^2 from the regression with observed controls (*medium regression*, Equation 8). Following Masten and Poirer (2023), we conduct the sensitivity analysis of our main results using two breakdown points of the parameter δ. First, we estimate the "explain away" breakdown point, $\hat{\delta}^{\text{bp,exact}}\left(R^2_{\text{long}}\right)$, often referred to as "Oster's delta," which has been widely reported in recent empirical papers (see Murtinu 2021, among others) as a measure of the robustness of the baseline model, as it indicates the smallest magnitude of the parameter δ that is consistent with $\beta_2 = 0$ (i.e., zero causal effect). For example, a value of $\hat{\delta}^{\text{bp,exact}}\left(R^2_{\text{long}}\right) = 2$, would suggests that the unobservables (W_{2t}) should be twice as important as the observables (W_{1t}) to produce $\beta_2 = 0$. In this respect, $\hat{\delta}^{\text{bp,exact}}\left(R^2_{\text{long}}\right) = 1$ has been considered an appropriate cutoff,[23] and estimated values above it would point to the main effect be robust to a substantial degree of selection on unobservables. Second, we estimate the "sign change" breakdown point, $\hat{\delta}^{\text{bp,sign}}\left(R^2_{\text{long}}\right)$, that is, the smallest value of δ such that β_2 shows the same sign as $\hat{\beta}_2$. As formally shown by Masten and Poirer (2023), the smallest value of delta for which β_2 has a different sign from $\hat{\beta}_2$ cannot be greater than one, and thus it may be much smaller than the "explain away" breakdown point in empirical applications. In other words, for omitted variables, it may be easier to reverse the sign of the coefficient of interest than to set it to zero. However, Masten and Poirer (2023) also show that this negative result can be overcome by hypothesizing a maximum value, M, for the omitted variable bias (i.e., the largest difference between the known $\hat{\beta}_2$ and the unknown β_2). For some values of M, the "sign change" breakdown point may well be larger than one. We follow Masten and Poirer (2023) and consider some multiples of $\hat{\beta}_2$ as possible values that M may take to check robustness of our estimates to sign changes.

5 | Results

5.1 | OLS Results

Based on Equation (8), we provide the baseline results for the effect of the extent of the market for knowledge on the average span of control, B_{ft}, in Table 4.[24]

In column 1, we include BS_SH_{rt} and $\text{BS\_SH}_{rt} \bullet w^{\text{BS}}_{j2004}$ together with a vector of industry–region-fixed effects and a vector of year dummies. The coefficient of BS_SH_{rt} is positive, indicating a positive relationship between the extent of the market for knowledge and the average span of control for a firm that is active in an industry whose business services intensity is zero. Given the negative coefficient of $\text{BS\_SH}_{rt} \bullet w^{\text{BS}}_{j2004}$, the relationship is weakened in industries that use KIBS intensively. However, the two coefficients are not statistically significant, which is probably due to omitted variables. To mitigate that risk and improve the precision of the estimates, we include in column 2 the vectors of controls at the industry–region and firm level, V'_{jrt} and X'_{ft}. While the coefficient on the extent of the market for KIBS is negative but not statistically different from zero (β_1), the coefficient of the interaction of BS_SH_{rt} with the intensity of use of business services by industries (β_2) is positive and significant. This means that the greater the availability of solutions from KIBS, the lower the average span of control, if the firm is active in an industry whose business services intensity is zero. This negative relationship is moderated and even reversed in industries that are highly intensive in the use of KIBS.

As shown in column 3 of Table 4, we control for linear time trends by region to account for specific trajectories in the span of control that firms in different NUTS 2 regions may have taken during the period analyzed. Column 4 includes both linear and quadratic region-specific time trends (Bitler and Carpenter 2016). As the coefficient for the extent of the market for KIBS is not statistically different from zero, the interaction term remains positive and significant. In column 5, we control for any remaining unobserved heterogeneity at the region–year level by including a vector of region–year fixed effects. That specification absorbs the coefficient of BS_SH_{rt}, while that of the interaction term β_2 can still be estimated. As shown in columns 2–5 in Table 4, as more controls are added, the magnitude of β_2 slightly decreases, but its estimate remains positive and significant.

With respect to column 5, the coefficient of 2.793 implies that the difference in terms of span of control (B_{ft}) between a highly KIBS-intensive industry (i.e., Manufacture of Machinery and Equipment, with $w^{\text{BS}}_{j2004} = 0.083$) and an industry that is slightly intensive in the use of KIBS (i.e., Manufacture of Food,

TABLE 4 | Average span of control, B_{ft}; OLS estimates.

	(1) (lin. reg.)	(2) (lin. reg.)	(3) (lin. reg. + lin tt)	(4) (lin. reg. + quad tt)	(5) (lin. reg.)	(6) (lin. reg.)
BS_SH_{rt}	0.163	−0.109	−0.113	−0.107		
	(0.159)	(0.096)	(0.095)	(0.095)		
$\text{BS\_SH}_{rt} \cdot w^{\text{BS}}_{j2004}$	−0.083	3.040***	2.819***	2.815***	2.793***	2.069**
	(1.710)	(0.959)	(0.878)	(0.871)	(0.844)	(0.926)
Investments in R&D$_{ft}$		−2.860***	−2.856***	−2.857***	−2.856***	−0.861***
		(0.263)	(0.263)	(0.263)	(0.263)	(0.257)
Product innovation$_{ft}$		−0.797***	−0.793***	−0.793***	−0.797***	−0.021
		(0.142)	(0.141)	(0.141)	(0.141)	(0.143)
Process innovation$_{ft}$		0.596***	0.589***	0.589***	0.587***	0.053
		(0.142)	(0.142)	(0.142)	(0.143)	(0.141)
% of temporary employees$_{ft}$		0.437	0.424	0.422	0.428	−0.142
		(0.289)	(0.288)	(0.288)	(0.289)	(0.274)
Firm size (#employees)$_{ft}$ (log + 1)		6.319***	6.316***	6.316***	6.318***	6.016***
		(0.110)	(0.110)	(0.110)	(0.110)	(0.195)
% of firms that invest in R&D$_{rjt}$		−0.705	−0.854	−0.640	−0.856	−2.141
		(1.366)	(1.221)	(1.233)	(1.257)	(1.447)
Median labor productivity$_{rjt}$ (log)		0.017	−0.069	−0.039	0.022	0.292
		(0.280)	(0.261)	(0.261)	(0.247)	(0.267)
% of exporters$_{rjt}$		2.975***	3.318***	3.337***	3.100***	0.872
		(0.966)	(0.914)	(0.915)	(0.982)	(0.955)
Industry relative size$_{rjt}$		−2.777*	−2.828**	−2.904**	−2.875**	0.200
		(1.430)	(1.247)	(1.245)	(1.275)	(1.225)
Constant	10.236***	−4.816***	−4.280***	98345.699*	−5.874***	−5.934***
	(1.112)	(1.546)	(1.532)	(53,067.245)	(1.285)	(1.376)
Industry−region FEs	Yes	Yes	Yes	Yes	Yes	No
Year FEs	Yes	Yes	Yes	Yes	No	No
Region time trends	No	No	Yes	Yes	No	No
Region−year FEs	No	No	No	No	Yes	Yes
Firm FEs	No	No	No	No	No	Yes
Adj. R^2	0.050	0.395	0.396	0.396	0.395	0.763
Log-likelihood	−191,936.7	−179,338.2	−179,306.9	−179,306.1	−179,293.7	−103,977.5
#Observations	46,763	46,241	46,241	46,241	46,241	32,883
#Industry−region−year	697	684	684	684	684	682

Note: Coefficients of industry−region FEs, year FEs, region time trends, region−year FEs and firm FEs are not reported to save space. Full tables are available from authors upon request. The subscripts *f, r, j*, and *t* denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry−region−year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Beverages and Tobacco, with $w^{\text{BS}}_{j2004} = 0.047$) is larger by about 0.371 employees per supervisor in a region at the 90th percentile of the distribution of BS_SH_{rt} in 2017 (value of about 12.3% of the regional employment) compared to a region at the 10th percentile of the distribution of BS_SH_{rt} in 2017 (value of about 8.61% of the regional employment).[25] If one assumes further that the extent of the market for KIBS has no effect in a hypothetical industry whose natural intensity in the use of KIBS is zero, we can also derive a rough estimate of the direct effect of the extent of the market for KIBS for the "average" industry, by multiplying $\hat{\beta}_2$ obtained in column 5 by the average sample value of w^{BS}_{j2004} (equal to 0.0674). When the extent of the market for knowledge increases by 1 percentage point,[26] the span of control B_{ft} of the firms in the "average" industry also increases by 0.188 employee per supervisor. That result indicates an average increase of 1.57% in the span of control relative to the sample's mean (11.985 layers, shown in Table 2). Thus, Prediction 1 is supported by our findings.

In column 6 of Table 4, we substitute industry–region-fixed effects with a vector of firm-fixed effects, α_f, to control for all time-invariant and unobserved characteristics of firms. Although the coefficient of $\text{BS\_SH}_{rt} \bullet w_{j2004}^{\text{BS}}$ remains positive and significant, it shrinks in magnitude and is less precisely estimated. A specification with firm-fixed effects is demanding because of the short time dimension and the low average number of observations per firm (i.e., about 1.8 in the sample of 46,241 observations reported in column 5) and the low within-firm variability of B_{ft} and L_{ft} over time. For these reasons, column 5 serves as our baseline specification.

Most of the correlations between the control variables and the span of control have the expected signs, supporting Prediction 3. Both investments in R&D and product innovation are negatively correlated with the average span of control, thereby indicating a lower degree of predictability (λ) in the production of firms that make these investments. Process innovation is positively and significantly associated with a larger span of control, as is consistent with the idea that process innovation may lead to lower costs of communicating knowledge (h). The higher the share of temporary employees, the larger the average span of control, even if the estimate is not statistically different from zero. If temporary contracts were a good proxy for a lower accumulation of knowledge to solve firm-specific problems, we would have expected a decrease in the span of control, via an increase in c. Nonetheless, temporary employees may be used when firms employ more routinized production processes, with a consequent need for a fewer supervisors per employee (Arrighetti et al. 2022). As expected, larger firms have larger span of control. The correlations between most of the industry–region control variables and B_{ft} also have the expected signs, although not all are statistically significant. In industry–region pairs in which most firms invest in R&D, the predictability of production (λ) is lower and the span of control thus smaller. Firms show a larger span of control in industry–region pairs characterized by a higher share of exporters, which may indicate the competitive pressure exerted by foreign markets in a firm's hierarchy, as suggested by Guadalupe and Wulf (2010). The median value of labor productivity is not statistically associated with the average span of control, and firms in industries that are relatively larger in the regional economy tend to have a narrower span of control.

In Table 5, we analyze the effect of the extent of the market for knowledge on the number of layers.

As shown in column 1, we include BS_SH_{rt} and $\text{BS\_SH}_{rt} \bullet w_{j2004}^{\text{BS}}$ together with a vector of industry–region fixed effects and a vector of year dummies. Despite a positive relationship between the extent of the market for KIBS and the number of layers for a firm that is active in an industry whose business services intensity is zero (as shown by the coefficient β_1), the relationship is moderated and even reversed in industries that use KIBS intensively. Those results may have been affected by omitted variable bias. As shown in column 2, we thus introduce the vectors V'_{jrt} and X'_{ft}. Whereas β_1 approaches zero in terms of magnitude and is not statistically significant, β_2 is statistically significant and captures the relative effect of the extent of the market for KIBS on the firm's number of layers in industry characterized by high intensity of the use of KIBS with respect to industry characterized by low intensity in the use of KIBS. Including region-specific time trends—to account for regional trajectories in firms' depth—, either linear (i.e., in column 3) or both linear and quadratic (i.e., in column 4) does not alter the results. The specification that better controls for region–year unobserved factors is the one in column 5, which includes a vector of region–year fixed effects. In that specification, the coefficient β_1 is absorbed by the fixed effects, and the coefficient β_2 remains negative and significant, as predicted by our theoretical framework in Section 2.

As shown in column 5 of Table 5, the coefficient of −0.114 implies that the difference in terms of the number of layers (L_{ft}) between an industry that is highly intensive in the use of KIBS (i.e., Manufacture of Machinery and Equipment) and an industry that is slightly intensive in the use of KIBS (i.e., Manufacture of Food, Beverage and Tobacco) is lower by about 0.015 layers in a region at the 90th percentile of the distribution of BS_SH_{rt} in 2017 with respect to a region at the 10th percentile of the distribution of BS_SH_{rt} in 2017. If one assumes further that the extent of the market for KIBS has no effect in a hypothetical industry whose natural intensity in the use of KIBS is zero, we can derive a rough estimate of the direct effect of the extent of the market for KIBS for the "average" industry, by multiplying $\hat{\beta}_2$ obtained in column 5 by the average sample value of w_{j2004}^{BS}. When the extent of the market for knowledge arises by 1 percentage point, L_{ft} of the firms in the "average" industry shrinks by 0.008 layers. That result indicates an average 0.35% decrease in the number of layers relative to the sample's mean (2.313 layers, shown in Table 2). To obtain an effect on the number of organizational layers comparable to those associated with pro-competitive market deregulation episodes and trade shocks (Cooke, Fernandes, and Ferreira 2021; Barba Navaretti et al. 2024), one should consider 10 percentage points increase in the extent of the regional market for knowledge.[27] This would correspond to a 0.08 reduction in the number of layers, or a 3.5% reduction. This result supports our Prediction 2. Last, we employ a linear model with firm-fixed effects to fully control for unobserved heterogeneity at the firm level (column 6). Similar to the evidence shown in Table 4, though the sign of the relationship is confirmed, the coefficient is far less precisely estimated and is not statistically significant. Thus, column 5 serves as our baseline specification. Controls show the expected sign and confirm Prediction 4. Ceteris paribus, investments in R&D and product innovation are associated with a higher number of layers (due to a lower predictability, λ), while the opposite is true for process innovation, which may reduce communication costs (h) within the firm. Larger firms and firms with a higher share of temporary workers have a higher number of layers. Among the controls at the industry region level, higher median productivity is associated with deeper hierarchies, in line with Garicano and Hubbard (2016). Larger industries (relative to the regional economy) show firms with a higher number of layers. Neither the share of firms investing in R&D nor the share of exporters are statistically associated with the number of layers.

TABLE 5 | Number of layers, L_{ft}; OLS estimates.

	(1) (lin. reg.)	(2) (lin. reg.)	(3) (lin. reg. + lin tt)	(4) (lin. reg. + quad tt)	(5) (lin. reg.)	(6) (lin. reg.)
BS_SH_{rt}	0.023**	0.005	0.005	0.004		
	(0.010)	(0.005)	(0.005)	(0.005)		
$BS\_SH_{rt} \cdot w^{BS}_{j2004}$	−0.353***	−0.138**	−0.117**	−0.117**	−0.114**	−0.071
	(0.119)	(0.061)	(0.057)	(0.057)	(0.053)	(0.067)
Investments in $R\&D_{ft}$		0.129***	0.129***	0.129***	0.130***	0.035***
		(0.010)	(0.010)	(0.010)	(0.010)	(0.012)
Product $innovation_{ft}$		0.035***	0.034***	0.034***	0.034***	−0.005
		(0.006)	(0.006)	(0.006)	(0.006)	(0.008)
Process $innovation_{ft}$		−0.032***	−0.032***	−0.032***	−0.032***	0.001
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
% of temporary $employees_{ft}$		0.117***	0.117***	0.117***	0.116***	0.076***
		(0.016)	(0.016)	(0.016)	(0.016)	(0.018)
Firm size (#employees) $_{ft}$ (log + 1)		0.406***	0.406***	0.406***	0.406***	0.444***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.009)
% of firms that invest in $R\&D_{rjt}$		0.068	0.026	0.014	−0.003	0.093
		(0.068)	(0.066)	(0.066)	(0.070)	(0.089)
Median labor $productivity_{rjt}$ (log)		0.038***	0.040***	0.038***	0.039***	0.013
		(0.014)	(0.013)	(0.013)	(0.013)	(0.019)
% of $exporters_{rjt}$		−0.055	−0.071	−0.072	−0.067	−0.017
		(0.051)	(0.046)	(0.046)	(0.049)	(0.060)
Industry relative $size_{rjt}$		0.057	0.068	0.072	0.083*	−0.047
		(0.053)	(0.047)	(0.047)	(0.048)	(0.061)
Constant	2.323***	1.136***	1.133***	−5211.605**	1.185***	1.198***
	(0.060)	(0.074)	(0.072)	(2546.961)	(0.071)	(0.101)
Industry–region FEs	Yes	Yes	Yes	Yes	Yes	No
Year FEs	Yes	Yes	Yes	Yes	No	No
Region time trends	No	No	Yes	Yes	No	No
Region−year FEs	No	No	No	No	Yes	Yes
Firm FEs	No	No	No	No	No	Yes
Adj.R^2	0.071	0.616	0.617	0.617	0.616	0.782
Log-likelihood	−55,480.71	−34,429.59	−34,391.95	−34,390.72	−34,375.11	−6941.69
#Observations	46,763	46,241	46,241	46,241	46,241	32,883
#Industry–region−year	697	684	684	684	684	682

Note: Coefficients of industry−region FEs, year FEs, region time trends, region−year FEs, and firm FEs are not reported to save space. Full tables are available from authors upon request. The subscripts *f*, *r*, *j*, and *t* denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry−region−year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

5.2 | IV Results

Table 6 presents the results of the specification with the vectors of industry−region and region−year-fixed effects. For ease of comparison, column 1 shows the OLS estimation for the average span of control, (i.e., column 5 in Table 4). As shown in column 2, the two-step GMM estimator is implemented. The *p* value of the Kleibergen−Paap *rk* LM test rejects the null hypothesis, which reassures us about the identification of the model. The Kleibergen–Paap Wald *rk* weak-identification test confirms that the relationship between the instruments and the potentially endogenous regressors is strong, given a remarkably

TABLE 6 | Average span of control and number of layers; IV estimates.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; iv-gmm2s)	(3) (B_{ft}; iv-liml)	(4) (B_{ft}; iv-cue)	(5) (L_{ft}; lin. reg.)	(6) (L_{ft}; iv-gmm2s)	(7) (L_{ft}; iv-liml)	(8) (L_{ft}; iv-cue)
$\text{BS\_SH}_{rt} \cdot w^{\text{BS}}_{j2004}$	2.793***	5.914***	5.865***	5.919***	−0.114**	−0.222***	−0.227***	−0.221***
	(0.844)	(1.354)	(1.358)	(1.354)	(0.053)	(0.072)	(0.072)	(0.072)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes[a]	Yes[a]	Yes[a]	Yes	Yes[a]	Yes[a]	Yes[a]
Region−year FEs	Yes	Yes[a]	Yes[a]	Yes[a]	Yes	Yes[a]	Yes[a]	Yes[a]
Adj. R^2	0.395	0.359	0.359	0.359	0.616	0.584	0.584	0.584
#Observations	46,241	46,241	46,241	46,241	46,241	46,241	46,241	46,241
#Industry–region–year	684	684	684	684	684	684	684	684
Underidentification test: Kleibergen–Paap *rk* LM stat. (*p*-value)		0.000	0.000	0.000		0.000	0.000	0.000
Weak identification test: Kleibergen–Paap *rk* Wald *F* stat.		218.374	218.374	218.374		218.374	218.374	218.374
Hansen *J* (*p* value)		0.661	0.661	0.661		0.250	0.250	0.251

Note: Coefficients of the control variables are not reported to save space.
[a]Industry–region FEs and region−year FEs are "partialled out" from all the other variables. Full tables are available from authors upon request. The subscripts f, r, j, and t denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

high F statistic (218.374).[28] Last, the instruments are valid, given a Hansen J test statistic showing that that overidentified restrictions are not rejected ($p = 0.661$).

The IV estimate of the coefficient of $\text{BS\_SH}_{rt} \bullet w^{\text{BS}}_{j2004}$ is larger (column 2: 5.914) than its OLS counterpart (column 1: 2.793), and the positive relationship between the span of control and the extent of the market for knowledge in industries that use KIBS intensively is confirmed. The results are consistent across different estimators such as the limited information maximum likelihood (LIML) and the continuously updated estimators (CUE).[29] In the case of column (2), the differential effect in terms of span of control between a highly KIBS-intensive and a slightly KIBS-intensive industry is larger by about 0.786 employees per supervisor in a region at the 90th percentile of the distribution of BS_SH_{rt} in 2017 than in a region at the 10th percentile of the distribution of BS_SH_{rt} in 2017. The IV estimates for the number of layers (columns (6)–(8)), L_{ft}, confirm what we have shown in Table 5, even showing a higher coefficient (in absolute value) than its OLS counterpart. Firms in industries that use KIBS intensively have a lower number of layers than firms that use KIBS less intensively, in regions with a larger extent of the regional knowledge market.

Overall, the IV estimates produce an effect of the extent of the market for knowledge on the firm's hierarchy that is almost twice as large (in absolute value) as the OLS estimates, both in terms of the average span of control and the number of layers. Following the arguments presented in Section 4.3, we can expect the direction of the OLS bias to be negative due to (i) classical measurement error in BS_SH_{rt} and (ii) the underlying economic motivation that, ceteris paribus, firms with a larger span of control and a lower number of layers demand lower quantities of KIBS inputs. This leads to a smaller regional market for KIBS. Overall, in Section 4.3, we have done our best to clarify the mechanisms of endogeneity that may affect the OLS estimates, and we have provided some explanations that are consistent with an underestimation of the true effect of the extent of the market for knowledge on firm hierarchy by OLS. Moreover, in Section 4.3, we also discussed how our two proposed IVs should satisfy the relevance and exclusion restriction conditions. In summary, the IV results confirm that, after controlling for the endogeneity of the extent of the market for knowledge due to measurement errors and reverse causality, a thicker market for knowledge is positively (negatively) associated with firms' average span of control (number of layers) in industries that are more exposed to the use of KIBS.

However, while the statistical test of the relevance of the IVs shown at the bottom of Table 6 gives encouraging results, we cannot rule out slight violations of the exclusion restriction (Aleksin and Becker 2024). Moreover, the IV estimates recover the effect only for the subset of firms whose probability of being treated is affected by the instrument (local average treatment effect). This may not be representative (Jiang 2017) of the population treatment effect (average treatment effect). For these reasons, the OLS estimates contained in Tables 4 and 5 are more conservative and remain our reference point.

5.3 | Sensitivity of the OLS Results to Omitted Variables

Table 7 shows the estimates of the two breakdown ("explain away" and "sign change") points of the sensitivity parameter δ, following the method proposed by Oster (2019).[30] The specifications of the

TABLE 7 | Sensitivity of the baseline results to omitted variables.

	"Explain away" breakdown point Oster's delta	"Sign change" breakdown point	
Dependent variable of the *medium regression*	$\hat{\delta}^{\text{bp,exact}}\left(R^2_{\text{long}}\right)$	$\hat{\delta}^{\text{bp,sign}}\left(R^2_{\text{long}}\right)$	$\hat{\delta}^{\text{bp,sign}}\left(R^2_{\text{long}}\right)$
		$M = +\infty$	$M = 10$
Average span of control (B_{ft})	−3.432	0.999	3.432
Number of layers (L_{ft})	1.794	1.000	1.794

Note: The specifications of the *medium regression* are those in column (5) of Table 4 and column (5) of Table 5. The subscripts f, r, j, and t denote firm, region (NUTS 2), industry and year, respectively. We use V'_{jrt} and X'_{ft} to calibrate the magnitude of selection on unobservables, while the vectors of industry−region and region−year fixed effects are included in the analysis but not used for calibration. As suggested by Oster (2019), we assume $R^2_{\text{long}} = 1.\ 3{\cdot}R^2_{\text{medium}}$. $\boldsymbol{M}$ is the maximum hypothesized omitted variable bias in the empirical model (i.e., the difference, in absolute value, between $\hat{\beta}_2$ and β_2).

medium regression are those in column (5) of Table 4 and column (5) of Table 5. We show estimates for $R^2_{\text{long}} = 1.3{\bullet}R^2_{\text{medium}}$.

The second raw of Table 7 refers to the baseline results on the average span of control, while the third raw refers to the baseline results on the number of layers. The second column shows the estimates of the "explain away" breakdown points, $\hat{\delta}^{\text{bp,exact}}\left(R^2_{\text{long}}\right)$, that is, "Oster's delta." For both outcome variables, the absolute values of these estimates are well above one, which is the commonly used cutoff for robust results. Thus, our baseline results would be considered robust. However, as emphasized by Masten and Poirer (2023), the smallest value of delta at which β_2 has a different sign from $\hat{\beta}_2$, $\hat{\delta}^{\text{bp,sign}}\left(R^2_{\text{long}}\right)$, may be smaller than $\hat{\delta}^{\text{bp,exact}}\left(R^2_{\text{long}}\right)$. For this reason, column (3) and (4) show the estimated "sign change" breakdown points. These are shown for two magnitudes of the hypothesized maximum omitted variable bias in our empirical model: $M = +\infty$ (no restriction) and $M = 10$. Without restrictions (column 3), the "sign change" breakdown points—as formally shown by Masten and Poirer (2023)—are bounded above by one, but our estimates are quite close to one in the case of B_{ft} (0.999) and equal to one in the case of L_{ft}. If the maximum omitted variable bias is constrained to be less than 10 (column 4), it is strong enough to imply that the "explain away" and "sign change" breakdown points take the same value. Overall, the sensitivity analysis we conducted shows that although the "explain away" breakpoints are larger than the "sign change" breakpoints without imposing restrictions on the magnitude of the omitted variable bias, our baseline results are robust to unobservables when using the conventional robustness cutoffs and imposing some restrictions on M.

5.4 | Addressing Other Sources of Endogeneity and Potential Generalizability of Results

In the following sections (5.4.1–5.4.5), we address additional sources of endogeneity that may bias the estimates in Tables 4 and 5. In Section 5.4.6, we discuss the generalizability of our results to other countries.

5.4.1 | The Geography of the Market for Knowledge and Multiplant Firms

The identification strategy undergirding our empirical analysis begins with the assumption that firms can choose to buy the knowledge needed to solve exceptional problems in the regional market in which they are located. Three potential issues can weaken this strategy. First, firms with scarce managerial resources and low predictability (λ) may re-locate their headquarters in a region precisely because it offers access to more developed KIBS. Such self-selection in the regional market for knowledge, based on the specific hierarchy of those firms (i.e., fewer layers and a larger span of control, coupled with a lower degree of predictability), may cause an additional endogeneity in our baseline estimates, which could lead to an upward bias in the OLS estimates. Second, multiplant firms may exploit their structure to access knowledge available in regions other than the one where they are headquartered. That circumstance may generate a bias in our estimates if the firms modify their hierarchies according to changes in the extent of the market for KIBS in other NUTS 2 regions. There could be an upward bias if changes in the share of KIBS among regions in which establishments of multiplant firms are located are positively correlated. Third, information on the location of the headquarters of industrial enterprises is available (in the RIL database) at the NUTS 2 level. Thus, the geographical dimension of BS_SH_{rt} is the region. Based on the argument that KIBS needs geographical proximity to the customer to be able to offer tailor-made solutions (Section 3.3), in some regions of Italy, it can make a difference whether a company is located in one province or another (e.g., in Lombardia, whether it is located in Milano or Sondrio). In other words, some Italian NUTS 2 regions may show greater heterogeneity in KIBS provision across NUTS 3 provinces.[31]

The first potential problem should not have affected our model and estimates, as we have shown in Tables 4 and 5 that firms with a larger average span of control and fewer layers are characterized by higher levels of predictability. Thus, relocation due to the above reason should be minimal in our sample of firms. Moreover, the data cleaning steps (Supporting Information S1: Section A.2) involved excluding firms that changed the region of their headquarters during 2004–2017. By contrast, we addressed the second potential problem by rerunning our baseline specification in the sample without multiplant firms to make this omitted variable bias less severe (Table 8).

In columns 2 and 4 of Table 8, we have restricted the analysis to firms with only one plant throughout the period in which they are observable. Indeed, a firm may change its hierarchy either vertically or horizontally to anticipate a future opening of a plant in another region and benefit from the extent of the market for knowledge therein. The number

TABLE 8 | Average span of control and number of layers; excluding multiplant firms.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg.; no multiplant firms)	(3) (L_{ft}; lin. reg.)	(4) (L_{ft}; lin. reg.; no multiplant firms)
$\text{BS\_SH}_{rt} \cdot w^{BS}_{j2004}$	2.793***	2.570**	−0.114**	−0.084
	(0.844)	(1.151)	(0.053)	(0.071)
Constant	−5.874***	−5.439***	1.185***	1.218***
	(1.285)	(1.648)	(0.071)	(0.097)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.503	0.616	0.569
#Observations	46,241	21,088	46,241	21,088
#Industry–region–year	684	669	684	669

Note: Coefficients of the control variables, industry–region FEs, and region–year FEs are not reported to save space. Full tables are available from authors upon request. The subscripts *f, r, j*, and *t* denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

TABLE 9 | Excluding the two NUTS2 regions with the highest standard deviation in BS across NUTS 3 provinces.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg.; no NUTS2 with high KIBS SD)	(3) (L_{ft}; lin. reg.)	(4) (L_{ft}; lin. reg.; no NUTS2 with high KIBS SD)
$\text{BS\_SH}_{rt} \cdot w^{BS}_{j2004}$	2.793***	3.293***	−0.114**	−0.090
	(0.844)	(1.182)	(0.053)	(0.066)
Constant	−5.874***	−5.278***	1.185***	1.106***
	(1.285)	(1.473)	(0.071)	(0.079)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.414	0.616	0.603
#Observations	46,241	36,091	46,241	36,091
#Industry–region–year	684	614	684	614

Note: In this table, we show the main coefficient of our empirical model, after having excluded the two NUTS 2 regions with the highest standard deviations in the extent of KIBS across their NUTS 3 provinces. The excluded regions (see Supporting Information S1: Figure A.1) are: Lombardia and Lazio. Coefficients of the control variables, industry–region FEs, and region–year FEs are not reported to save space. Full tables are available from authors upon request. The subscripts *f, r, j*, and *t* denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

of observations is drastically reduced, but the results are in line with our expectations. Once multiplant firms are excluded from the analysis, the effect of the regional market for knowledge is smaller in magnitude in relation to both the average span of control and the number of layers. However, the sign of the relationship is confirmed for both the average span of control and the number of layers, even if the coefficients are less precisely estimated due to the notable reduction in the number of observations. Finally, we address the third issue by excluding the two NUTS 2 regions that show the highest standard deviation in the extent of KIBS across subregional territories (NUTS 3 provinces)—Lombardia and Lazio—and rerunning the baseline model on this subset of observations.[32] Results are shown in Table 9. Despite the relevant decrease in the number of observations, the baseline results are confirmed both for the average span of control and for the number of layers, in terms of the sign of the coefficients (which is even larger in magnitude in the case of B_{ft}, column 2). Thus, our baseline results are not entirely affected by the heterogeneity in KIBS supply, which is certainly more relevant in some Italian NUTS 2 regions than in others.

5.4.2 | Knowledge and Firm Boundaries: Industrial Groups and Mergers and Acquisitions (M&As)

A firm's knowledge set may be broader than its boundaries. The market is the fundamental way to increase the availability of the knowledge that we have in mind in our paper. However, firms may be part of or be involved in extraordinary operations in which their physical and intangible assets are transferred to or consolidated with other firms (i.e., M&As). For this reason, it is important to check the robustness of our results by excluding firms that belong to a group or have been involved in M&As. The results presented in Tables 10 and 11 are reassuring and confirm our expectations, with the sole exception of the average span of control, when groups represent an alternative means of accessing knowledge. These results suggest that, mostly for the number of layers, firms participating in groups can gather knowledge from other firms in the group, bypassing both the market and the hierarchy.

5.4.3 | Different I–O Coefficients and Alternative Definitions of KIBS

As explained in Sections 3.3 and 4.2, we proxy the use of KIBS by firms with the I–O *use* coefficient w^{BS}_{j2004}, which refers to the beginning of the period analyzed, but not to any NUTS 2 region in particular. However, I–O coefficients in 2004 may suffer from idiosyncratic shocks that hamper their effectiveness in capturing technological differences across firms active in different industries. For that reason, we have rerun our baseline specification by including the $w^{\mathrm{BS}}_{\overline{j2004-2017}}$ coefficients, obtained by averaging the intensity of the use of KIBS by industry j in the years 2004, 2006, 2009, 2014, and 2017. Table 12 shows the results.

In column 2 of Table 10, the coefficient of $\mathrm{BS\_SH}_{rt} \bullet w^{\mathrm{BS}}_{\overline{j2004-2017}}$ is positive, although slightly smaller in magnitude compared to the baseline results (column 1), when considering the average

TABLE 10 | Average span of control and number of layers; excluding firms that belong to industrial groups.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg., no group)	(3) (L_{ft}; lin. reg.)	(4) (L_{ft}; lin. reg., no group)
$\mathrm{BS\_SH}_{rt} \cdot w^{\mathrm{BS}}_{j2004}$	2.793***	1.120	−0.114**	−0.150**
	(0.844)	(1.044)	(0.053)	(0.060)
Constant	−5.874***	−6.073***	1.185***	1.207***
	(1.285)	(1.480)	(0.071)	(0.084)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.506	0.616	0.523
#Observations	46,241	31,734	46,241	31,734
#Industry–region–year	684	672	684	672

Note: Coefficients of the control variables, industry–region FEs, and region–year FEs are not reported to save space. Full tables are available from authors upon request. The subscripts *f, r, j,* and *t* denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

TABLE 11 | Average span of control and number of layers; excluding firms involved in M&As.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg.; no M& As)	(3) (L_{ft}; lin. reg.)	(4) (L_{ft}; lin. reg.; no M& As)
$\mathrm{BS\_SH}_{rt} \cdot w^{\mathrm{BS}}_{j2004}$	2.793***	2.663***	−0.114**	−0.151***
	(0.844)	(0.890)	(0.053)	(0.056)
Constant	−5.874***	−6.321***	1.185***	1.229***
	(1.285)	(1.309)	(0.071)	(0.076)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.422	0.616	0.597
#Observations	46,241	43,044	46,241	43,044
#Industry–region–year	684	684	684	684

Note: Coefficients of the control variables, industry–region FEs, and region–year FEs are not reported to save space. The subscripts *f, r, j,* and *t* denote firm, region (NUTS 2), industry and year, respectively. Full tables are available from authors upon request. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

TABLE 12 | Average span of control and number of layers; alternative definition of KIBS and average (over-time) I–O coefficients.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg.; $w^{BS}_{\overline{j2004-2017}}$)	(3) (B_{ft}; lin. reg., alternative def. of KIBS)	(4) (B_{ft}; lin. reg., $w^{BS}_{\overline{j2004-2017}}$, alternative def. of KIBS)	(5) (L_{ft}; lin. reg.)	(6) (L_{ft}; lin. reg.; $w^{BS}_{\overline{j2004-2017}}$.)	(7) (L_{ft}; lin. reg., alternative def. of KIBS)	(8) (L_{ft}; lin. reg., $w^{BS}_{\overline{j2004-2017}}$, alternative def. of KIBS)
$\text{BS\_SH}_{rt} \cdot w^{\text{BS}}_{j2004}$	2.793***				−0.114**			
	(0.844)				(0.053)			
$\text{BS\_SH}_{rt} \cdot w^{\text{BS}}_{\overline{j2004-2017}}$		2.413***				−0.102**		
		(0.793)				(0.047)		
BS_SH_{rt} (alt.) $\cdot w^{\text{BS}}_{j2004}$			2.500***				−0.097**	
			(0.677)				(0.042)	
BS_SH_{rt} (alt.) $\cdot w^{\text{BS}}_{\overline{j2004-2017}}$ (alt.)				2.203***				−0.090**
				(0.674)				(0.040)
Constant	−5.874***	−5.717***	−5.474***	−5.259***	1.185***	1.180***	1.166***	1.160***
	(1.285)	(1.304)	(1.240)	(1.251)	(0.071)	(0.069)	(0.067)	(0.066)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.395	0.395	0.395	0.616	0.616	0.616	0.616
#Observations	46,241	46,241	46,241	46,241	46,241	46,241	46,241	46,241
#Industry–region	684	684	684	684	684	684	684	684

Note: Coefficients of the control variables, industry–region FEs and region–year FEs are not reported to save space. The subscripts *f, r, j*, and *t* denote firm, region (NUTS 2), industry and year, respectively. Full tables are available from authors upon request. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

intensity of KIBS use by industry over the period 2004–2017. Column 6 shows that the negative relationship between the extent of the market for KIBS and the number of layers is also confirmed, but the coefficient is somewhat smaller than in the baseline results (column 5).

As we have introduced in Section 3.3 and discussed at length in Supporting Information S1: Section A.3.1, we are also concerned about the change in the NACE industrial classification that occurred in 2007. Indeed, Schnabl and Zenker (2013) have pointed out that the change has had consequences for the taxonomy of sectors that enter in the group of KIBS. For that reason, they proposed a new taxonomy to be adopted after 2007. Following their work, we built a stricter taxonomy of business services that excludes NACE (rev. 2) industries 74 (Other professional, scientific and technical activities) and 75 (Veterinary activities). These services may indeed not obviously be part of a knowledge-based hierarchy whose base level is production. The results shown in columns 3, 4, 7, and 8 align with the baseline estimates, and, again, using the $w^{BS}_{j2004-2017}$ I–O coefficients, reduces the magnitude of the relationship between the extent of the market for KIBS and the hierarchical dimensions of the firm. With respect to the baseline results, the sign and statistical significance of the relationship is confirmed for both the average span of control and the number of layers.

5.4.4 | Excluding One-Layer Firms

In all previous runs of our empirical model, we include one-layer firms in our analysis, that is, firms with an average span of control equal to 0 and that can be considered nonstandard (an owner acting as a self-employed worker). Even if those firms represent a small share of observations in the sample (about 11%), we rerun the baseline specification by excluding those firms. Table 13 shows the results, which align with the baseline estimates. Thus, one-layer firms did not drive the main results of our analysis.

5.4.5 | The Extent of the Market for Knowledge and the Cost of Buying Solutions in the Market

We acknowledge that BS_SH_{rt} cannot capture market concentration. A high share of regional employment in KIBS could indicate the presence of numerous small companies competing with each other, potentially leading to lower prices for the business services they provide to industrial companies. Nevertheless, the same high share could also result from a few large KIBS companies exerting market power, thereby making the purchase of solutions in the market for knowledge more expensive.[33] While we cannot directly test and rule out the possibility that KIBS firms exert market power and increase the price of solutions in the market for knowledge, we can provide two pieces of evidence that suggest that our findings are not entirely driven by this issue. First, consider Figure 5, which shows the distribution of the average size (number of employees per firm) of firms active in KIBS in relation to that of all businesses that populate a regional economy in a given year, calculated from regional data.

The average size of KIBS firms is typically smaller than the average size of all firms in the regional economy, and the central tendency of the distribution in Figure 5 is around 60%. This suggests that KIBS are relatively smaller than their clients and, therefore, may not easily exert market or bargaining power over the buyers of their services. Second, we rerun our empirical model restricting the analysis to observations belonging to region–year pairs where the average firm size in KIBS is between 50% and 70% of the average firm size of all firms in the regional economy, as indicated by the vertical lines in Figure 5. This restriction allows us to consider only region–year pairs with similar (though not equal) relative size and bargaining power of KIBS, while still exploiting some variability in terms of BS_SH_{rt}. Higher values of BS_SH_{rt} should now be more closely associated with a higher number of KIBS firms relative to the number of all firms in the region, indicating more competition and a lower cost p. The results in Table 14 are encouraging.

TABLE 13 | Average span of control and number of layers; excluding one-layer firms.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg.; no 1-lyr firms)	(3) (L_{ft}; lin. reg.)	(4) (L_{ft}; lin. reg.; no 1-lyear firms)
$BS\_SH_{rt} \cdot w^{BS}_{j2004}$	2.793***	2.761***	−0.114**	−0.095*
	(0.844)	(0.885)	(0.053)	(0.055)
Constant	−5.874***	−8.311***	1.185***	1.330***
	(1.285)	(1.480)	(0.071)	(0.081)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.356	0.616	0.475
#Observations	46,241	41,095	46,241	41,095
#Industry–region–year	684	684	684	684

Note: Coefficients of the control variables, industry–region FEs, and region–year FEs are not reported to save space. Full tables are available from authors upon request. The subscripts f, r, j, and t denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

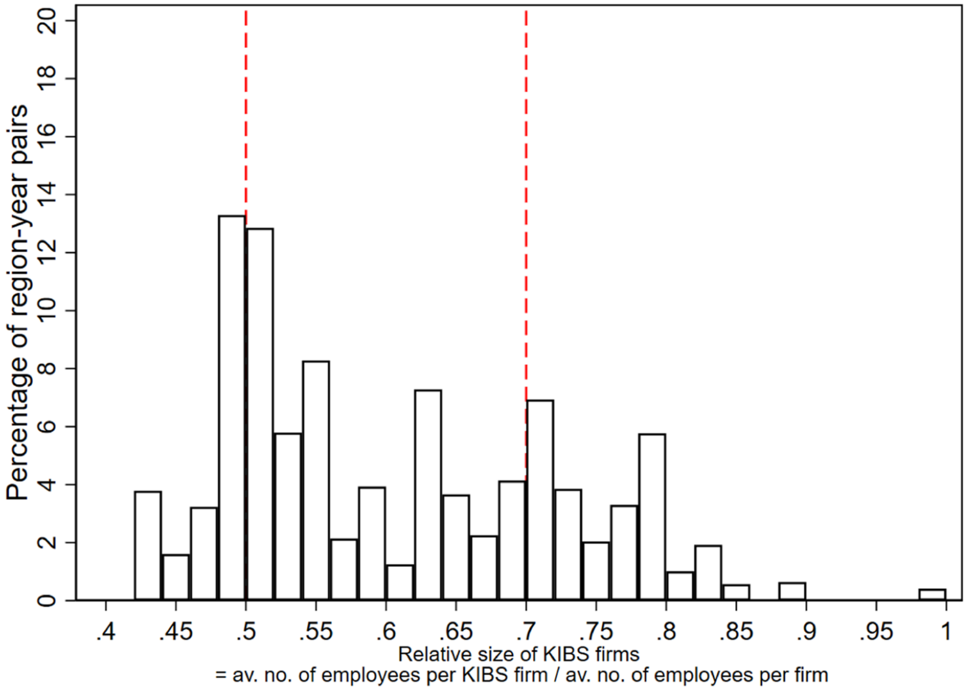

FIGURE 5 | Average size of KIBS firms in relation to that of all firms in the NUTS 2 region. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 14 | Average span of control and number of layers; excluding region–year pairs where the average size of KIBS firms is less than 50% or more than 70% of the average size of all firms in the regional economy.

	(1) (B_{ft}; lin. reg.)	(2) (B_{ft}; lin. reg.; restricted sample)	(3) (L_{ft}; lin. reg.)	(4) (L_{ft}; lin. reg.; restricted sample)
$\text{BS\_SH}_{rt} \cdot w^{BS}_{j2004}$	2.793***	5.643*	−0.114**	−0.077
	(0.844)	(2.906)	(0.053)	(0.116)
Constant	−5.874***	−5.200**	1.185***	1.106***
	(1.285)	(2.640)	(0.071)	(0.124)
Firm controls$_{ft}$	Yes	Yes	Yes	Yes
Industry–region controls$_{rjt}$	Yes	Yes	Yes	Yes
Industry–region FEs	Yes	Yes	Yes	Yes
Region–year FEs	Yes	Yes	Yes	Yes
Adj. R^2	0.395	0.385	0.616	0.618
#Observations	46,241	23,008	46,241	23,008
#Industry–region–year	684	356	684	356

Note: Coefficients of the control variables, industry–region FEs, and region–year FEs are not reported to save space. Full tables are available from authors upon request. The subscripts *f, r, j,* and *t* denote firm, region (NUTS 2), industry and year, respectively. Cluster (industry–region–year)—robust standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Indeed, the coefficient $\hat{\beta}_2$ in the regression when the dependent variable is the average span of control (column 2) is of the same sign and larger magnitude than the baseline estimates (column 1). For the number of layers, we obtain a coefficient with the same (negative) sign as the baseline estimate (column 4), with a smaller magnitude. The coefficients in columns (2) and (4) are less precisely estimated than those in the baseline estimates, but it should be kept in mind that the number of observations has been halved, making this robustness check very demanding. While we cannot rule out the possibility that a larger extent of the market for KIBS reflects a larger size for firms selling business services, our results are not entirely driven by this factor.

5.4.6 | Generalizability of the Results

In this section, we discuss the generalizability of our results to other countries. Few recent empirical papers allow us to compare some descriptive statistics about the hierarchies of firms in

TABLE 15 | Comparison with other studies on firms' hierarchy.

Paper	**Our study**	**Cooke, Fernandes and Ferreira (2021), Table B.2**	**Gumpert, Steimer and Antoni (2022), Table A.3**	**Caliendo, Monte and Rossi-Hansberg (2015), Table 3**
Data	Italian data (RIL), years: 2004, 2006, 2009, 2014, 2017; sectors: industrial firms (see Supporting Information S1: Table A.1)	Portuguese data, years: 2002–2009; sectors: manufacturing and services firms	German data, years: 2000–2010 sectors: all sectors	French data, years: 2002–2007; sectors: manufacturing firms
Number of layers, L_{ft}	**% of firm-year (%)**	**% of firm-year (%)**	**% of firm-year (%)**	**% of firm-year (%)**
$L_{ft} = 1$	11	10	32	18
$L_{ft} = 2$	59	52	40	28
$L_{ft} = 3$	18	25	27	35
$L_{ft} = 4$	12	13	2	19

Note: When comparing the percentage of firms with a given number of layers, it is important to note some differences in the definition of each layer across comparable studies. For Cooke, Fernandes, and Ferreira (2021), $\boldsymbol{l = 1}$ gathers production workers, $l = 2$ higher-skilled professionals, $l = 3$ intermediary executives, and $l = 4$ top executives; for Gumpert, Steimer, and Antoni (2022), $l = 1$ gathers production workers, $l = 2$ supervisors, $l = 3$ middle managers, and $l = 4$ CEOs; for Caliendo, Monte, and Rossi-Hansberg (2015), $l = 1$ gathers production workers, $l = 2$ supervisors, $l = 3$ senior staff, and $l = 4$ CEOs and firm directors.

our sample with those in other European countries. We refer to Caliendo, Monte, and Rossi-Hansberg (2015) for data on French firms, Cooke, Fernandes, and Ferreira (2021) for data on Portuguese firms, and Gumpert, Steimer, and Antoni (2022) for data on German firms. We focus on the vertical dimension of the hierarchy, the number of layers, L_{ft} (for which more information is available in these papers), by analyzing the percentage of firms characterized by different numbers of layers in different countries. Table 15 shows the comparison.

The majority of Italian firms in our sample have a number of layers, L_{ft}, equal to 2 (59% of observations). Similarly, the majority of Portuguese firms studied by Cooke, Fernandes, and Ferreira (2021) have a number of layers equal to 2 (52% of observations). In both countries, the least frequent values of L_{ft} are one layer (11% and 10% of the cases, respectively in Italy and Portugal) and four layers (12% and 13% of the cases, respectively in Italy and Portugal). Italian and Portuguese companies are, therefore, quite similar in terms of depth. Conversely, German and French hierarchies look different from Italian and Portuguese hierarchies. On the one hand, Germany has a higher percentage of enterprises with only one layer ($L_{ft} = 1$) and a lower percentage of enterprises with four layers ($L_{ft} = 4$) compared to Italy and Portugal. Thus, German enterprises are flatter than Italian and Portuguese enterprises. On the other hand, French companies have, on average, deeper hierarchies (about 54% of the companies have either three or four layers) compared to the other European firms. Based on the main findings of our study, we can expect KIBS to have a larger scope in providing solutions for German companies than for Italian, Portuguese and French companies. As shown in Schnabl and Zenker (2013), Table 4 and discussed in Deza and López (2014), in terms of employment, KIBS are more relevant in Germany than in Portugal and Italy, while they are similar in France.

However, generalizing the results to other countries must be done with great caution. In fact, our empirical analysis aims to assess the role of the market for knowledge once the predictability of the production process (λ), learning costs (c) and communication costs (h) (plus other controls) are taken into account. In other words, it is a *ceteris paribus*analysis. It is known that Italian firms are different, on average, from their French and German counterparts in terms of the complexity, innovativeness, and digital content of the adopted production processes (Bugamelli et al. 2018). These differences may explain, for example, why French firms have deeper hierarchies compared to Italian firms, even if KIBS are also well-developed in France.

6 | Conclusion

Deeper hierarchies offer the advantage of increasing the utilization rate of knowledge. Unsolved problems are passed upward through the hierarchy until the conditional probability of solving them is enough to bear the associated learning and communication costs. By adding layers of supervisors, the firm increasingly facilitates problem-solving in the productive process.

In our work, we add the market for knowledge to Garicano (2000), model and inquire into its role in shaping knowledge-based hierarchies. Instead of relaying exceptional and unsolved problems to supervisors or managers (i.e., "making" solutions via hierarchy), a firm may resort to the market for knowledge, typically by "buying" solutions from providers of specialized business services (KIBS). To our knowledge, that factor has not received enough attention, for it may affect firms' hierarchical structure in addition to other well-studied determinants of firm hierarchy. We predict that the cost of buying solutions in the market for knowledge will affect firm hierarchy (i.e., span of control and number of layers) and, in particular, that firms will be flatter as the extent of the market for knowledge increases due to the greater availability of solutions at lower cost.

We test our predictions with data from a rich database representing industrial firms in Italy from 2004 to 2017. The database offers information on firms' employment per occupational category and allows us to construct proxies for the supervisors' span of control and the number of layers in a firm. We enrich those firm-level data with information about the extent of the market for knowledge, proxied by the relevance of KIBS in the region where firms are headquartered. Using a DiD approach inspired by Rajan and Zingales (1998b), which exploits cross-regional variability in the development of the market for KIBS and cross-industry differences in the intensity of the use of KIBS as inputs, we find that a thicker regional market for knowledge is associated with flatter firms in industries that use KIBS more intensively compared to industries that use KIBS less intensively. Since measurement error in the explanatory variable, reverse causality, and omitted variables can lead to endogeneity and biased estimates, we strengthen our analysis with an IV approach. As IVs, we use information at the NUTS 2 level on the share of firms with more than 10 employees with broadband connections and the share of regional employment in cooperative businesses. Both OLS and IV estimates suggest that the extent of the market for knowledge in a given area affects an important economic outcome of firms located there, namely their hierarchical organization. The results are robust to a test of the sensitivity of the estimates to omitted variable bias and a series of robustness checks.

Acknowledgments

We are grateful to the Editor (Ramon Casadesus-Masanell), an anonymous coeditor and an anonymous reviewer for helpful comments. We are indebted to Filippo Belloc, Lorenzo Caliendo, Mauro Caselli, Stefano Dughera, Andrea Fracasso, Nicola Meccheri, Daniela Piazzalunga, Anna Rosso, Chiara Tomasi, Alexandra Tsvetkova, Francesco Venturini, Lynn Wu, and participants at the workshops *Firms and Workers at the Crossroad: Automation and Market Power* (Trento, December 2020), and *KIIS* (Valencia, November 2022) for their suggestions. We would like to express our gratitude to Prof. Oliver Hart for his encouragement in this project. We also thank the *Istituto Nazionale per l'Analisi delle Politiche Pubbliche* (INAPP) for providing us with the five waves of the *Rilevazione Longitudinale Imprese e Lavoro* (RIL) survey. Fabio Pieri: the topics covered in this paper are further developed in the framework of the EU—Next Generation EU/PRIN 2022 project entitled “REWIND—Resilient Enterprises and Workers: leveraging INtangibles to address Disruptions.” Massimiliano Vatiero gratefully acknowledges financial support from the “Brenno Galli” Fondo (Fondo Ricerca e Sviluppo, USI). The usual disclaimer applies. Open access publishing facilitated by Universita degli Studi di Trento, as part of the Wiley - CRUI-CARE agreement.

Data Availability Statement

The firm-level data are sourced from the RIL survey, accessible upon request from INAPP. To obtain access to the waves of the RIL survey, please follow the instructions for submitting a request at https://www.inapp.gov.it/rilevazioni/microdati. Third-party restrictions apply to the availability of these data, which have been used under license for this study. We integrate these firm-level data with regional and industry-level data that are publicly available from the Italian National Institute of Statistics, http://dati.istat.it. We can provide the.do files and output (.log) files to researchers interested in replicating our study.

Endnotes

[1] Garicano (2000) has provided the example of the shop floor of a production plant, where machinists deal with most problems related to the standard operation of the machines and where unexpected downtime involving exceptional problems may require the attention of a mechanical supervisor.

[2] There are a few notable exceptions. Garicano (2000) has provided a preliminary description of the market for knowledge in Section F of “Implications and Discussion” (p. 901), and Garicano and Rossi-Hansberg (2006, 2012) have offered an analysis of how a knowledge-based hierarchy can be decentralized into knowledge transactions. In Section 2, we refer to their theoretical framework and explain how it differs from ours.

[3] According to Muller and Doloreux (2009, 65), KIBS are “mainly concerned with providing knowledge-intensive inputs to the business process of other organizations”. Evangelista, Lucchese and Meliciani (2015, 965) have added that the sector “includes all industries that provide intangible intermediate inputs to the rest of the economy.”

[4] This empirical approach has been used in subsequent applications. See Bassanini and Garnerno (2013); Haltiwanger, Scarpetta, and Schweiger (2014); Andrews and Cingano (2014); Cardullo, Conti, and Sulis (2015); Franco, Pieri, and Venturini (2016); Bottasso, Conti, and Sulis (2017); Fiszbein et al. (2024), among others.

[5] In the theoretical literature, the hierarchical structure has also been studied and justified as a mechanism to lower transaction costs (Coase 1937), coordinate tasks (Williamson 1967), provide employers with power over their employees (Marglin 1974), furnish motivations (Calvo and Welisz 1978, 1979), gather information (Bolton and Dewatripont 1994), allocate risks and decision-making power (Wulf 2007; De Varo and Prasad 2015) and encourage specific investments (Hart 1995; Rajan and Zingales 1998a). Chen and Suen (2019) have provided a stylized representation of hierarchies that is general enough to be consistent with multiple explanations.

[6] For example, the 2005 wave contains information on firms that refer to the end of 2004 and, for some variables, to the period 2002–2004. The same lag applies to the other waves. In Supporting Information S1: Section A.2, we provide additional information regarding the representativeness of the database and the distribution of firms by industry and region.

[7] The Nomenclature of Units for Territorial Statistics (NUTS) indicates a classification of administrative areas used by the European Statistical Office (i.e., Eurostat). The RIL database contains information on the geographical location of the firm' headquarters at the NUTS 2 level, which, in Italy, correspond to regions. Additional information about the classification is available at https://ec.europa.eu/eurostat/web/nuts/overview.

[8] Using the NACE (rev 1.1) classification, we consider the following seven industrial sectors of the Italian economy: Mining and Quarrying (Section C); Manufacture of Food, Beverages, and Tobacco (Section DA); Manufacture of Textile and Wearing; Wood; Paper and Reproduction (Sections DB + DC + DD + DE); Manufacture of Coke, Chemicals, and Metals (Sections DF + DG + DH + DI + DJ); Manufacture of Machinery and Equipment (Sections DK + DL + DM); Other Manufacturing (Section DN); Supply and Distribution of Electricity, Gas, Steam, and Water Supply (Section E). In the RIL survey, the aggregated taxonomy of industries is based on the NACE rev. 1.1 (waves 2005 and 2007) and NACE rev. 2 (waves from 2010 onwards) statistical classifications. Using the conversion matrix proposed by Perani and Cirillo (2015), the 2-digit industry codes in the third, fourth and fifth waves of RIL have been converted to NACE Rev. 1.1 and then aggregated into the consistent taxonomy of seven sectors. Supporting Information S1: Table A.1 describes the distribution of firms across the industrial sectors.

[9] The decision to merge occupational categories (iii) and (iv) into a single layer is due to the similar knowledge that employees in those categories have (see also Caliendo, Monte, and Rossi-Hansberg 2015).

[10] A nonempty layer is one with a nonzero number of employees.

[11] In the theoretical model, we derive the effect of the market for knowledge on the span of control for a single layer l of the hierarchy,

B_{lft} (Prediction 1). In our empirical analysis, however, we examine the effect of the market for knowledge on the average span of control of the firm, as shown in Section 5.1, to gain insight into how firms change their average horizontal dimension of the hierarchy as the regional market for knowledge expands.

[12]Even if our work does not specifically address the phenomenon of hierarchical flattening (i.e., the long-term tendency towards hierarchies with fewer layers and a larger span of control), our descriptive evidence regarding the dynamics of the two hierarchical dimensions aligns with past findings from specific industries or countries (Colombo and Delmastro 1999; Delmastro 2002; Rajan and Wulf 2006; Guadalupe and Wulf 2010; Caliendo, Monte, and Rossi-Hansberg 2015).

[13]The double list (up to 2006 and from 2009 onwards) of sectors that we consider as KIBS is due to the change in the NACE industrial classification occurred in 2008, with the introduction of NACE rev.2 taxonomy. This change affects the identification of the sectors that are in the KIBS and produce a general downward shift in the share of KIBS common to all NUTS 2 regions over time (Schnabl and Zenker 2013). Supporting Information S1: Section A.3.1 explains this general shift (due to the occurred change in industrial classification) and how we have coped with that. In our analysis, we take that measurement issue into account by inserting year (or region–year) fixed effects in all specifications.

[14]In Supporting Information S1: Section A.3.2, we provide additional details about the definitions of the control variables, while Supporting Information S1: Table A.5 provides descriptive statistics regarding both firm and industry–region controls.

[15]In the context of cross-industry cross-country interaction models, this approach has recently been discussed in detail by Ciccone and Papaioannou (2023). A basic assumption is that the relevant industry characteristic must be independent of all country characteristics. If this is the case, the use of the industry characteristic of a benchmark country (usually the United States), which is an equally imperfect proxy for the technology industry characteristic of all other countries, introduces measurement error in its classical form, resulting in an *attenuation bias*. Conversely, if this assumption does not hold, the use of an industry characteristics of the benchmark-country may lead to biased results in the form of amplified estimates (*amplification bias*). However, as interestingly suggested by Fiszbein et al. (2024), this relevant concern should be less of an issue in our cross-regional (NUTS 2) setting, where the relevant industry characteristic, the intensity in the use of KIBS as inputs, w^{BS}_{j2004}, is measured at the country level and thus represents a national average. In a context of mobility of production factors across regions, a national industry characteristic should be a better proxy for the technology to which industrial firms have access.

[16]In some specifications, we substitute industry–region–fixed effects with a vector of firm fixed effects, α_f, to control for any time-invariant and unobserved firm characteristic. While desirable to get closer to the ideal experiment, a specification with firm fixed effects is very demanding in our setting, because of the short temporal dimension of the panel, the relatively low average number of observations per firm at our disposal, and the low within-firm variability of B_{ft} and L_{ft} over time.

[17]We provide evidence that our results are robust to different levels of clusterization (region–year; industry–year; region–industry; two-way cluster: region–year and industry–year; firm) in Supporting Information S1: Tables A.6 and A.7.

[18]These theoretical predictions are also supported by our empirical analysis. Assuming that investments in R&D and the introduction of new products are associated with less predictable production processes, the results in Tables 4 and 5 show that firms that invest in R&D and introduce new products tend to have a lower span of control and a higher number of layers, ceteris paribus.

[19]To further lessen simultaneity between the excluded instruments and the endogenous variables we lag both instrumental variables by 1 year with respect to BS_SH_{rt}. As explained in footnote 6, due to the time structure of the RIL survey, the 2005 wave contains data on firm hierarchy and other controls that refer to the year 2004, and we merge these data with the extent of the market for KIBS in 2004 and the instrumental variables in 2003. The same time structure applies to the other waves in the sample (i.e., 2007, 2010, 2015, and 2018).

[20]Evidence of the convergence toward European targets in terms of the percentage of firms with broadband connections is also appreciable in the Italian regions. Supporting Information S1: Figure A.2 shows the trend in the diffusion of broadband connections across regions during the period analyzed.

[21]Supporting Information S1: Figure A.3 shows cross-regional heterogeneity in the relevance of cooperative businesses.

[22]Oster (2019) definition of the proportional selection relationship would apply to our case as follows: $\delta \frac{\mathrm{Cov}\left(W_{1t}, BS\_SH_{rt} \cdot w^{BS}_{j2004}\right)}{\mathrm{Var}(W_{1t})} = \frac{\mathrm{Cov}\left(W_{2t}, BS\_SH_{rt} \cdot w^{BS}_{j2004}\right)}{\mathrm{Var}(W_{2t})}$.

[23]Both Oster (2019) and Altonji, Elder, and Taber (2005b) do not consider the researcher's choice of control variables (in our case X'_{ft}, V'_{jrt}) to be random, because it is reasonable to assume that the researcher initially chooses the most important controls (those that explain a significant proportion of the variation in the dependent variable) based on theory or previous empirical results. Therefore, $\delta = 1$ can be considered a reasonable cutoff.

[24]We use the Stata commands `reghdfe` and `ivreghdfe` developed by Correia (2014) to estimate the coefficients of the linear regression, and IV regressions.

[25]As suggested by Bottasso, Conti, and Sulis (2017), the exact formula of the differential effect is $\hat{\beta}_2 \cdot \left(w^{BS}_{ManMach\&Eq} - w^{BS}_{ManFoodBevTob}\right) \cdot \left(BS\_SH^{90}_{rt} - BS\_SH^{10}_{rt}\right)$. In the case of column 5 in Table 4, this back-on-the-envelope calculation is: $2.793 \cdot (0.083 - 0.047) \cdot (12.300 - 8.610) = 0.371$.

[26]The change—in terms of the share of regional employment represented by KIBS—corresponds to region Molise in 2006 (i.e., approximately 11.4% of regional employment) increasing the extent of the market for knowledge to the level of region Valle d'Aosta in 2006 (i.e., approximately 12.4%). The reader is cross-referred to Supporting Information S1: Table A.3.

[27]The change—in terms of the share of regional employment represented by KIBS—corresponds to region Veneto in 2004 (approximately 11.1% of regional employment) increasing the extent of the market for knowledge to the level of region Lazio in 2004 (approximately 21.8%). The reader is cross-referred to Supporting Information S1: Table A.3.

[28]Critical values tabulated by Stock and Yogo (2005) are well below the reported value.

[29]Both the LIML and CUE estimators are less affected by the weak instrument problem than the GMM estimator (Baum, Schaffer, and Stillman 2007).

[30]We use the Stata command `regsensitivity`, developed by Masten and Poirier (2023) and Diegert, Masten and Poirier (2023) to perform the sensitivity analysis of the baseline results to omitted variables.

[31]We thank an anonymous reviewer for bringing this issue to our attention.

[32]We refer the reader to Supporting Information S1: Section A.3.1.1 for a description of the data used to calculate the means and standard deviations of the extent of KIBS across NUTS 3 provinces within NUTS 2 regions.

[33]We thank an anonymous reviewer for bringing this issue to our attention.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *Quarterly Journal of Economics* 138, no. 1: 1–35.

Akerman, A., I. Gaarder, and M. Mogstad. 2015. "The Skill Complementarity of Broadband Internet." *Quarterly Journal of Economics* 130, no. 4: 1781–1824.

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005a. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113, no. 1: 151–184.

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005b. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling." *Journal of Human Resources* 40: 791–821.

Andrews, D., and F. Cingano. 2014. "Public Policy and Resource Allocation: Evidence From Firms in OECD Countries'." *Economic Policy* 29, no. 78: 253–296.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arora, A., A. Fosfuri, and A. Gambardella. 2001. *Markets for Technology: The Economics of Innovation and Corporate Strategy*. Cambridge: MIT Press.

Arrighetti, A., E. Bartoloni, F. Landini, and C. Pollio. 2022. "Exuberant Proclivity Toward Non-Standard Employment: Evidence From Linked Employer-Employee Data." *ILR Review* 75, no. 4: 1024–1053.

Atalay, E., A. Hortaçsu, and C. Syverson. 2014. "Vertical Integration and Input Flows." *American Economic Review* 104, no. 4: 1120–1148.

Barba Navaretti, G., L. Fontagné, G. Orefice, G. Pica, and A. C. Rosso. 2024. "TBTs, Firm Organization and Labor Structure." *Review of International Economics* 32, no. 3: 958–992. https://doi.org/10.1111/roie.12714.

Barone, G., and F. Cingano. 2011. "Service Regulation and Growth: Evidence From OECD Countries." *Economic Journal* 121, no. 555: 931–957.

Bassanini, A., and A. Garnero. 2013. "Dismissal Protection and Worker Flows in OECD Countries: Evidence From Cross-Country/Cross-Industry Data." *Labour Economics* 21: 25–41.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2007. "Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing." *Stata Journal: Promoting Communications on Statistics and Stata* 7, no. 4: 465–506.

Aleksin, G., and S. Becker 2024. "Using Instrumental Variables to Establish Causality." *IZA World of Labor 250*. https://doi.org/10.15185/izawol.250.v2.

Belloc, F. 2022. "Profit Sharing and Innovation Across Organizational Layers." *Journal of Economic Behavior & Organization* 197: 598–623.

Bitler, M. P., and C. S. Carpenter. 2016. "Health Insurance Mandates, Mammography, and Breast Cancer Diagnoses." *American Economic Journal: Economic Policy* 8, no. 3: 39–68.

Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen. 2014. "The Distinct Effects of Information Technology and Communication Technology on Firm Organization." *Management Science* 60, no. 12: 2859–2885.

Bolton, P., and M. Dewatripont. 1994. "The Firm as a Communication Network." *Quarterly Journal of Economics* 109: 809–839.

Bottasso, A., M. Conti, and G. Sulis. 2017. "Firm Dynamics and Employment Protection: Evidence From Sectoral Data." *Labour Economics* 48: 35–53.

Brusoni, S., A. Prencipe, and K. Pavitt. 2001. "Knowledge Specialization, Organizational Coupling, and the Boundaries of the Firm: Why Do Firms Know More Than They Make?" *Administrative Science Quarterly* 46, no. 4: 597–621.

Bugamelli, M., F. Lotti, M. Amici, et al., 2018. "No. 422—Productivity Growth in Italy: A Tale of a Slow-Motion Change." Questioni di Economia e Finanza (Occasional Papers), Bank of Italy.

Caliendo, L., G. Mion, L. D. Opromolla, and E. Rossi-Hansberg. 2020. "Productivity and Organization in Portuguese Firms." *Journal of Political Economy* 128, no. 11: 4211–4257.

Caliendo, L., F. Monte, and E. Rossi-Hansberg. 2015. "The Anatomy of French Production Hierarchies." *Journal of Political Economy* 123, no. 4: 809–852.

Caliendo, L., and E. Rossi-Hansberg. 2012. "The Impact of Trade on Organization and Productivity." *Quarterly Journal of Economics* 127, no. 3: 1393–1467.

Calvo, G. A., and S. Wellisz. 1978. "Supervision, Loss of Control, and the Optimum Size of the Firm." *Journal of Political Economy* 86, no. 5: 943–952.

Calvo, G. A., and S. Wellisz. 1979. "Hierarchy, Ability, and Income Distribution." *Journal of Political Economy* 87, no. 5: 991–1010.

Cardullo, G., M. Conti, and G. Sulis. 2015. "Sunk Capital, Unions and the Hold-Up Problem: Theory and Evidence From Cross-Country Sectoral Data." *European Economic Review* 76, pp. : 253–274.

Casadesus-Masanell, R., and T. Khanna. 2003. "Globalization and Trust: Theory and Evidence From Co-Operatives." William Davidson Institute Working Paper 592, University of Michigan.

Castaldi, C. 2009. "The Relative Weight of Manufacturing and Services in Europe: An Innovation Perspective." *Technological Forecasting and Social Change* 76, no. 6: 709–722.

Chen, C., and W. Suen. 2019. "The Comparative Statistics of Optimal Hierarchies." *American Economic Journal: Microeconomics* 11, no. 2: 1–25.

Ciccone, A., and E. Papaioannou. 2023. "Estimating Cross-Industry Cross-Country Interaction Models Using Benchmark Industry Characteristics." *Economic Journal* 133, no. 649: 130–158.

Coase, R. H. 1937. "The Nature of the Firm." *Economica* 4: 386–405.

Colombo, M. G., and M. Delmastro. 1999. "Some Stylized Facts on Organization and Its Evolution." *Journal of Economic Behavior & Organization* 40: 255–274.

Colombo, M. G., and M. Delmastro. 2004. "Delegation of Authority in Business Organizations: An Empirical Test." *Journal of Industrial Economics* 52, no. 1: 53–80.

Colombo, M. G., and M. Grilli. 2013. "The Creation of a Middle-Management Level By Entrepreneurial Ventures: Testing Economic Theories of Organizational Design." *Journal of Economics & Management Strategy* 22, no. 2: 390–422.

Conner, K. R., and C. K. Prahalad. 1996. "A Resource-Based Theory of the Firm: Knowledge Versus Opportunism." *Organization Science* 7, no. 5: 477–501.

Cooke, D., A. P. Fernandes, and P. Ferreira. 2021. "Entry Deregulation, Firm Organization, and Wage Inequality." *International Journal of Industrial Organization* 77: 102763.

Correia, S. 2014. "reghdfe: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-dimensional Fixed Effects." Statistical Software Components S457874, Boston College Department of Economics.

Delmastro, M. 2002. "The Determinants of the Management Hierarchy: Evidence From Italian Plants." *International Journal of Industrial Organization* 20: 119–137.

Deza, X. V., and M. G. López. 2014. "Regional Concentration of Knowledge-Intensive Business Services in Europe." *Environment and Planning C: Government and Policy* 32, no. 6: 1036–1058.

Diegert, P., M. A. Masten, and A. Poirier, 2023. *Assessing Omitted Variable Bias When the Controls Are Endogenous*. https://doi.org/10.48550/arXiv.2206.02303.

Duranton, G., and D. Puga. 2005. "From Sectoral to Functional Urban Specialisation." *Journal of Urban Economics* 57, no. 2: 343–370.

European Commission. 2009. "Challenges for EU Support to Innovation in Services—Fostering New Markets and Jobs Through Innovation." *Commission Staff Working Document* SEC 1195.

European Court of Auditors. 2018. "Broadband in the EU Member States: Despite Progress, Not All the Europe 2020 Targets Will Be Met" (European Court of Auditors Special Report 12/2018).

Evangelista, R., M. Lucchese, and V. Meliciani. 2015. "Business Services and the Export Performances of Manufacturing Industries." *Journal of Evolutionary Economics* 25: 959–981.

Fiszbein, M., J. Lafortune, E. G. Lewis, and J. Tessada. 2024. "Powering Up Productivity: The Effects of Electrification on U.S. Manufacturing." NBER Working Paper 28076. http://www.nber.org/papers/w28076.

Franco, C., F. Pieri, and F. Venturini. 2016. "Product Market Regulation and Innovation Efficiency." *Journal of Productivity Analysis* 45: 299–315.

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

Garicano, L. 2010. "Policemen, Managers, Lawyers: New Results on Complementarities Between Organization and Information and Communication Technology." *International Journal of Industrial Organization* 28: 355–358.

Garicano, L. and T. N. Hubbard. 2012. "Learning About the Nature of Production From Equilibrium Assignment Patterns." *Journal of Economic Behavior & Organization* 84, no. 1: 136–153.

Garicano, L., and T. N. Hubbard. 2016. "The Returns to Knowledge Hierarchies." *Journal of Law, Economics & Organization* 32, no. 4: 653–687.

Garicano, L., and E. Rossi-Hansberg. 2004. "Inequality and the Organization of Knowledge." *American Economic Review* 94, no. 2: 197–202.

Garicano, L., and E. Rossi-Hansberg. 2006. "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics* 121: 1383–1435.

Garicano, L., and E. Rossi-Hansberg. 2012. "Organizing Growth." *Journal of Economic Theory* 147: 623–656.

Garicano, L., and E. Rossi-Hansberg. 2015. "Knowledge-Based Hierarchies: Using Organizations to Understand the Economy." *Annual Review of Economics* 7: 1–30.

Garicano, L., and Y. Wu. 2012. "Knowledge, Communication, and Organizational Capabilities." *Organization Science* 23, no. 5: 1382–1397.

Grant, R. M. 1996. "Toward a Knowledge-Based Theory of the Firm." *Strategic Management Journal* 17: 109–122.

Guadalupe, M., and J. Wulf. 2010. "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies." *American Economic Journal: Applied Economics* 2, no. 4: 105–127.

Gumpert, A., H. Steimer, and M. Antoni. 2022. "Firm Organization With Multiple Establishments." *Quarterly Journal of Economics* 137, no. 2: 1091–1138.

Halac, M., and A. Prat. 2016. "Managerial Attention and Worker Performance." *American Economic Review* 106, no. 10: 3104–3132.

Haltiwanger, J., S. Scarpetta, and H. Schweiger. 2014. "Cross Country Differences in Job Reallocation: The Role of Industry, Firm Size and Regulations." *Labour Economics* 26: 11–25.

Hart, O. 1995. *Firms, Contracts, and Financial Structure*. New York: Oxford University Press.

Jiang, W. 2017. "Have Instrumental Variables Brought Us Closer to the Truth." *Review of Corporate Finance Studies* 6, no. 2: 127–140.

Kogut, B., and U. Zander. 1992. "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology." *Organization Science* 3, no. 5: 383–397.

Koschatzky, K. 1999. "Innovation Networks of Industry and Business-Related Services—Relations Between Innovation Intensity of Firms and Regional Inter-Firm Cooperation." *European Planning Studies* 7, no. 6: 737–757.

Mack, E. A., and S. J. Rey. 2014. "An Econometric Approach for Evaluating the Linkages Between Broadband and Knowledge Intensive Firms." *Telecommunications Policy* 38, no. 1: 105–118.

Marglin, S. A. 1974. "What Do Bosses Do?: The Origins and Functions of Hierarchy in Capitalist Production." *Review of Radical Political Economics* 6: 60–112.

Masten, M. A., and A. Poirer. 2023. *The Effect of Omitted Variables on the Sign of Regression Coefficients*. https://doi.org/10.48550/arXiv.2208.00552.

Meliciani, V., and M. Savona. 2015. "The Determinants of Regional Specialisation in Business Services: Agglomeration Economies, Vertical Linkages and Innovation." *Journal of Economic Geography* 15, no. 2: 387–416.

Miles, I. 2005. "Innovation in Services." In *The Oxford Handbook of Innovation*, edited by J. Fagerberg, D. Mowery, and R. Nelson, 433–458. Oxford: Oxford University Press.

Muller, E., and D. Doloreux. 2009. "What We Should Know About Knowledge-Intensive Business Services." *Technology in Society* 31, no. 1: 64–72.

Muller, E., and A. Zenker. 2001. "Business Services As Actors of Knowledge Transformation: The Role of KIBS in Regional and National Innovation Systems." *Research Policy* 30: 1501–1516.

Murtinu, S. 2021. "The Government Whispering to Entrepreneurs: Public Venture Capital, Policy Shifts, and Firm Productivity." *Strategic Entrepreneurship Journal* 15: 279–308.

O'Farrell, P. N., and L. A. R. Moffat. 1995. "Business Services and Their Impact Upon Client Performance: An Exploratory Interregional Analysis." *Regional Studies* 29, no. 2: 111–124.

OECD. 2008. "Broadband and the Economy." (Ministerial Background Report DSTI/ICCP/IE(2007)/FINAL).

Oster, E. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37, no. 2: 187–204.

Perani, G., and V. Cirillo. 2015. "Matching Industry Classification: A Method for Converting NACE Rev. 2 to NACE Rev. 1." Working Paper Series in Economics, Mathematics and Statistics, University of Urbino.

Pilat, D., A. Cimper, K. B. Olsen, and C. Webb. 2006. "The Changing Nature of Manufacturing in OECD Economies." OECD Science, Technology and Industry Working Papers, 2006/09, Paris: OECD Publishing.

Rajan, R., and L. Zingales. 1998a. "Power in a Theory of the Firm." *Quarterly Journal of Economics* 113: 387–432.

Rajan, R., and L. Zingales. 1998b. "Financial Dependence and Growth." *American Economic Review* 88, no. 3: 559–586.

Rajan, R. G., and J. Wulf. 2006. "The Flattening Firm: Evidence From Panel Data on the Changing Nature of Corporate Hierarchies." *Review of Economics and Statistics* 88, no. 4: 759–773.

Rajan, R. G., and L. Zingales. 2001. "The Firm As a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms." *Quarterly Journal of Economics* 116: 805–851.

Rubalcaba, L., and H. Kox. 2007. *Business Services in European Economic Growth*. Palgrave-MacMillan.

Sabatini, F., F. Modena, and E. Tortia. 2014. "Do Cooperative Enterprises Create Social Trust?" *Small Business Economics* 42: 621–641.

Schnabl, E., and A. Zenker. 2013. "Statistical Classification of Knowledge-Intensive Business Services (KIBS) With NACE Rev. 2." evoREG Research Note#25, Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe.

Stock, J. H., and M. Yogo. 2005. "Testing for Weak Instruments in Linear iv Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothemberg*, Chap. 5, edited by D. W. K. Andrews, and J. H. Stock, 80–108. Cambridge: Cambridge University Press.

Tranos, E., and E. A. Mack. 2015. "Broadband Provision and Knowledge-Intensive Firms: A Causal Relationship?" *Regional Studies* 50, no. 7: 1113–1126.

De Varo, J., and S. Prasad. 2015. "The Relationship Between Delegation and Incentives Across Occupations: Evidence and Theory." *Journal of Industrial Economics* 63, no. 2: 279–312.

Williamson, O. E. 1967. "Hierarchical Control and Optimum Firm Size." *Journal of Political Economy* 75, no. 2: 123–138.

Williamson, O. E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press.

Wulf, J. 2007. "Authority, Risk, and Performance Incentives: Evidence From Division Manager Positions Inside Firms." *Journal of Industrial Economics* 55, no. 1: 169–196.

Xu, X., A. Watts, and M. Reed. 2019. "Does Access to Internet Promote Innovation? A Look at the U.S. Broadband Industry." *Growth and Change* 50: 1423–1440.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.