Technological Change and Domestic Outsourcing*

Antonin Bergeaud Clément Malgouyres
Clément Mazet-Sonilhac Sara Signorelli

August 2021

Abstract

Domestic outsourcing has grown substantially in developed countries over the past two decades. This paper addresses the question of the technological drivers of this phenomenon by studying the impact of the staggered diffusion of broadband internet in France during the 2000s. Our results confirm that broadband technology increases firm productivity and the relative demand for high-skill workers. Further, we show that broadband internet led firms to outsource some non-core occupations to service contractors, both in the low and high-skill segments. In both cases, we find that employment related to these occupations became increasingly concentrated in firms specializing in these activities, and was less likely to be performed in-house within firms specialized in other activities. As a result, after the arrival of broadband internet, establishments become increasingly homogeneous in their occupational composition. Finally, we provide suggestive evidence that high-skill workers experience salary gains from being outsourced, while low-skill workers lose out.

JEL classification: G14, G21, O33
Keywords: Broadband, Firm organization, Labor market, Outsourcing.

*Addresses: Bergeaud: Banque de France and CEP; Malgouyres: Institut des Politiques Publiques and PSE; Mazet-Sonilhac: Banque de France and Sciences Po; Signorelli: University of Amsterdam. We would like to thank without implicating Philippe Aghion, Adrien Bilal, Johannes Boehm, Mathilde Le Moigne and seminar participants at College de France, Sciences Po, X-CREST, LAGV, JMA, AFSE and RCEA for helpful comments. Corresponding author: Antonin Bergeaud: 31 rue croix des petits champs 75001 Paris, FR. Email: antonin.bergeaud@banque-france.fr

1 Introduction

The implications of international outsourcing (or offshoring), i.e., the process of using a third-party firm based abroad to perform services that would otherwise be performed in-house by local employees, have been at the center of recurrent public debates and the focus of a large body of economic research and policy analysis.[1] More recently, empirical work has highlighted the pervasiveness of *domestic* outsourcing among developed economies (Dube and Kaplan, 2010; Goldschmidt and Schmieder, 2017; Bernhardt et al., 2016; Dorn et al., 2018) and has started to document its implications for the distribution of earnings. However, this nascent literature is silent on the factors underlying these trends and in particular, on whether the rising prevalence of domestic outsourcing is primarily driven by changes in (labor market) institutions or whether the adoption of new technologies played a role in this development.

In this paper, we look at the role played by innovation in Information and Communication Technology (ICT) in fostering the rise of domestic outsourcing. To do so, we estimate the impact associated with the diffusion of a new General-Purpose Technology (GPT), namely broadband internet (BI hereafter), on employment outsourcing by French firms. As explained in Abramovsky and Griffith (2006), GPTs such as ICTs improve the adaptability of firms which results in important organizational changes. In particular, to the extent that BI has reduced communication and transaction costs, its diffusion is likely to have modified the optimal boundaries of firms and affected the terms of the make-or-buy decisions that they face (see e.g. Lewis and Sappington, 1991; Garicano and Rossi-Hansberg, 2006; Bloom et al., 2014; Aghion et al., 2019b). It is therefore natural to assess the impact it has had on the allocation of workers to firms, in particular through the rise of domestic outsourcing. In addition, BI is likely to directly improve the productivity of the impacted firms through various channels, in a way that is biased towards workers with a more formal education (Akerman et al., 2015) and those employed in more complementary occupations (Aghion et al., 2019a).

To better understand the global impact of BI on the propensity to outsource labor, we build a partial equilibrium model in which firms are defined as a collection of different tasks that can either be performed in-house or by outsourced workers. In this model, BI represents a positive shock on a representative firm's revenue productivity and/or

[1] See Hummels et al. (2018) for a review of recent empirical works and Biscourp and Kramarz (2007) for a seminal work on French administrative data.

a reduction in the cost of outsourcing.[2] We show that, in the presence of imperfections in the labor market and under a broad set of conditions, BI expansion leads firms to outsource a greater share of their labor force—so that the ratio of outsourcing expenditure over sales or in-house labor cost goes up—and become more specialized in terms of their in-house employment—in particular they spend a relatively larger share of their (in-house) labor cost on a small set of core occupations.

We then take the main predictions of the model to the data. We use individual- and firm-level administrative data on the universe of French workers from 1994 to 2010. This time interval covers the full roll-out of broadband connections across the French territory, which essentially took place between 2000 and 2007. We first confirm the skill-biased nature of BI. As cities become progressively connected, the average productivity of local firms increases and so does the share of high-skill workers. This increasing demand for high-skill workers is confirmed by a wage regression on individual panel data which shows a positive average effect on wage dynamics that is more pronounced for high-skill workers, even after controlling for a rich set of observed and unobserved heterogeneity. These results are in line with Akerman et al. (2015), who study the effect of broadband expansion in Norway. We go one step further and show evidence that this skill-biased technological change is accompanied by important changes in the distribution of workforce as BI led firms to outsource some non-core occupations to service contractors. We start by documenting how BI increases both the expenditure in outsourcing services and the occupational segregation of establishments, as captured by a Herfindahl-Hirschman Index (HHI) of concentration. This effect is driven by both changes within existing establishments and changes in establishment composition through entries and exits in the cities connected to the internet.[3]

Interestingly, outsourcing does not only affect low-skill workers such as cleaners or drivers, but also skilled professionals. We establish a list of sectors specializing in services to other firms, and we identify the occupations that are likely to be outsourced

[2]An additional impact of BI would include a shift in the output elasticity of high skill occupations relative to low skilled occupations resulting in increasing wage returns for high skill workers. In this paper, we show that BI is indeed a skill biased technology, but we do not explicitly model it through an impact in the output elasticities. We refer the reader to Akerman et al. (2015) for such a model.

[3]This evidence is in line with the trends described by Godechot et al. (2020), who, based on administrative data in several high-income countries, document a steep increase in both earnings and occupational segregation at work, which, especially with regard to occupation, is particularly pronounced in France.

to such sectors.[4] We show that following BI expansion, workers employed in outsourceable occupations in both the high- and low-skill segment become increasingly concentrated in establishments specializing in these particular services, and thus less likely to be employed in-house by establishments active in other sectors. Finally, we provide suggestive evidence that the impact of domestic outsourcing on wages is heterogeneous across skills: high-skill workers experience wage increases after outsourcing while low-skill workers see their wages decrease, implying that outsourcing tends to widen preexisting disparities in earnings.

The analysis in this paper contributes to several strands of literature. First, we relate to a recent empirical literature exploiting BI expansion as a plausibly exogenous technological and informational shock. In France, Malgouyres et al. (2019) show that the BI expansion was associated with an increase in imports among firms treated by the policy.[5] More closely related to our study, Akerman et al. (2015) evaluate the skill-bias of BI in Norway. They find that broadband availability increases both firm productivity and the skill wage premium. We contribute to this literature by documenting one of the mechanisms behind the BI skill-bias. In particular, we show that it operates through the reallocation of high and low-skill workers across firms –consistent with an increase of productivity and a reduction of outsourcing costs– and not only through a shift in the firm's production function. These reorganizations affect heterogeneous workers in a way that amplifies preexisting wage inequalities.

Second, our paper is linked to the literature on domestic outsourcing. Weil (2014) describes how the nature of work has changed in the 21st century as a result of large companies switching to a "fissured workplace" business model. Goldschmidt and Schmieder (2017) show how Germany has experienced an explosion of domestic outsourcing of low-skill non-core activities since the early 1990s.[6] In their paper, outsourcing results in wage reductions for outsourced jobs, mainly driven by the loss of firm-specific rents.[7] Song et al. (2019) underline how the increase in sorting and

[4] A detailed description of these definitions is reported in Appendix B. The definition of low-skill outsourcing follows closely Goldschmidt and Schmieder (2017) and includes cleaning, security, driving and logistics. For high-skill outsourcing we select the two largest industry categories that provide professional services to other firms: IT and consulting (which includes strategy consulting, HR and advertising).

[5] Consistently, Akerman et al. (2018) show that BI narrows the role of distance in explaining bilateral trade in Norway. In the United Kingdom, DeStefano et al. (2018) find that a similar shock did not affect firms' productivity but only their size, which suggests that local institutions matter in explaining the effect of broadband expansion.

[6] They focus on four emblematic activities: cleaning, logistics, security and food services and introduce an innovative measure to capture outsourcing of such tasks.

[7] Drenik et al. (2020) quantify this loss and show that the share of the firm-specific wage premium that

the segregation of low-skill workers contributed to the surge in income inequality. Technology has also been related to outsourcing: using data on US firms, Fort (2017) shows that technology lowers the cost of coordination within a firm and leads to an increase in the fragmentation of production. Cortes and Salvatori (2019) show that firms in the UK have become increasingly specialized over the past 20 years, and that this was partly driven by an increase in domestic outsourcing in high-skill non-routine cognitive tasks. Aghion et al. (2019a) characterize so-called "good jobs" that are protected from outsourcing as technology advances, even among low-educated workers. We contribute to this literature in two main aspects. First, we look explicitly at the effect of a specific technology and extend the spectrum of affected workers to high-skill occupations.[8] Our estimate represents intent-to-treat as we do not track a firm-level measure of actual ICT adoption. As such, our results complement the survey evidence (Abramovsky and Griffith, 2006) which points to a positive association between direct measures of firm-level ICT-intensity and the propensity to outsource and purchase services on the market. Second, we capture margins of adjustment largely ignored by the existing literature. Although most work focuses on either individual transitions from in-house to outsourced activities (Dube and Kaplan, 2010) or on plant-level outsourcing events (Goldschmidt and Schmieder, 2017), domestic outsourcing is likely to become more prevalent through the allocation of new labor-market entrants to firms operating in the business services sectors and the entry of new firms in such sectors. Our city-level approach to analyzing the local evolution of the allocation of workers and occupations across firms in reaction to a change in the set of available technologies allows us to capture such margins.

In a recent work, Bilal and Lhuillier (2021) consider the aggregate implications of domestic outsourcing in terms of wage and productivity in France. They develop a model where firms face labor market frictions when hiring workers and can instead buy the services from outsourcing firms. Their structurally estimations suggest that that the rise of outsourcing in France have positively contributed to global output and reduced the labor share. We depart from this work mostly by estimating in a reduced-form framework how the propensity to outsource respond to a change in available technology.[9]

outsourced workers earn is about half of the premium that an insider earns.

[8] Abramovsky et al. (2017) discuss the evolution of high-skill occupations' *offshoring*. The literature has also identified occupations that are more likely to be domestically outsourced to the growing business service sector, which can include some high-skill occupations such as advertisers, accountants, IT specialists and legal professionals (Ono, 2003; Berlingieri, 2014; Goldschmidt and Schmieder, 2017).

[9] Our theoretical framework builds partly on the same intuition—namely that rent-sharing when hiring

Our paper is also more generally connected to the literature on technology and firm organization of labor. In France, Caliendo et al. (2015) analyze how the hierarchical structure changes as a firm grows, underlining the importance of considering layers when studying the dynamics of employment and wages. Le Moigne (2020) describes how, over the period from 1995 to 2015, French plants have become increasingly prone to removing their "middle layer", and show that these reorganizations are associated with greater productivity for firms but lower internal promotion opportunities for workers at the bottom of the job ladder. While theses papers do not formally consider new technologies, their role in skill-biased organizational change has been documented in previous papers (see e.g. Autor et al., 1998, 2008, 2003; Michaels et al., 2014) and has been found to be associated with a reduction in layers by accentuating the "flatness" of a firm (Caroli and Van Reenen, 2001; Aghion et al., 2019a). The literature on the effects of ICT usually considers separately the effect of a reduction in the cost of communication and in the cost of information on the organizational structure of a firm (see e.g. Bloom et al., 2014). For instance, Garicano and Rossi-Hansberg (2006) makes a clear distinction between the cost of communication and the cost of acquiring knowledge. In their model, a reduction in communication cost fosters a firm's reorganization, where more problems are solved at the top of the hierarchy (typically by managers) and the knowledge content of production work shrinks. Garicano (2000) and Borghans and Ter Weel (2006) describe how improved communication increases the specialization of workers within a firm. In our paper we consider the effect of the reduction in both communication and information costs brought about by BI on reorganization beyond the boundaries of the firm, through the subcontracting of non-core activities to high and low-skill service providers.

The rest of the paper is organized as follows. Section 2 presents a simple model that will guide the empirical analysis. Section 3 present the data and our empirical methodology. Section 4 presents results regarding the effect of BI on outsourcing. Finally, Section 5 looks more closely at the consequence of outsourcing on workers' wage trajectories. Section 6 concludes.

in-house workers provides an incentive to outsource—, but departs from this work by introducing an explicit distinction between core and non-core occupations depending on how substitutable outsourcing and in-house labor services are within an occupation. We further focus on the qualitative, partial equilibrium predictions of our model in order to guide the interpretation of our event-study results rather than carry out the structural estimation of a general equilibrium model.

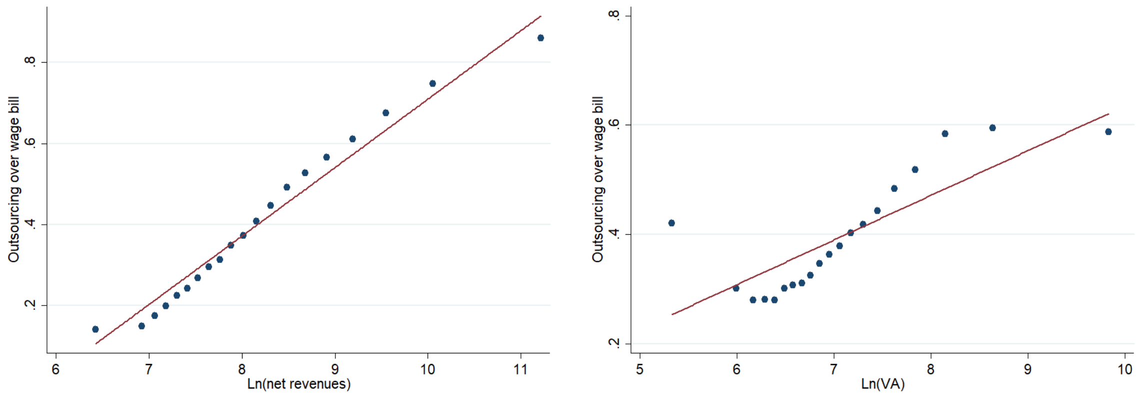
2 A Theory of technology-led domestic outsourcing

In this section, we propose a simple partial equilibrium model to highlight through what mechanisms could access to broadband internet lead to higher levels of domestic outsourcing. The main goal of this exercise is to show that our empirical results are rationalizable in a model with a fairly standard set of assumptions.

A key empirical regularity which will guide the the theory is the positive relationship between a firm size (as measured by revenue) and its propensity to outsource as shown in Figure I. Figure I and the corresponding Table A1 in the Appendix show the cross-sectional correlation between the outsourcing intensity of the firm and its size, measured as net revenue and value added. The intensity of outsourcing is measured as the ratio of subcontracting expenditure to in-house worker wage bill.[10]

FIGURE I. Correlation between outsourcing intensity and firm size

(a) Outsourcing intensity and net revenues (b) Outsourcing intensity and value added



Notes: The figure presents the binned scatter plot for the correlation between firm outsourcing intensity and the log of firm revenues (panel a), a,d between firm outsourcing intensity and the log of value added (panel b).

In our model, firms choose employment across several occupations to maximize profit. Within each occupation, workers can be hired either in-house or through a subcontractor.

2.1 Model setup

Production technology: aggregation across occupations. Firms combine several occupations or tasks to produce output using a Cobb-Douglas function with constant

[10] Such correlation is also a feature of Bilal and Lhuillier (2021)'s model.

returns to scale.[11] We denote output of j as a function of each occupation output $H_{i,j}$ as:

$$Y_j = \theta_j \prod_{i \in \mathbf{N}} H_{i,j}^{\alpha_i} \tag{1}$$

where θ denotes a Hicks neutral productivity shifter, $\mathbf{N}$ is the set of occupations and we have $\sum_i \alpha_i = 1$, and $\alpha_i \in [0, 1]$.

Production technology: in-house and outsourced workers within occupation. Each occupation i can be carried out by a mix of in-house workers which are directly employed and of outsourced workers whose labor services are hired through a third party (subcontractor). Each occupation is characterized by a specific elasticity of substitution between in-house and outsourced workers. Output by occupation i depends on the number of in-house and outsourced workers denoted n_i and s_i respectively, and is expressed as:[12]

$$H_i = \left(\mu_i^{\frac{1}{\sigma_i}} n_i^{\frac{\sigma_i - 1}{\sigma_i}} + (1 - \mu_i)^{\frac{1}{\sigma_i}} s_i^{\frac{\sigma_i - 1}{\sigma_i}} \right)^{\frac{\sigma_i}{\sigma_i - 1}} \tag{2}$$

where σ_i is the elasticity of substitution between the services provided by in-house and outsourced workers and μ_i affects the relative productivity of the groups. A *core occupation* is defined as an occupation with a low elasticity of value for σ_i. It could be either a high or a low productivity occupation as measured by α_i, i.e. the elasticity of overall output Y to the occupational output H_i.

How does this modeling approach relates to influential theories of outsourcing? The knowledge-based view of the firm (Kogut and Zander, 1992) suggests that firms should keep in-house tasks (here occupations can be seen as bundle of tasks) that (1) are hard to codify and (2) have higher economic value to the firm. Similarly, a strand of management literature dealing with "core-competencies" recommend to fo-

[11] Throughout the paper we make the assumption that one occupation is a fixed bundle of tasks. We acknowledge that this is a simplification since the literature has shown that the task content of jobs evolves over time also due to technological change (e.g. Spitz-Oener, 2006, Deming and Noray, 2020). We are obliged to make this assumption due to the limits of administrative data, which does not allow to observe tasks beyond their occupational dimension.

[12] To be more consistent with the empirical analysis, we adopt an "occupation" approach. We can see each occupation as a continuum of tasks, some of which will be performed by outsourced workers and other by in-house workers. At the equilibrium, an occupation is therefore characterized by its relative level of outsourced workers, which in turns is determined by the elasticity of substitution σ_i.

cus on tasks with strategic value to the firm and outsource non-strategic routine tasks (Lepak and Snell, 1999).

The model encapsulates the key ideas from this literature through the parameter σ_i. The elasticity of substitution σ_i will determine the extent to which firms wish to increase outsourcing as the relative cost of doing so goes down. A core occupation in that set-up is a bundle of tasks that is hard to codify and therefore has a low σ_i, which in turn implies that a decline in the cost of outsourcing will not provide a strong incentive to outsource this task. The economic value produced by the task is captured by its weight in the Cobb-Douglas aggregation (α_i). A core task with high economic potential is therefore an occupation with a high α_i and a low σ_i. On the contrary, a non-core task is characterized by a low α_i and a high σ_i. As we will see below, in the model, profit maximization implies that a decline in the relative cost of outsourcing, or an increase in the optimal scale of the firm, will lead to a refocusing of the firm on core tasks. As a consequence, the share of core occupations in the overall wage bill increases.

(Labor) market structure for in-house and outsourced workers. A key difference between the hiring of in-house and outsourced workers is that each firm disposes of some wage setting power when hiring in-house workers but are price-takers with respect to the firms from which they outsource (which we call the agencies). We micro-found (occupation-specific) firm-level labor supply curves as resulting from a discrete choice modeling and in keeping with the recent literature on monopsony (Card et al., 2018; Lamadon et al., 2019). In this set-up, the labor supply curve that individual firms face (within a given occupation) is not perfectly elastic because of idiosyncratic tastes among workers for the amenities offered by the firms (for instance working conditions, commute, corporate culture). Due to asymmetric information regarding the valuation by individual workers of such amenities, firms are not able to perfectly discriminate and fully price these amenities into individual-specific wages. As we will see this assumption naturally generates the positive correlation between outsourcing intensity and size which showed in Figure I that is key in this model.[13]

This idea is summarized by equation (3) which gives the labor supply curve of a firm j hiring a (in-house) worker in occupation i. Namely:

[13] We assume the market for outsourcing services is competitive and that consequently there is not rent-sharing between employers and their outsourced workers. This does not preclude the possibility that outsourced workers benefits from rent-sharing with respect to their direct employer (the agencies) due to frictions on the labor market.

$$n_{i,j} = N_i \frac{w_{i,j}^{1/\rho_i}}{\sum_j w_{i,j}{}^{1/\rho_i}} = a_i w_{i,j}^{\frac{1}{\rho_i}}, \tag{3}$$

where N_i is the measure of the population of workers in occupation i with iid extreme value type-1 preferences across firms with shape parameters ρ_i.[14] We consider a standard atomistic monopsonistic competition setting as firms ignore their own impact on the competition index which is captured by a_i. For simplicity, we consider in our model that all occupations i have the same value of $\rho_i > 0$ and a_i which we denote as ρ and a.

Profit maximization. We consider a set-up with monopolistic competition and CES demand. Each firm j faces the demand function $Y^D = p^{-\varepsilon} I$ which yields the following revenue function:

$$R(Y) = Yp = Y^{\frac{\varepsilon-1}{\varepsilon}} I^{\frac{1}{\varepsilon}} \tag{4}$$

The cost associated with hiring a vector of workers $\{n_{i,j}, s_{i,j}\}_{i\in\mathbf{N}}$ writes as :

$$C(\{n_i, s_i\}_{i\in\mathbf{N}}) = \sum_{i\in\mathbf{N}} n_{i,j} w_i(n_{i,j}) + \sum_{i\in\mathbf{N}} s_{i,j} \cdot \gamma_{i,j} r_i \tag{5}$$

where $w_i(n_{i,j})$ is the inverse labor supply function faced by firm j when hiring in occupation i. The variable r_i is the market price for outsourcing services in occupation i and $\gamma_{i,j}$ is the firm-specific cost shifter of outsourcing.

$$\max_{\{n_{i,j}, s_{i,j}\}_{i\in\mathbf{N}}} \pi_{i,j} = Y(\{n_{i,j}, s_{i,j}\}_{i\in\mathbf{N}})^{\frac{\varepsilon-1}{\varepsilon}} I^{\frac{1}{\varepsilon}} - \left(\sum_{i\in\mathbf{N}} n_{i,j} w_i(n_{i,j}) + \sum_{i\in\mathbf{N}} s_{i,j} \cdot \gamma_{i,j} r_i \right) \tag{6}$$

It is fairly straightforward to show that the problem defined in Equation (6) admits a unique positive solution $\{n^*_{i,j}, s^*_{i,j}\}_{i\in\mathbf{N}}$.[15]

Unlike what would occur under a competitive labor market, occupation-firm specific optimal wage $w^*_{i,j}$ depend on the level of labor demand $n^*_{i,j}$. This dependence precludes any closed form solution for $n^*_{i,j}$ but under some conditions on ρ allows us to

[14] This labor supply function arises from worker k in occupation i having utility: $u_{i,j,k} = w_{i,j} + e_{i,j,k}$, where $e_{i,j,k}$ follows an extreme-value type I distribution with scale parameters ρ_i.

[15] The problem with fixed wages $w_{i,j}$ is entirely standard and $\pi_{i,j}$ is strictly concave in $\{n_{i,j}, s_{i,j}\}$ so that any first order condition correspond to a global maximum. Allowing $w_{i,j}$ to increase with respect to $n_{i,j}$ make the profit function more concave and does not alter the uniqueness and existence of the solution.

derive our main predictions.

2.2 Outsourcing motives in our model

At this point it is worth recalling what the literature has presented as the most important rationales for outsourcing and assess the extent to which they are captured by our model.[16]

Abraham and Taylor (1996) consider three main sets of considerations underpinning a firm's decision to outsource some tasks to subcontractors : (i) the wage and benefit savings it could realize, (ii) the availability of specialized skills possessed by the outside contractor (iii) the volatility of its output demand. Our model emphasizes the first motive. As in Card et al. (2018); Lamadon et al. (2019); Kline et al. (2019); Aghion et al. (2018), the labor market power enjoyed by employers lead to some rent-sharing with in-house workers despite the absence of explicit bargaining. This rent-sharing creates an incentive to outsource some occupations to outside contractors as the market price for these services is independent of the firm's rents. This incentive becomes naturally more intense as firms become more productive or enjoy larger rents. The second motive is also incorporated in our setting, although less directly. The ability of outside contractor to provide cheaper services on a quality adjusted basis (i.e. low $\gamma_i r_i$) might derive from their access to more specialized skills.[17] Finally, the third rationale is intrinsically dynamic and is not explicitly modeled.

2.3 Broadband internet expansion in our model

By its nature, BI can impact the organization of a firm through different channels:

i In line with the literature (Akerman et al., 2015) and our own findings, a first effect of BI is to increase firm-level productivity, i.e. θ in our model.

ii To the extent that broadband lowers inter-organization frictions, another potential effect would be to lower the relative cost of outsourced versus in-house labor,

[16]Our review is primarily based on Abraham and Taylor (1996) and Appelbaum et al. (2015).

[17]As subcontractors become more competitive, firms will primarily choose to outsource non-core activities (high σ_i) while in core activities (low σ_i) employment in-house and outsourced will tend to grow hand-in-hand. In that sense, our model is inline with management theories recommending firms focus resources and talent on their core competencies and outsource whatever outside contractors are able to do better.

i.e. a decline in the term $\gamma_{i,j}$'s (see Abramovsky and Griffith, 2006; Abramovsky et al., 2017).

iii A third effect, again in line with most of the literature, would be to potentially generate some degree of skill-biased technical change (Akerman et al., 2015), which we model as a shift to the output elasticity α_i of high skill versus low-skill occupations.

iv A fourth potential channel through which broadband would affect firm organization would be a particularly strong productivity-impact of the IT outsourcing sector. Given the increase in the demand for IT services among producers and enhanced efficiency of IT subcontractors, the revenue marginal productivity of the IT subcontractors has likely gone up in the aftermath of broadband expansion.[18]

2.4 Predictions from the model

Without lost of generality, we consider that:

$$1 \leq \sigma_1 < \sigma_2 < \ldots < \sigma_N.$$

The first-order conditions with respect to s_i and n_i for all occupations $i \in \mathbf{N}$ leads to the following relationship:

$$s_{i,j} = \lambda_{i,j} n_{i,j}^{\rho\sigma_i+1}, \text{ where } \lambda_{i,j} = \frac{1-\mu_i}{\mu_i}\left[\frac{\rho+1}{a_i^{\rho} r_i \gamma_{i,j}}\right]^{\sigma_i}. \tag{7}$$

Because of this relationship, a firm can only increase its size by increasing its number of in-house workers as well as its outsourcing expenditures. $\lambda_{i,j}$ is a coefficient that measure the relative cost of these two types of labor and the level of complementarity. We assume that the parameters are distributed such that:

$$\lambda_{1,j} < \lambda_{2,j} < \ldots < \lambda_{N,j}.$$

[18] A last potential channel to mention is that broadband internet might increases workers' search efficiency thanks to online job adds, which will in turn decrease firms' monopsony power (Bhuller et al., 2019). This channel pushes in the opposite direction of the others, given that lower monopsony power would decrease the link between growth in productivity and internal labor cost, thus would work against our main findings.While acknowledging this potential effect, for the rest of the model we assume monopsony power to be fixed over time for simplicity.

In the baseline version of the model, we keep the setup as simple as possible and assume that there are only two types of occupations 1 and 2. Occupation 1 is the "core" occupation which is associated with a value of $\sigma_1 = 1$ and a value $\alpha_1 > 1/2$. By contrast, occupation 2 is the "non-core" occupation where in-house workers are more easily substituable by outsourcing workers ($\sigma_2 > 1$). We primarily present results pertaining to an increase in productivity. In the Appendix E, we show how the results are impacted if we allow for more than two types of occupations and we discuss the case of a decline in the cost of outsourcing. We also present a numerical resolution of the model.

Proposition 1. *A positive increase in θ raises the cost share of outsourcing for each occupation i for which $\sigma_i > 1$*

Proof. See Appendix E.1 □

The intuition for this result comes from the fact that firms respond to a positive productivity shock by increasing their workforce. As long as the elasticity of substitution between the two types of workers is larger than 1, then the firm will adjust both its number of in-house and outsourced workers. Yet, because $\rho > 0$, as the firm grows, it is more and more costly to hire in-house workers and the ratio η_c of the labor cost coming from outsourcing over the total labor force increases.

Proposition 2. *Following an increase in θ, the increase in the cost share of outsourcing is larger for the non-core occupations*

Proof. See Appendix E.2 □

This Proposition shows that all occupations will not be affected equally by the BI shock. The high σ_i (non-core) occupations will become increasingly composed of outsourced workers. While we cannot directly identify these occupations, in the empirical part of the paper we show that workers are more likely to move to a service firm specialized in tasks that are typically considered as non-core (cleaning services, driving, security...) following the BI shock.

Proposition 3. *Following an increase in θ, the concentration of in-house workers increases within firms.*

Proof. See Appendix E.3 □

Proposition 3 is easy to look at in the data as we directly observe in-house occupation composition (while we do not have direct measure of outsourcing expenditures by occupation). This result predicts that when a firm is connected to BI, its HHI of concentration should increase.

In this baseline version of the model, there are only two types of occupations. We did this to keep the model as simple as possible while keeping the core economic intuition. In Appendix E.7, we provide a numerical illustration of the comparative statics of the model. We solve the profit maximization problem of the firm for a specific case (with 4 occupations) and show how the optimal choices vary as productivity increases and the cost of outsourcing decreases. Results from this simple exercise show that outsourcing intensity increases with size (as measured by sales) and that a positive productivity shock or a decline in outsourcing cost is associated with rising HHI. In Appendix E.5, we show how to extend the model to include more than two type of occupations.

2.5 Link to empirical analysis

Our theoretical framework implicitly adopts a model of the core-competency of the firm, where cost-minimizing firms will concentrate on in-house occupations which are the hardest to outsource. Monopsony power in the in-house labor market implies some degree of rent-sharing with in-house workers, which provides an incentive for firms to outsource that increase with the level of productivity. A motivating consideration for this set of assumptions is that it predicts a positive correlation between scale and outsourcing intensity[19] which turns out to be a salient feature of the data (Figure I).

In that setting, a technological shock such as broadband expansion, which decreases the cost of outsourcing directly and/or boosts firms' productivity, will lead to more outsourcing. Under a broad set of conditions, it will also push firms to become more occupationally segregated meaning that their total labor cost will be concentrated on a smaller set of occupations (which we measure with the HHI). We will take most of these predictions to the data.

Before we move on to the empirical sections of the paper, let us briefly discuss the case of high-skill outsourcing. In this version of the model, we have assumed that there is a negative correlation between the σ_i and α_i, i.e. that non-core occupations

[19] See Appendix E.6 for a formal proof.

are low skilled and conversely. However, the proofs presented in the Appendix show that most of the results hold if some non-core occupations have high α_i (we typically think of IT workers or accountants). We further present numerical results assuming a null correlation between α and σ and obtain (for a particular value of the parameters) predictions similar to those presented here.[20]

3 Data and Empirical Strategy

3.1 Data

Administrative data on labor market outcomes. Our data comes from two main administrative sources. The first is the matched employer-employee dataset covering all workers based in France since 1994 (*Déclaration annuelle des données sociales* or DADS). This dataset constitutes the primary source used to compute payroll tax data and gives detailed individual information including salary, hours worked, occupation, age, gender and the identifier of the employing firm and establishment. It is not possible to follow workers over time, except for a random sample of 1 out of every 24 workers, which we refer to as the DADS panel. We clean these data and restrict our sample to the French private sector. Appendix B provides details on how we depart from the raw dataset. One information that is missing from such administrative data is the individual level of education. We therefore rely on occupation codes to classify workers into skill groups.

The second main source is the administrative financial records made available by the French Ministry of Finance (*FIchier Complet Unifié de Suse* or FICUS), which report performance indicators for each private sector firm for the same period of time. We need this data to extract information relative to each business' value added, which is used to compute our measure of labor productivity. In the FICUS records the value of outsourcing expenditure is included within a broader category of spending, making it impossible to isolate the impact of broadband internet on this particular sub-category. In order to recover a more precise measure of outsourcing expenditure at the firm level, we thus take advantage of an alternative administrative dataset which is curated by the Bank of France and is called *Fichier inter-Bancaire des Entreprises* (or FiBEn). The

[20]However, if most high-skill occupations are considered non-core, the revenue share would increase with σ_i which results in an opposite effect for the concentration index. We do not view this case as relevant however.

raw files used to build FiBEn and FICUS are the same, but the coverage is slightly different as FiBEn excludes firms that did not reach 700,000 euros in yearly turnover. Results on outsourcing expenditure are thus representative of larger firms. Contrary to the DADS, neither FICUS nor FiBEn provide information at the establishment level but only at the more aggregate firm level. In the analysis we thus present evidence that our results are robust to including or excluding multi-establishment firms.[21]

We put forward several measures that serve as a proxy for outsourcing. The most direct one is firm-level outsourcing expenditure scaled by the wage bill, which closely matches the outcome of the theoretical model. Secondly, to have a broader sense of how BI affects the segregation of occupations across establishments, we construct a measure of establishment occupational concentration based on an HHI computed using shares in the total wage bill.[22] The theory predicts that the HHI should increase as a result of the BI shock.[23] Finally, we define typical outsourcing sectors and outsourceable occupations as follows, consistent with Goldschmidt and Schmieder (2017):

- **Outsourceable occupations.** Non-core occupations that are likely to be outsourced, offshored or subcontracted. We divide them into high-skill outsourceable occupations (IT specialists and consultants), and low-skill outsourceable occupations (cleaners, security workers, drivers, and logistic workers). The full list of occupations and their definition is detailed in Appendix B.

- **Outsourcing sectors.** Administrative and business support service sectors. High-skill outsourcing sectors: IT and consulting sectors. Low-skill outsourcing sectors: cleaning, security, driving and logistics services. The full list of sectors and their definition is detailed in Appendix B.

[21] Our preferred model for the productivity results assigns to each plant within multi-establishment firms the same value added over wage bill, as measured at the aggregate firm level. Our preferred model for the outsourcing expenditure results keeps multi-establishment firms and assigns outsourcing intensity to the location of the headquarter, since information on other plant locations is not available in FiBEn. In both cases, we provide robustness tests excluding all multi-establishment firms from the sample.

[22] The concentration of occupations within each establishment i is computed with the following formula: $HHI_i = \sum_{o=1}^{18} \psi_{oi}^2$, where o indexes each of the 18 occupational categories that compose the French private sector at the 2 digit (also known as the PCS classification), and ψ_{oi} is the share of the wage bill of the establishment attributed to workers employed in a given occupation. At the city level c, we compute the average establishment concentration as following: $HHI_c = \sum_{i=1}^{N} \omega_{ic} HHI_{ic}$, where ω_{ic} is the share of employment in the city accounted for by each establishment i.

[23] See Proposition 3. In addition, we show numerical results based on a special case of the model in Online Appendix Section E.7, see in particular Figure E3 and associated comments.

At the city level, we construct a measure of the expansion of outsourcing sectors by looking at their share of overall employment in the city. We also check whether their corresponding outsourceable occupations are increasingly concentrated within establishments specializing in these particular services, and thus less likely to be employed in-house by establishments active in other sectors. To do that we compute the share of total employment in high- and low-skill outsourceable occupations that is employed by the high- and low-skill outsourcing sectors within the city. At the individual level, we test whether workers are increasingly likely to change establishment once the city is connected to BI, and we measure whether this effect is stronger for mobility towards outsourcing sectors, both in low-skill services and in high-skill services.

Data on broadband expansion. ADSL (Asymmetric Digital Subscriber Line) is a data communication technology that enables fast data transmission over copper telephone lines. Eligibility for ADSL depends on the distance between the final customer and a Local Exchange (LE), since the intensity and the quality of the analogue signal decreases as it is routed over the copper lines. LEs are telephone exchanges owned by the incumbent operator France Télécom into which subscribers' telephone lines connect. A key feature of ADSL technology is that one can supply high-speed internet by upgrading the LE while relying on the existing (copper) local loop to connect the premises of the final customers. The upgrading of local LEs is the key source of variation that we use in our empirical analysis.[24] Our main information of interest documents the date when each local exchange unit was upgraded in mainland France.[25] While we do not have data on firm level use of BI, suggestive evidence points towards the fact that the ADSL was widely adopted by local companies once it became available to them.[26] We additionally obtained data from the regulatory agency (ARCEP) on the geographical coverage of each local exchange unit. Each city in France is partitioned into census blocks, and the data document the area of each

[24] See Appendix C for a more detailed account of the ADSL expansion.

[25] Throughout the paper, broadband or ADSL refers to first generation ADSL that is associated with speed of 512 kbit/s. The historical operator was compelled by law to make this data available to other operators as well as websites allowing consumers to gauge the quality of their line. The data was collected through one such website by Malgouyres et al. (2019) and validated manually.

[26] A 2016 survey showed that in that year 73% of small and medium enterprises used ADSL technology (Arcep, 2016). The large take-up reflects the fact that ADSL was a massive improvement in terms of speed as well as in terms of connection cost and time. While there is no administrative data on firm-level use of broadband, based on repeated survey data, firms located in cities that received broadband internet earlier experienced higher growth in the proportion of employees that used internet on a regular basis between 1999 and 2004. This statistical association cannot be interpreted causally, but it is however strongly suggestive of an impact from broadband availability on broadband adoption.

census block (IRIS) covered by a given local exchange unit. Combining both datasets, we construct a continuous measure of broadband access for city i at year t. This measure, which we denote $\widetilde{Z}_{it}$, is a time-weighted percentage of area covered in city i. It is formally defined as:

$$\widetilde{Z}_{it} = \sum_{b \in i} D_{b,t} \frac{\mathcal{A}_{b,t}}{\sum\limits_{b' \in i} \mathcal{A}_{b',t}}, \tag{8}$$

where $b \in i$ denotes the census tracks included in city i, $D_{b,t}$ share of the days of year t with access in b since January 1st of year t normalized by the number of days in year t. Finally, $\mathcal{A}_{b,t}$ denotes the area covered by census track b. The variable $\widetilde{Z}_{it}$ is continuous with support between 0 and 1, reflecting both the area and time dimensions of local broadband availability.[27] We refer to $\widetilde{Z}_{it}$ as the *degree of connection* of a city.

In our study of the dynamics of the labor market following BI expansion, it will sometimes be useful to consider a discrete analog of $\tilde{Z}$. We thus define the year of treatment as the first year where $\tilde{Z}$ becomes positive, thus signaling that the first portion of the city was connected to BI. We denote $C_{i,t}$ the corresponding discrete (binary) variable, and we refer to it as the "year of connection". The discretization results in little loss of information, given the underlying distribution of the continuous variable, and allows us to implement a transparent before/after comparison through the estimation of event studies. The latter ranges from 1999 for a handful of experimental cities to 2007 for the most remote areas. Figure A2 in Appendix A shows the evolution of $\tilde{Z}$ after the year of first connection. On average the first year is characterized by a degree of connection of 0.3, which jumps to 0.6 the following year and quickly reaches 0.8.

Table I reports the summary statistics for the main outcomes of interest depending of the year of connection. Panel A describes the city-level data while panel B describes the establishment level data. Firms and cities connected to the internet during the first three years of ADSL expansion are larger, have a slightly higher share of high-skill workers and spend more in outsourcing relative to their wage bill. All other measures of outsourcing are also larger in earlier cohorts, except for the occupational concentration, which is slightly lower in large cities.

[27] $\widetilde{Z}_{it}$ will be equal to one if all of its areas have had access for the entire year. It will be equal to 1/2 if the entire city has had access to broadband over half the year t. In practice however, it is very strongly concentrated on 0 and 1, with very few intermediate observations, see Figure A1 in Appendix A.

TABLE I. Summary Statistics

	All sample	By cohort of ADSL arrival		
		99-01	02-04	05-07
VARIABLES	mn/(sd)	mn/(sd)	mn/(sd)	mn/(sd)
Panel A: City level				
Share high skill workers	0,07	0,08	0,06	0,05
	(0,07)	(0,08)	(0,06)	(0,06)
Occupational concentration (HHI)	0,38	0,37	0,39	0,40
	(0,16)	(0,14)	(0,17)	(0,19)
Value added over wage bill (2010 euros)	2,08	2,12	2,07	2,00
	(1,84)	(1,59)	(2,13)	(1,08)
Outsourcing exp. over wage bill (2010 euros)	0,30	0,34	0,27	0,24
	(0,63)	(0,62)	(0,65)	(0,48)
Share of empl. in high-skill outs. services	0,02	0,03	0,01	0,01
	(0,09)	(0,10)	(0,09)	(0,08)
Sh. outsourceable workers in HS outs. Services	0,08	0,11	0,05	0,04
	(0,22)	(0,23)	(0,19)	(0,16)
Share of empl. in low-skill outs. services	0,11	0,11	0,11	0,11
	(0,24)	(0,22)	(0,26)	(0,28)
Sh. outsourceable workers in LS outs. Services	0,25	0,29	0,20	0,17
	(0,38)	(0,38)	(0,37)	(0,36)
Average establishment size	77,05	109,12	52,09	42,66
	(262,76)	(372,31)	(103,60)	(102,53)
N. of establishments	13,17	24,65	3,96	2,63
	(59,68)	(87,33)	(6,90)	(6,77)
Observations	172'132	77'385	82'593	12'154
Panel B: Establishment level				
Share high skill workers	0,11	0,13	0,06	0,06
	(0,17)	(0,18)	(0,09)	(0,09)
Occupational concentration (HHI)	0,38	0,37	0,39	0,39
	(0,19)	(0,18)	(0,20)	(0,20)
Value added over wage bill (2010 euros)	2,19	2,21	2,09	2,05
	(5,78)	(6,26)	(2,92)	(2,63)
Outsourcing exp. over wage bill (2010 euros)	0,39	0,42	0,27	0,28
	(1,45)	(1,55)	(0,86)	(0,82)
Observations	2'439'340	2'052'701	352'038	34'601

Notes : This table presents the summary statistics of the main variables used in the regression. The first column provides the means (and standard deviation) over the whole sample, while the subsequent three columns split the sample according to the cohorts of BI arrival. Averages are computed over the full period from 1997 to 2007.

3.2 The diffusion of Broadband Internet in France

As evidenced by Malgouyres et al. (2019), the deployment of broadband internet technology beyond France's largest cities began very slowly at the dawn of the 2000's and continued up to 2007, due to multiple reasons. First, France Télécom, the monopolistic telecom supplier, was uncertain regarding the future regulations on the wholesale price that it was going to be able to charge. Second, at the same time that France Télécom had to invest massively in upgrading its LEs to ADSL, it went through a debt crisis that ended with what was essentially a government bailout in 2002. Urged on by the government – which increased its stake in the firm during the 2002 bailout of the firm – in 2003 France Télécom pledged to cover 90% of the French (mainland) population by the end of 2005. Between 2004 and 2007, local governments got involved in broadband internet deployment by subsidizing the expansion and favoring competition among providers. We account for the role of local government in our empirical analysis by including département-year fixed effects. Overall, the account of the broadband expansion in France over the period suggests that it was slow until 2002. After 2002, with strong encouragement from the government, France Télécom gradually started covering more secondary areas with a focus on the overall number of lines per LE, with only limited attention paid to local economic potential (see Appendix C for more details). Table C1 in Appendix shows the percentage of cities, establishments and workers that were connected each year. In 2000, only 2% of cities were connected, although this corresponds to a much larger share of workers and establishments (respectively 25% and 22%). By 2003, 80% of workers and 76% of establishments were connected. At the end of 2005, 96% of French workers and 80% of cities were covered.

Explaining broadband expansion Because our main effects of interest are identified out of the gradual diffusion of the new technology in different LEs over space and time, addressing the endogeneity of the decision to "treat" one LE before another deserves special consideration. Malgouyres et al. (2019) describe how that broadband expansion occurred to maximize population coverage with no special consideration for economic potential. Our identification strategy relies on the assumption that the coverage of cities was mostly determined by city population density – which is almost fixed over time and can be controlled for – and did not take into account underlying local trends in productivity or propensity to outsource activities. As a result, conditional on time and city fixed effects, we consider the variation in broadband access to be as good as random. In order to assess the plausibility of this assumption, we explore the extent to which broadband coverage over time can be explained by different types of lagged city-level covariates. We group those covariates into three main groups:

1. **Density**: population in 1999 per square km (log), interacted with a full set of year dummy variables.

2. **Industry dynamics**: shares of employment in 42 economic sectors at $t-1$ as well as changes in shares between $t-1$ and $t-2$.

3. **Changes in politics**: share of votes for left-wing candidates at the presidential

election of 1995 and a dummy for change in political majority between 1995 and 2002, both interacted with a full set of year dummy variables.[28]

We then estimate the following specification:

$$\tilde{Z}_{it} = \mathbf{dens}'_{it}\rho_1 + \mathbf{indyn}'_{it}\rho_2 + \mathbf{pol}'_{it}\rho_3 + \text{FE}_i + \text{FE}_{r(i),t} + \varepsilon_{it}, \tag{9}$$

where $\tilde{Z}_{it}$ is the time-weighed share of city i that is covered by broadband internet as described in Equation (8). As we are mostly interested in the explanatory power of these different groups of observable variables, we only report the R-squared of each regression in Table II.

TABLE II. EXPLAINING CITY BROADBAND COVERAGE: PANEL ANALYSIS

VARIABLES	(1) Baseline	(2) Baseline + covariates	(3) Baseline + dep x yr FE	(4) (3) + density	(5) (3) + industry	(6) (3) + politics	(7) (3) + all cova.
R^2	0.759	0.765	0.782	0.784	0.783	0.785	0.787
Industry F-stat		111.5			20155		1742
Density F-stat		84.49		8.490			10.96
Politics F-stat		2.96				9.9	9

Notes : This table presents the R-square of panel regressions following equation (9). Baseline (Column 1) refers to the baseline model with only city fixed effects and year fixed effects. Column (2) adds all the covariates to column (1): Density (including 1999 population density at the city level interacted with year dummies), industrial structure (including the lagged share and their changes of industry shares using 40 sectors), and politics (including the share of left votes in the city in the 1995 presidential election interacted by year dummies and an indicator for a change in majority from left to write or vice-versa between '95 and '02, again interacted with year dummies. Column (3) adds department × year fixed effects to column (1). The remaining columns report different combinations of these controls.

Our results are summarized in Table II. We start by regressing broadband internet coverage on time and city fixed effects, which we call baseline. As reported in column (1) of Table II, we obtain an R-square of 75.9% indicating that these variables capture a substantial share of the variation in treatment status. Column (2) presents the R-square of the baseline model with the addition of all covariates described in equation 9, which only increases to 76.5%. Column (3) adds department × year fixed effects to the baseline model, to account for the fact that local governments got involved in the BI diffusion decision after 2002. This model absorbs 78.2% of the variance in treatment intensity. Columns (4) to (7) add different combinations of the covariates to the

[28] These variables are proxy for the political orientation of local politicians, since candidates in local elections often do not have an explicit political affiliation. We categorize all the national candidates that run in the 1995 and 2002 presidential election into either left-wing or right-wing. Left-wing candidates include Lionel Jospin, Robert Hue, Arlette Laguiller, Dominique Voynet, Jeacques Cheminade, Jean-Pierre Chevènement, Noël Mamère, Olivier Besancenot, Christiane Taubira, Daniel Gluckstein. Right-wing candidates include Jacques Chirac, Edouard Balladur, Jean-Marie Le Pen, Philippe De Villiers, François Bayrou, Jean Saint-Josse, Alain Madelin, Bruno Mégret, Christine Boutin, Corinne Lepage.

specification with department × year fixed effects. Interestingly, the set of industry dynamic variables barely increases the fit of the model, indicating that, conditional on city and department-year fixed effects, they are roughly unrelated to the timing of internet coverage. Population density and city level political orientation have a slightly larger explanatory power, also considering the fact that they reduce much less the degrees of freedom relative to the industry controls, but adding all the covariates leads to little improvement relative to column (3). We decide to retain Column (5) as our preferred set of controls since density has a slightly larger F-stat than local politics, and we know from official documents that reaching a wide portion of the population was among the main criteria used to decide which cities to connect first.

3.3 Empirical Strategy

To identify the causal effect of broadband internet, we exploit the fact that the dissemination of ADSL across municipalities was staggered over a period of nearly 10 years, from 1999 to 2007. While the timing of diffusion is clearly dependent on the size and density of the city, as described in the previous subsection, after controlling for these elements we can exploit quasi-random variation in access across cities of similar densities that are located within the same department. Formally, we use the panel of cities c observed for each year t from 1997 to 2007 to run the following econometric model:

$$Y_{c,t} = \sum_{\tau=-k}^{k'} \alpha_\tau \mathbb{1}\left\{t = t_c + \tau\right\} + \gamma_t \boldsymbol{X_{c,t}} + \nu_c + \varepsilon_{c,t} \tag{10}$$

where Y is the variable of interest, t_c is the year of ADSL arrival of a city c and X is a vector of control variables. X contains a time unvarying measure of the density of the city in 1999 interacted with year dummies and a set of department-year intercepts. Finally, we add city-level fixed effects such that the coefficients α_τ can be interpreted as changes within a given city resulting from the arrival of the ADSL, compared to cities within the same department and with similar densities that have yet to be connected to BI. We exclude two dummies from the regression, respectively for $\tau = -1$ and $\tau = -6$ as suggested in Borusyak et al. (2021) in this type of event study setting with year and individual fixed effects. The regression is run over the sample of cities that have more than 100 inhabitants at the beginning of the period, to avoid capturing an effect driven by small villages, and over the years 1997-2007, which include the full period of ADSL expansion.

A recent literature underlines the caveats of using two-way fixed effects models in staggered adoption contexts (Callaway and Sant'Anna, 2020; De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2018; Borusyak et al., 2021). The main issue resides in the fact that α_τ for the post-treatment period is partly estimated using previously treated observations as controls. This might introduce biases in the presence of dynamic or heterogeneous treatment effects across cohorts. Given that in our context there is no pure control, since all cities are treated by the end of the period, we adopt the stacked difference in difference approach employed by Vannutelli (2020), which is also similar to Cengiz et al. (2019) and Deshpande and Li (2019).[29] The latter consists in constructing a *rolling control group* for each treated cohort. In practice, for each treated cohort between 2000 and 2006, we construct a separate sample where we define the time to treatment relative to that specific cohort, and where all the cohorts treated in later years serve as controls. The observations within control cohorts are only considered before their treatment. The cohort receiving BI in 2007 serves as pure control, since there is no cohort treated afterwards to create an additional sample. Finally, we append the 8 samples constructed using this method to run the following regression model:

$$Y_{c,s,t} = \alpha_0 Treat_{c,s} + \sum_{\tau=-k}^{k'} \alpha_\tau D^\tau \times Treat_{c,s} + \sum_{\tau=-k}^{k'} \beta_\tau D^\tau + \gamma_t \boldsymbol{X_{c,t}} + \nu_c + \pi_s + \varepsilon_{c,s,t} \quad (11)$$

Where $Treat_{c,s}$ takes the value of 1 if city c is treated in sample s. This parameter is identified despite the city fixed effects ν_c because the same city appears in multiple samples with both treated and control status. D^τ are dummies for time relative to treatment (equal to $\mathbb{1}\{t = t_c + \tau\}$ of equation 10) and the α_τ identify the pre- and post-treatment dynamic effects. We also add fixed effects for each one of the samples stacked (π_s). Standard errors are clustered at the department level, which also account for the error correlation generated by the repeated appearance of the same cities across different samples. This constitutes our preferred strategy that we apply throughout the paper, but in Appendix A we also report the results obtained from the standard two-way fixed effects model presented in equation (10) for comparison.

[29] Baker et al. (2021) show that the stacked difference in differences method gives similar results to the method proposed by Callaway and Sant'Anna (2020).

4 Empirical Evidence on BI and Outsourcing

In our theoretical model, BI acts as a general purpose technology. Once connected, firms become more productive and this raises their incentives to outsource occupations with the smaller degree of complementarity in production. In Appendix D we present evidence that confirms that BI is a skilled-biased technology. We use our empirical strategy described in the previous section to show that cities and firms that are connected become more productive, as measured by their value added scaled by their wage bill, and increase their share of high-skill workers (see Figure D1 and Table D1). Furthermore, individuals increase their wage once their firm receives access to BI, even conditional on a rich set of fixed-effects (commuting zone $\times$ year and individual). In particular, results in Table D2 show that, consistent with Akerman et al. (2015), wage dynamics are positively impacted by broadband expansion (+3% in the baseline specification) but that this average effect conceals a strong degree of heterogeneity between low skill (+1%) and high-skill workers (+5.2%).

In this section, we go beyond the predictions of the skilled-biased technological change literature and provide direct causal evidence that the arrival of BI is associated with an increased reliance to outsourced labor, which in turn increases the level of segregation of workers with different occupations across firms. The theory of skill-biased technological change states that technology substitutes routine workers while it increases the demand for high-skill workers, but it does not speak directly about workers' reshuffling across firms.

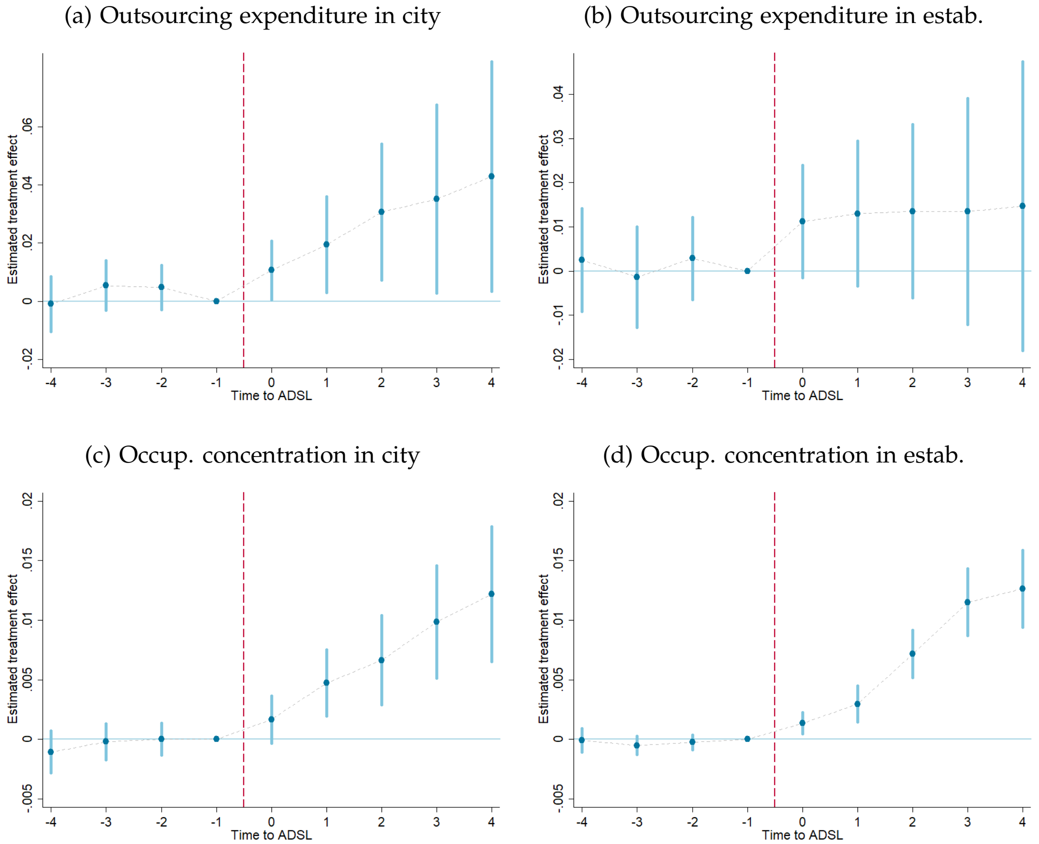
4.1 At the city and establishment level

We start by looking at the causal effect of BI on outsourcing expenditure and occupational segregation across firms. The first measures the value of outsourcing expenditure scaled by the wage bill of the firm. The second captures the HHI for the concentration of employment across occupations within each establishment, computed using the occupational share of the wage bill. At the city level, we compute the weighted average among the firms operating in the municipality.

Figure A3 in Appendix A summarizes the evolution of these two outcomes over the period. The expenditure on outsourcing over wage bill grew very steeply over the end of the 1990s, going from 0.38 in 1997 to 0.44 in 2000 (a growth of 16%), and then stabilized at this high plateau. The occupation concentration in establishments increased steadily from 0.33 in 1997 to 0.39 in 2007 (a growth of 18%). To test whether these trends are (partly) linked to the diffusion of BI, we run the empirical specification reported in equation (11) on these outcomes. Figure II presents the event study graphs at the city level and at the establishment level. Municipalities belonging to different cohorts of broadband expansion followed very similar trends before the arrival of the internet, but started spending more on outsourcing and became increasingly sorted after connection. This indicates that, after the arrival of BI, establishments within the city progressively specialize and employ fewer types of occupations in-house. At the same time, these establishments increasingly buy these services from other firms. Consequently, workers become increasingly segregated into firms that primarily hire their type. The establishment level results on outsourcing expenditure follow similar patterns but are smaller in magnitude, and become noisy after the first few years. This suggests that some of the effect operates through composition: newly created firms spend more in outsourcing relative to disappearing firms. On the other hand, establishment level results on occupational segregation are of similar magnitude than the city level ones.

Table A2 in Appendix A quantifies the effect at the city and establishment level by presenting the dynamic post-BI coefficients. Five years after the arrival of the ADSL in the city, we observe an average increase in outsourcing expenditure over wage bill of about 0.025, which corresponds to a growth of 9% with respect to the baseline levels. Occupational concentration increases by about 0.006 (1.6% growth relative to baseline). The effects at the establishment level are smaller and less precise on outsourcing expenditure: the latter increases by 0,012 (+3%) but it is marginally not significant, while occupation concentration increases by 0.007 (+1.9%) within existing establishments and the coefficient remains strongly significant. Table A3 and A4 show that these results are qualitatively similar if we control for the trends explained by pre-ADSL productivity differences between cities, which suggests that differences in the timing of ADSL diffusion across cities of similar density within the same department are not correlated with preexisting differences in city-level productivity growth. They also show the results obtained from a standard dynamic two-way fixed effects model as reported in equation (10) (coefficients are smaller and less significant), and report regressions excluding multi-establishment firms for the outsourcing intensity outcome (broadly consistent). Table A10 reports the coefficients obtained from static regressions where the treatment status is interacted with the post-ADSL period. The magnitude is slightly smaller than the average of the dynamic coefficients but remains in line with the main results.

FIGURE II. Outsourcing intensity and occupational concentration within establishments



(a) Outsourcing expenditure in city (b) Outsourcing expenditure in estab.

(c) Occup. concentration in city (d) Occup. concentration in estab.

Notes: This Figure shows regression coefficients and the 95% confidence intervals from a stacked event study design. The city-level specification follows equation 11, while the establishment-level specification follows the same logic but replaces city fixed effects with firm fixed effects.

Next, we focus on the evolution of the business service sector. We consider two main measures at the city level, which are computed separately for the high- and low-skill segments: (i) the share of total employment in the city concentrated within outsourcing sectors; and (ii) the share of outsourceable workers in the city that are employed in outsourcing sectors.[30] While the first is a measure of supply of business services that might react even in the absence of an increase in outsourcing, the second captures the change in the concentration of outsourceable workers within their respective services regardless of total volume change. All measures are computed on a full-time equivalent basis.

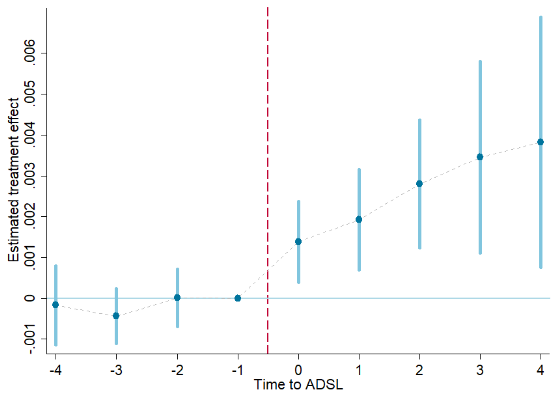
Figure A4 in Appendix A describes the evolution of these outcomes over time. Em-

[30] See Appendix B for a formal definition of the categories included in each group. High skill outsourcing services include management, advertising and HR consulting as well as IT services. Low-skill outsourcing services include security, cleaning, driving and logistics. We match these sectors to their corresponding occupations.

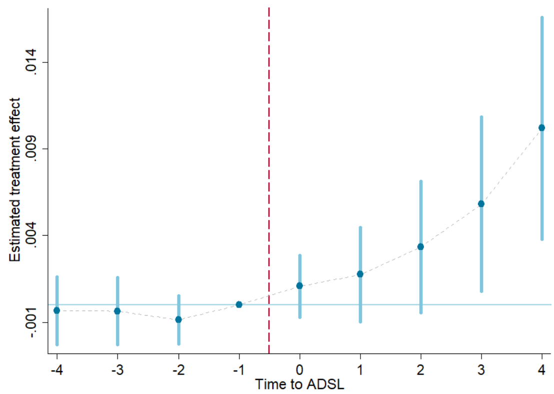
ployment in high-skill outsourcing services increased by almost 50% over the years of BI diffusion, going from 6.2% of the labor force in 1997 to 9.2% in 2007. The share of employment in low-skill outsourcing services increased by 30% during the same period, going from 9% in 1997 to 11.5% in 2007. Beyond the overall growth of these two emblematic outsourcing sectors, we also observe a shift of outsourceable occupations towards these sectors and away from the rest of the economy. The share of IT specialists and consultants employed by the high-skill service sector went from 17% in 1997 to 27% in 2007 (+59%). The share of security workers, cleaners, drivers and logistic workers employed in the low-skill service sector went from 30% in 1997 to 53% in 2007 (+77%). These changes are substantial. In what follows we test whether they have been, at least in part, facilitated by the diffusion of the ADSL technology over the French territory.

FIGURE III. Effect of ADSL on high- and low-skill outsourcing

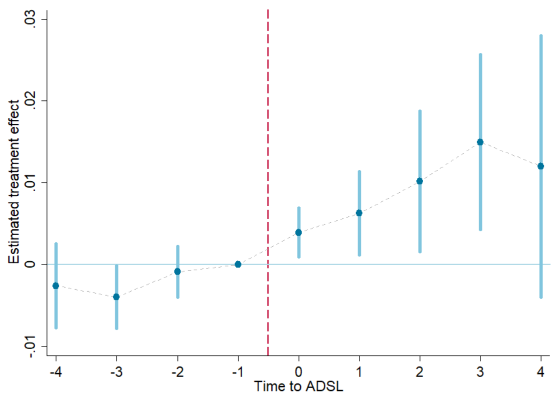
(a) Share of employment in HS outs. services



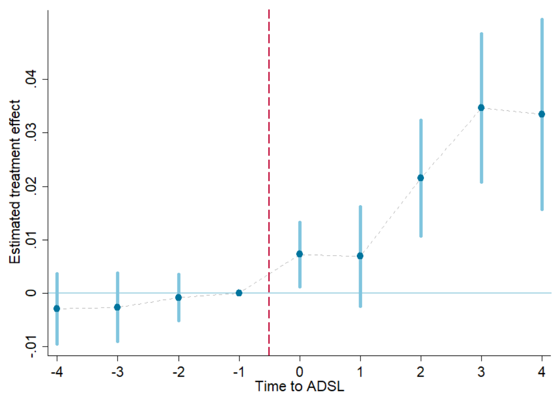
(b) Share of employment in LS outs. services



(c) Sh. outs. workers in HS outs. services



(d) Sh. outs. workers in LS outs. services



Notes: This Figure shows regression coefficients and 95% confidence intervals from a stack event study where the dependent variables are the share of workers employed in high-skill (low-skill) outsourcing services in a city at t (Figures 3(a) and 3(b)) and the share of outsourceable high-skill (low-skill) workers employed in their respective services in a city at t (Figures 3(c) and 3(d)). The regression follows equation 11.

Figure III reports the event study graphs of the two main outcomes, and shows them separately for high- and low-skill outsourcing. Table A5 in Appendix A shows the corresponding regression coefficients, Tables A6 and A7 show the coefficients obtained from breaking down even further each industry group into its sub-categories, Tables A8 and A9 present their robustness to the inclusion of city level productivity growth interacted with year dummies and the robustness to running a standard dynamic model. Table A10 reports the coefficients obtained from static regressions. We see that all of these four proxies for outsourcing were evolving similarly across cities connected at different times prior to the BI arrival, while the arrival of the ADSL increases them significantly afterwards.

While the magnitude of the coefficients might appear small, the effect is actually far from trivial when compared to baseline (pre-treatment) values.[31] The average share of city employment accounted for by high-skill outsourcing services was about 1.7% at the beginning of the period, such that the average impact of ADSL in the five years after its arrival amounts to growth of 16%. Low-skill outsourcing services accounted for about 10% of total city employment in 1997, and the BI connection led to a growth of 4.6%. When it comes to the concentration of workers in outsourceable occupations within their service sector, the baseline value is 6.5% for the high-skill ones, and the effect of the ADSL amounts to 1 additional percentage point (growth of 16%). For the low-skill ones, the baseline value is 18% and the effect of the ADSL is of 2 additional percentage points (growth of 12%). Given these magnitudes, we can infer that the arrival of the internet generated a structural change in the way these services are used by their business customers. This is an indication that BI catalyzed growth in domestic outsourcing, and this for non-core activities situated both at the low- and high-skill ends of the spectrum.

Finally, to get a more concrete sense of the magnitude of these results, we can wonder how outsourcing would have evolved in the absence of broadband internet. Since the main driver of increased outsourcing in our model is a productivity shock, we can imagine that other forces beyond BI, such as the diffusion of robots and computers for instance, are also behind the increased popularity of this practice. To answer this question we do some back-of-the-envelope calculations to compute the predicted trends in the main outcomes of interest after the subtraction of our estimated effects

[31] Here we report the baseline averages computed across cities, which are our units of interest in this analysis. The magnitudes differ from the evidence presented in the summary graphs above because the latter report the values computed in the overall population (weighting larger cities more).

of BI. We follow the procedure adopted in Malgouyres et al. (2019) and we compute the predicted outcome as the actual outcome minus the dynamic effects predicted by our semi-dynamic specification reported in Table A2. More specifically, we compute an average effect of broadband internet expansion for each year as:

$$\bar{a}_t = \sum_{t_0=1999}^{2007} w^y_{t_0,t'} \hat{\alpha}_{t-t_0}$$

where $w^y_{t_0,t'}$ represents the outcome y measured in t' for firms located in cities where BI became available at t_0, weighted by the employment share of these cities in the entire national economy. We postulate that the observed outcome y is given by a baseline level $y_t(0)$ that would have occurred in the absence of broadband diffusion multiplied by the predicted effect of BI: $y_t = \exp(\bar{a}_t) y_t(0)$. We then obtain the trends in the main outcomes in the absence of our estimated BI effect by inverting this relationship: $y_t(0) = (-\bar{a}_t) y_t$. This exercise is obviously not a proper counterfactual analysis since it assumes the complete absence of spill-overs and general equilibrium effects in the economy, which we suspect might be important in this case. However, we still believe that it can provide a useful benchmark to interpret the magnitude of our results. Figure IV shows that outsourcing expenditure over wage bill decreases after the beginning of the 2000s if we subtract the estimated BI effect, and by the end of the period is 9% smaller than what is actually observed. The occupational segregation across firms and the share of high- and low-skill outsourceable occupations in outsourcing sectors increase less steeply over the period when we subtract the estimated BI effect, reaching in 2007 a level that is respectively 2.5%, 4.9% and 6.6% smaller than what is observed in the data.

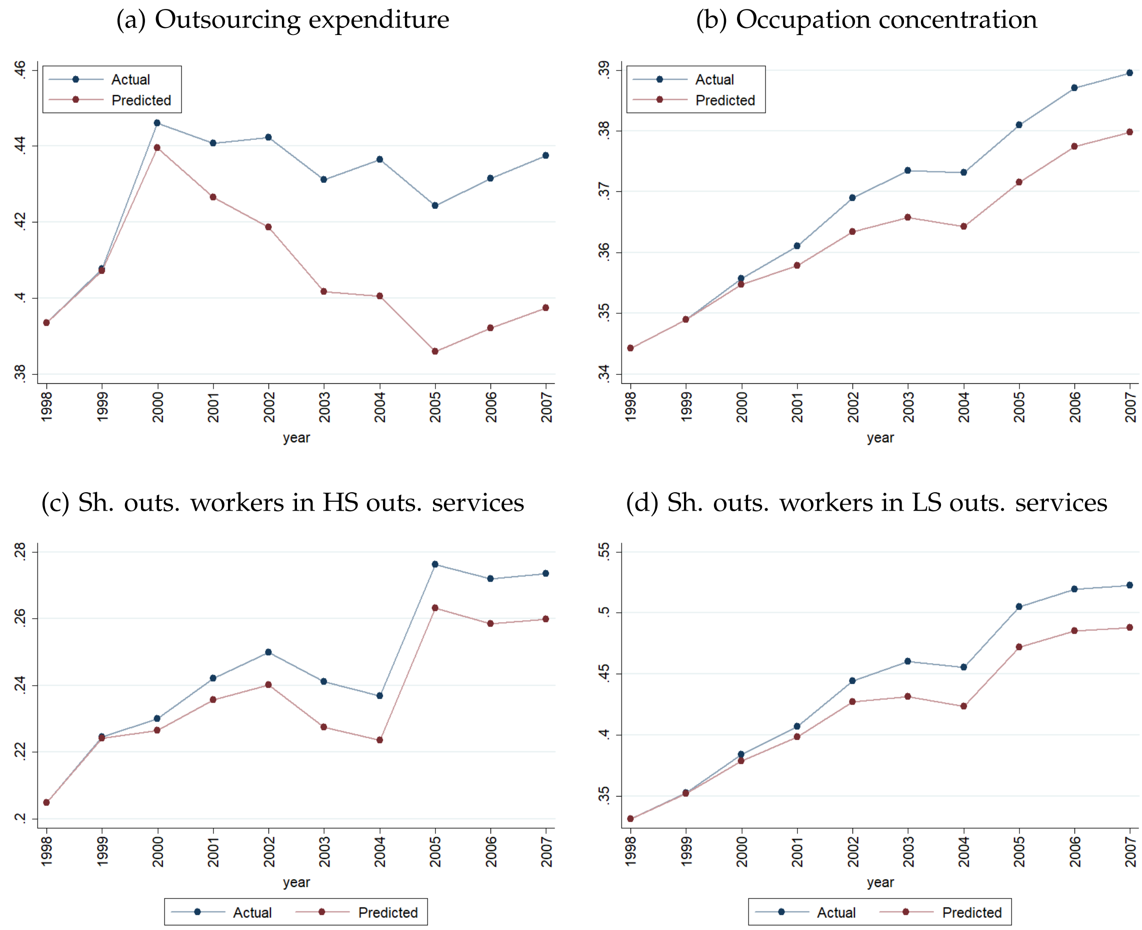
4.2 At the individual level

In this section, we leverage our panel data to follow the workers over time and study their mobility decisions. More precisely, we look at the effect of BI expansion on the probability that workers switch jobs. We consider different types of mobility: within/between the same city or labor market area (*Zone d'Emploi*) and/or from an establishment outside the outsourcing sector to an establishment in the outsourcing sector.

We therefore estimate the following linear probability model:

$$Move_{i,t} = \beta \tilde{Z}_{c(i),t} + X\gamma + \psi_{k,t} + \zeta_{s(i)} + \varepsilon_{i,t}, \tag{12}$$

FIGURE IV. Computation of the predicted aggregate trends after substracting our estimated BI effect



Notes: The actual outcomes are the weighted averages of the outcomes observed in our sample aggregated at the economy level. The predicted outcomes are obtained by subtracting the predicted effect of broadband internet to the actual outcomes. The latter is computed using a weighted average of the estimated α_τ, for $\tau > 0$ where the weights correspond to the share in national employment of each cohort of firms (i.e. all firms for which broadband expansion occurs the same year) measured the year of broadband expansion. The weights are normalized so that they sum to one.

where $Move_{i,t}$ is a binary variable equal to 1 if the worker i has moved on year t and 0 otherwise.[32] $\tilde{Z}$ has been defined above and measures the exposition to BI. X is a vector of standard time-varying individual characteristics usually included in wage regressions: age, age squared, an indicator of whether the job is part-time (as opposed to full-time). Finally, $\psi_{k,t}$ and $\zeta_{s(i)}$ are a set of labor market area k times year t fixed effects and sector $s(i)$ fixed effects. ε is an idiosyncratic error that we assume can be correlated within labor market areas but not across. β captures the effect of being connected to BI on the probability of moving, controlling for observable and

[32]Specifically, we set this dependent variable to 1 in year t if a worker is not in the same establishment in year $t+2$ when compared to year t. This is because workers sometimes disappear from the sample the year immediately following a mobility.

time-varying unobservable worker characteristics.

TABLE III. Effect of ADSL on workers' mobility

	High-skill workers			Low-skill workers		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Mobility	Any move	To outsourcing	Other	Any move	To outsourcing	Other
Any	0.063***	0.177**	0.066***	0.070***	0.086*	0.064***
	(0.020)	(0.081)	(0.022)	(0.012)	(0.046)	(0.013)
Same city	-0.013	-0.189	-0.004	0.257***	0.367***	0.262***
	(0.097)	(0.138)	(0.100)	(0.040)	(0.115)	(0.043)
Same LMA	-0.004	0.062	0.020	0.129***	0.164**	0.130***
	(0.038)	(0.084)	(0.033)	(0.017)	(0.066)	(0.017)
Different city	0.102**	0.314***	0.096	-0.021	0.033	-0.015
	(0.048)	(0.097)	(0.067)	(0.018)	(0.046)	(0.019)
Different LMA	0.174***	0.328**	0.132**	-0.040*	-0.003	-0.036
	(0.054)	(0.130)	(0.053)	(0.022)	(0.058)	(0.024)
Fixed Effects						
LMA × year	✓	✓	✓	✓	✓	✓
Sector	✓	✓	✓	✓	✓	✓
Obs.	1,054,063	1,054,063	1,054,063	5,143,010	5,143,010	5,143,010

Notes: This Table presents the point estimate and standard errors of coefficient β in equation (12). The dependent variable is a binary variable equal to 1 if worker i has changed establishment between t and $t+2$. Each line conditions on different type of mobility (any move, within/between city and within/between labor market area - *Zone d'emploi*). Columns (1) and (4) consider any move, columns (2) and (5) consider mobility to a firm in the outsourcing sector and columns (3) and (6) consider mobility to firms excluding the outsourcing sectors. The outsourcing sectors respectively denote the high-skill outsourcing sectors in column (2) (e.g. IT services) and the low-skill outsourcing sectors in column (5) (e.g. cleaning services). The dependent variable has been standardized by its sample mean. Columns (1), (2) and (3) restrict to high-skilled occupation workers and columns (4), (5) and (6) to low-skilled occupation workers. OLS estimator with standard errors clustered at the LMA level and robust to heteroskedasticity. Time period: 1995-2008.

Table III presents our results. We report coefficient β from equation (12) for different types of mobility. Columns (1) to (3) refer only to high-skill workers. In column (1), the dependent variable is equal to 1 in the case of a mobility, regardless of the sector of the destination firm. Column (2) is conditional on moving to a firm in any of the high-skill outsourcing sectors (IT, accounting etc.) and column (3) is conditional on moving to a firm outside these outsourcing sectors. Columns (4) to (6) do the same but for low-skill workers, and for the corresponding outsourcing sectors (cleaning, driving, security, etc.). Each line then considers different types of mobility: within a city, within a labor market area (*Zone d'emploi*), between cities and between labor market areas. We have standardized the dependent variable by its sample mean for each regression. Hence, the coefficient should be interpreted as a deviation in percentage points from the average probability of moving.

These results show that being connected to BI is associated with a greater propensity to move, both for high and low-skill workers with a larger relative effect for mobility to an outsourcing sector. Yet, the patterns of mobility are different across skill groups. Low-skill workers tend to move locally while high-skilled workers are more mobile and can move across different labor market areas. While these results suggest that control cities might be in part affected by the treatment through an increase in labor mobility, we argue that this effect is of one order of magnitude smaller relative to the main outcomes of interest at the city level, and thus can be considered as negligible. New entrant low-skill workers represent on average 0.5% of the overall employment of receiving cities and 1.3% of the outsourceable employment. New entrant high-skill workers represent on average 3.7% of the overall employment of receiving cities and 11% of outsourceable employment. In addition, the majority of the moves of both high-skill and low-skill workers take place towards cities that are already connected to BI starting from 2001, and thus do not affect controls (see Appendix table A11).

5 Effect of outsourcing on wages

Our findings suggest that BI expansion increased mobility of workers across establishments, with a particularly strong effect on the mobility of workers in outsourceable occupations towards firms in the outsourcing sectors. We define *individual* outsourcing events as follows:

- ***Individual outsourcing event.*** *The mobility of one worker from a firm outside the outsourcing sector to a firm belonging to the outsourcing sector.*

This definition allows us to capture many outsourcing events for high and low-skill workers respectively. We thus look in greater detail at the dynamics in wage and working conditions of workers around these events.

Table A12 in Appendix A describes the average wage and employment of outsourceable occupations across sectors, aiming at comparing their working conditions in the outsourcing sectors with the ones elsewhere. IT specialists and consultants earn a gross hourly wage of about 30€ in their outsourcing services, which is slightly higher than the one they gain in other services and in manufacturing (25€ and 28€ respectively). Security workers, cleaners, drivers and logistic workers, on the other hand, earn less when they are employed in their outsourcing services (12€) than when they are employed in other services or manufacturing (13€ and 15€ respectively). These characteristics are consistent with the idea that high-skill outsourcing may be voluntary, while low-skill outsourcing might not be.

To test the direct effect of outsourcing on individual wages, we leverage our individual panel data and follow outsourced workers before and after an individual outsourcing event. We restrict our analysis to the subset of workers that experience only one outsourcing event over the period of observation, which allows us to define our event study. We then look at the evolution of their hourly wage by estimating the following dynamic model:

$$\log(w_{it}^{ow}) = \sum_{\tau=-v}^{v'} \alpha_\tau \mathbb{1}\left\{t = t_i + \tau\right\} + X\gamma + \psi_{k,t} + \nu_i + \varepsilon_{i,t} \tag{13}$$

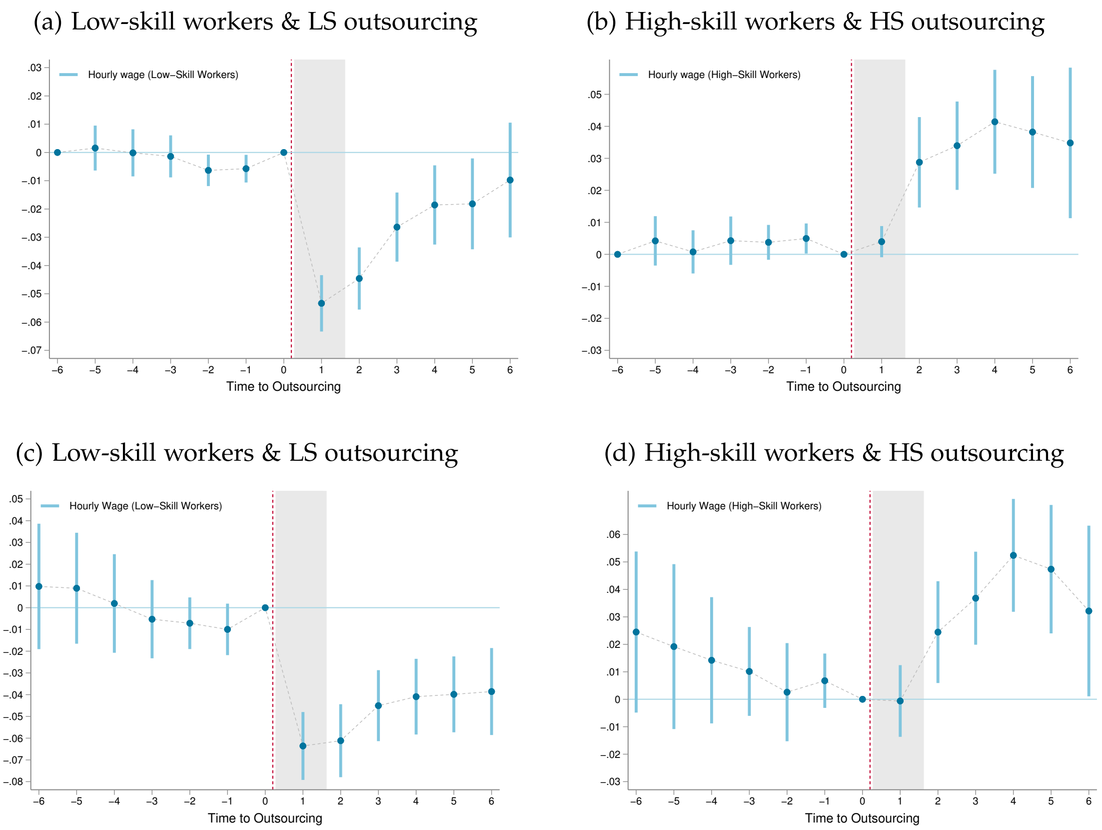
where w_{it}^{ow} represents the hourly wage of workers that are being outsourced by their firms, and t_i is the year of the event. X is a vector of time-varying individual characteristics: age, age squared and an indicator of whether the job is part-time and $\psi_{k,t}$ and ν_i are a set of labor market area k times year t fixed effects, and individual fixed effects.[33] ε is an idiosyncratic error that we assume can be correlated within labor market areas but not across. We estimate this regression separately for low-skill workers and high-skill workers, with a respective sample of 7,995 and 10,404 different workers. We consider $v = v' = 6$ which means that we follow the workers over 13 years. Finally, given that in this setting, every worker is treated, we standardize $\alpha_{-6} = \alpha_0 = 0$ (see Borusyak et al., 2021). We interpret the findings from this exercise as suggestive, as they rely on the relatively strong assumption that the timing of the outsourcing events are exogenous from the point of view of workers' wage trajectory, after controlling for observable characteristics and fixed effects. Results are presented in Figures 5(a) and 5(b).

This Figure shows that the hourly wage of outsourced low-skill workers sharply decreases after the outsourcing event. This finding is in line with the results evidenced by Goldschmidt and Schmieder (2017), which explain this phenomenon by the fact that firms in the outsourcing sectors benefit from lower rents, on average, than another 25 companies. This translates into lower wage premia for their employees.[34] We find an average effect that ranges from -4%, immediately after the outsourcing event, to -2% five years later, gradually converging towards the pre-treatment level. By contrast, outsourced high-skill workers enjoy a 4% gain in hourly wage in the long run after the outsourcing transition, a pattern that is consistent with broadband stimulating demand for IT services and thus resulting in an increase in IT workers' outside options and voluntary job-to-job transitions. These particular high-skill occupations indeed continue to be in high demand and can easily be contracted, especially as

[33] Adding sector $s(i)$ fixed effects does not affect our results.

[34] The shaded area corresponds to the year of the move, when the outsourced worker quit its previous employer to join an outsourcing firm. This transition period is known to create reporting errors of the number of hours worked and the wage in the data, resulting in a possible noisy estimate for the dummy α_1.

FIGURE V. Wage of outsourced workers before and after the outsourcing event



Notes: This Figure shows regression coefficients and confidence intervals at ± 2 standard errors from a dynamics event study where the dependent variables are the log hourly wage of workers being outsourced from the establishment at time t, and the regressors are dummies for the number of years before/after the establishment experiences an outsourcing event as well as control variables: age, age squared, short-time dummy and a set of year times labor market area and individual fixed effects. Shaded area denotes the year of the mobility which is associated with noisy measures of work duration. Top panels restrict to workers with one and only one mobility from a non-outsourcing firm to an outsourcing firms. Bottom panels add a control groups composed of workers who move from a any firm to a non-outsourcing firm. Left-hand side panels restrict to low skill workers (39,659 workers in panel (a) and 173,788 in panel (c)) and right-hand side to high skill workers (10,404 workers in panel (b) and 39,659 in panel (d)). Standard errors are computed using an heteroskedastic robust variance covariance estimators allowing for autocorrelation at the labor market area level.

communication technologies improve. By regrouping these occupations, specialized firms can serve different clients and generate more profit, first by maximizing the utilization rate of their inputs (mainly labor) and second by reducing fixed costs. This mechanism is similar in certain respects to some low-skill outsourcing, but in the case of IT specialists or consultants, the outsourced workers have greater bargaining power and can capture part of the rent by commanding a higher wage. Overall, these results emphasize a heterogeneous impact of outsourcing on workers' wages. While low-skill workers suffer from a significant wage loss, high-skill workers seem to benefit from outsourcing, as they are able to capture part of the increased profit made by firms as a result of this cost reduction.

To disentangle the specific effect of moving towards the outsourcing sector relative to other types of mobility, in Figures 5(c) and 5(d) we reproduce the same exercise but using the set of workers who transition to a different firm that do not belong to one of the outsourcing sectors as a control group. Then the coefficients should be interpreted as the difference in hourly wage compared to workers of similar skill level who move to another firm outside the outsourcing sector at the same point in time. In both cases, the results are consistent with the view that outsourced low skilled workers are negatively affected by an outsourcing event, while high skill outsourced workers are not, and that this effect goes beyond what could be observed in other types of moves.

6 Conclusion

The diffusion of the internet has fostered many changes in the way firms operate, some of which differ from those observed in the previous waves of technological change. In this paper we examine the role that broadband internet played in incentivizing firms to outsource some non-core activities, both in the high and low-skill segment, and we describe its consequences on the affected workers. We start by showing that these effects can be rationalized using a partial equilibrium model under a fairly minimal set of conditions. In the empirical analysis, we leverage the staggered roll-out of broadband connections across the French territory to adopt an event study design and estimate the causal effect of broadband internet diffusion. The latter compares similarly dense municipalities within a given department, which gain access to BI at different times. Our results show that the internet is not only skill biased, but also increases the degree of occupational concentration within establishments by pushing firms to outsource activities with lower degrees of complementarity in production. This phenomenon touches both low-skill and high-skill occupations. Finally, we provide suggestive evidence that high-skill workers experience wage gains through outsourcing, while low-skill workers experience wage losses. These findings confirm that the impact of internet technology is not homogeneous across the skill distribution, and reveal that domestic outsourcing is an additional mechanism through which this effect plays out. More broadly, it appears that these forces contribute to increasing the segregation of workers in the labor market of advanced countries, which might have detrimental effects on the level of trust and cohesion in our societies.

References

Abraham, Katharine G and Susan K Taylor, "Firms' use of outside contractors: Theory and evidence," *Journal of Labor Economics*, 1996, *14* (3), 394–424.

Abramovsky, Laura and Rachel Griffith, "Outsourcing and offshoring of business services: How important is ICT?," *Journal of the European Economic Association*, 2006, *4* (2-3), 594–601.

_ , _ , and Helen Miller, "Domestic effects of offshoring high-skilled jobs: Complementarities in knowledge production," *Review of International Economics*, 2017, *25* (1), 1–20.

Aghion, Philippe, Antonin Bergeaud, Richard Blundell, and Rachel Griffith, "The Innovation Premium to Soft Skills in Low Skill Occupations," Technical Report 14102, CEPR 2019.

_ , _ , Timo Boppart, Peter J Klenow, and Huiyu Li, "A theory of falling growth and rising rents," Technical Report w26448, National Bureau of Economic Research 2019.

_ , Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen, "On the returns to invention within firms: Evidence from Finland," in "AEA Papers and Proceedings," Vol. 108 2018, pp. 208–12.

Akerman, Anders, Edwin Leuven, and Magne Mogstad, "Information Frictions, Internet and the Relationship between Distance and Trade," Memorandum, Oslo University, Department of Economics February 2018.

_ , Ingvil Gaarder, and Magne Mogstad, "The skill complementarity of broadband internet," *Quarterly Journal of Economics*, 2015, *130* (4), 1781–1824.

Appelbaum, Eileen, Rosemary Batt, Annette Bernhardt, and Susan Houseman, "Domestic Outsourcing in the US: A Research Agenda to Assess Trends and Effects on Job Quality," in "Prepared for The Future of Work Symposium, US Department of Labor, Washington DC" 2015.

Autor, David H., Frank Levy, and Richard J. Murnane, "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

Autor, David H, Lawrence F Katz, and Alan B Krueger, "Computing inequality: have computers changed the labor market?," *The Quarterly journal of economics*, 1998, *113* (4), 1169–1213.

_ , _ , and Melissa S Kearney, "Trends in US wage inequality: Revising the revisionists," *The Review of economics and statistics*, 2008, *90* (2), 300–323.

Baker, Andrew, David F Larcker, and Charles CY Wang, "How Much Should We Trust Staggered Difference-In-Differences Estimates?," *Available at SSRN 3794018*, 2021.

Berlingieri, Giuseppe, "Outsourcing and the Rise in Services," Technical Report 1199, Centre for Economic Performance 2014.

Bernhardt, Annette, Rosemary Batt, Susan N Houseman, and Eileen Appelbaum, "Domestic outsourcing in the United States: a research agenda to assess trends and effects on job quality," Technical Report 16-253, Upjohn Institute Working Paper 2016.

Bhuller, Manudeep, Andreas Kostol, and Trond Vigtel, "How broadband internet affects labor market matching," *Available at SSRN 3507360*, 2019.

Bilal, Adrien and Hugo Lhuillier, "Outsourcing, Inequality and Aggregate Output," 2021. Mimeo Harvard.

Biscourp, Pierre and Francis Kramarz, "Employment, skill structure and international trade: Firm-level evidence for France," *Journal of International Economics*, 2007, *72* (1), 22–51.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen, "The distinct effects of information technology and communication technology on firm organization," *Management Science*, 2014, *60* (12), 2859–2885.

Borghans, Lex and Bas Ter Weel, "The division of labour, worker organisation, and technological change," *The Economic Journal*, 2006, *116* (509), F45–F72.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event study designs: Robust and efficient estimation," 2021. Unpublished working paper, version May.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg, "The anatomy of French production hierarchies," *Journal of Political Economy*, 2015, *123* (4), 809–852.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 2020.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Caroli, Eve and John Van Reenen, "Skill-biased organizational change? Evidence from a panel of British and French establishments," *Quarterly Journal of Economics*, 2001, *116* (4), 1449–1492.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Cortes, Guido Matias and Andrea Salvatori, "Delving into the demand side: Changes in workplace specialization and job polarization," *Labour economics*, 2019, *57*, 164–176.

Deming, David J and Kadeem Noray, "Earnings dynamics, changing job skills, and STEM careers," *The Quarterly Journal of Economics*, 2020, *135* (4), 1965–2005.

Deshpande, Manasi and Yue Li, "Who is screened out? application costs and the targeting of disability programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–48.

DeStefano, Timothy, Richard Kneller, and Jonathan Timmis, "Broadband infrastructure, ICT use and firm performance: Evidence for UK firms," *Journal of Economic Behavior & Organization*, 2018, *155*, 110–139.

Dorn, David, Johannes F Schmieder, and James R Spletzer, "Domestic Outsourcing of Labor Services in the United States: 1996-2015," 2018. Department of Labor.

Drenik, Andres, Simon Jäger, Miguel Pascuel Plotkin, and Benjamin Schoefer, "Paying outsourced labor: Direct evidence from linked temp agency-worker-client data," Technical Report 26891, National Bureau of Economic Research 2020.

Dube, Arindrajit and Ethan Kaplan, "Does outsourcing reduce wages in the low-wage service occupations? Evidence from janitors and guards," *ILR Review*, 2010, *63* (2), 287–306.

Fort, Teresa C, "Technology and production fragmentation: Domestic versus foreign sourcing," *Review of Economic Studies*, 2017, *84* (2), 650–687.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 2000, *108* (5), 874–904.

_ **and Esteban Rossi-Hansberg**, "Organization and inequality in a knowledge economy," *Quarterly Journal of Economics*, 2006, *121* (4), 1383–1435.

Godechot, Olivier, Paula Apascaritei, István Boza, Lasse Folke Henriksen, Are Skeie Hermansen, Feng Hou, Naomi Kodama, Alena Křížková, Jiwook Jung, Marta M Elvira et al., "The great separation: Top earner segregation at work in high-income countries," Technical Report, MaxPo Discussion Paper 2020.

Goldschmidt, Deborah and Johannes F Schmieder, "The rise of domestic outsourcing and the evolution of the German wage structure," *Quarterly Journal of Economics*, 2017, *132* (3), 1165–1217.

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," Technical Report, National Bureau of Economic Research 2018.

Hummels, David, Jakob R Munch, and Chong Xiang, "Offshoring and labor markets," *Journal of Economic Literature*, 2018, *56* (3), 981–1028.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who profits from patents? rent-sharing at innovative firms," *Quarterly Journal of Economics*, 2019, *134* (3), 1343–1404.

Kogut, Bruce and Udo Zander, "Knowledge of the firm, combinative capabilities, and the replication of technology," *Organization science*, 1992, *3* (3), 383–397.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials and rent sharing in the US labor market," Technical Report w25954, National Bureau of Economic Research 2019.

Le Moigne, Mathilde, "Exploring the "Fissured workplace": Internal job ladders' fragmentation and its effect on plants and workers," Technical Report, Unpublished Manuscript 2020.

Lepak, David P and Scott A Snell, "The human resource architecture: Toward a theory of human capital allocation and development," *Academy of management review*, 1999, *24* (1), 31–48.

Lewis, Tracy R and David EM Sappington, "Technological Change and the Boundaries of the Firm," *American Economic Review*, 1991, pp. 887–900.

Malgouyres, Clément, Thierry Mayer, and Clément Mazet-Sonilhac, "Technology-induced Trade Shocks? Evidence from Broadband Expansion in France," CEPR Discussion Papers 13847, C.E.P.R. Discussion Papers 2019.

Michaels, Guy, Ashwini Natraj, and John Van Reenen, "Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years," *Review of Economics and Statistics*, 2014, *96* (1), 60–77.

Oi, Walter Y and Todd L Idson, "Firm size and wages," *Handbook of labor economics*, 1999, *3*, 2165–2214.

Ono, Yukako, "Outsourcing business services and the role of central administrative offices," *Journal of Urban Economics*, 2003, *53* (3), 377–395.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *Quarterly Journal of Economics*, 2019, *134* (1), 1–50.

Spitz-Oener, Alexandra, "Technical change, job tasks, and rising educational demands: Looking outside the wage structure," *Journal of labor economics*, 2006, *24* (2), 235–270.

Vannutelli, Silvia, "From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance in Italy," *Job Market Paper*, 2020.

Weil, David, *The fissured workplace*, Harvard University Press, 2014.

Appendix

Outline

- Appendix A presents additional empirical results
- Appendix B presents the data in more details
- Appendix C presents the roll-out of broadband internet in France
- Appendix D presents evidence that broadband internet is skill-biased
- Appendix E presents the proofs and extension of the theoretical model

A Additional Results

A.1 Tables

TABLE A1. Correlation between outsourcing intensity and firm size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Outsourcing / wage bill				Outsourcing		Outsourcing / (wage bill + outsourcing)			
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
ln (net revenues)	0.173***	0.206***								
	(0.00461)	(0.00474)								
ln (Value added)			0.0809***	0.0472***						
			(0.00367)	(0.00377)						
Value added					0.138***	0.134***				
					(0.00189)	(0.00187)				
Value added / wage bill							0.0199***	0.0264***		
							(0.000720)	(0.000678)		
Value added / (wage bill + outsourcing)									-0.144***	-0.130***
									(0.00103)	(0.000929)
Observations	1,040,578	1,040,578	1,030,298	1,030,298	1,072,414	1,072,414	1,042,898	1,042,898	1,043,250	1,043,250
R-squared	0.017	0.080	0.003	0.062	0.214	0.261	0.007	0.157	0.257	0.333
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sector FE		✓		✓		✓		✓		✓

Notes: Regressions (1) to (4) correlate the size of the firm, measured as the natural logarithm of net revenues or value added, and the outsourcing intensity of the firm. Regressions (5) and (6) correlate value added and outsourcing expenditure in absolute terms. Regressions (7) and (8) correlate value added over wage bill and the share of expenditure accounted for by outsourcing. Finally, regressions (9) and (10) correlate value added over total labor cost and the share of expenditure accounted for by outsourcing. Standard errors clustered at the firm level, period of analysis: 1995 - 2007. Columns (1), (3), (5), (7) and (9) only control for year fixed effects, columns (2), (4), (6), (8) and (10) add sector fixed effects.

TABLE A2. Effect of ADSL on outsourcing expenditure and occupational sorting

	(1)	(2)	(3)	(4)
	Outsourcing / wage bill		Occup. concentration (wage bill HHI)	
VARIABLES	City level	Estab. Level	City level	Estab. Level
T = 0	0.00839	0.0104*	0.00128	0.00157***
	(0.00539)	(0.00597)	(0.00118)	(0.000471)
T = +1	0.0171**	0.0121	0.00418**	0.00318***
	(0.00838)	(0.00808)	(0.00162)	(0.000797)
T = +2	0.0284**	0.0127	0.00693***	0.00736***
	(0.0120)	(0.00991)	(0.00196)	(0.00104)
T = +3	0.0328*	0.0126	0.00906***	0.0117***
	(0.0166)	(0.0130)	(0.00273)	(0.00145)
T = +4	0.0404**	0.0139	0.00981***	0.0128***
	(0.0203)	(0.0167)	(0.00321)	(0.00168)
Average effect	0.0254**	0.0123	0.00625***	0.00733***
	(0.0119)	(0.00996)	(0.00202)	(0.000988)
Observations	233,656	1,267,316	423,770	3,077,125
R-squared	0.701	0.760	0.779	0.846

Notes: Columns (1) and (3) run the regression at the city level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level. Columns (2) to (4) run the same specification on the outcome computed at the establishment level, replacing city fixed effects by establishment fixed effects.

TABLE A3. Robustness of effect of ADSL on outsourcing expenditure

	(1)	(2)	(3)	(4)	(5)	(6)
	Outsourcing / wage bill					
VARIABLES	City level			Estab. Level		
T = 0	0.00838	-0.00507	0.00742	0.0104*	0.00448	0.00967*
	(0.00539)	(0.00567)	(0.00496)	(0.00597)	(0.00485)	(0.00577)
T = +1	0.0171**	-0.00482	0.0132*	0.0121	0.00468	0.0139*
	(0.00837)	(0.00838)	(0.00678)	(0.00808)	(0.00806)	(0.00775)
T = +2	0.0284**	-0.00327	0.0186*	0.0127	0.00234	0.0128
	(0.0120)	(0.0108)	(0.0102)	(0.00991)	(0.00996)	(0.0101)
T = +3	0.0329**	-0.00859	0.0207	0.0126	0.00153	0.00702
	(0.0165)	(0.0137)	(0.0125)	(0.0130)	(0.0127)	(0.0131)
T = +4	0.0405**	-0.00670	0.0239	0.0138	0.000293	0.00896
	(0.0203)	(0.0163)	(0.0152)	(0.0167)	(0.0147)	(0.0175)
Control: pre-BI prod. growth	✓			✓		
Staggered DiD strategy		✓			✓	
Mono-estab. only			✓			✓
Average effect	0.0255**	-0.00569	0.0168*	0.0123	0.00266	0.0105
	(0.0119)	(0.0101)	(0.00908)	(0.00996)	(0.00954)	(0.0100)
Observations	233,656	102,012	223,482	1,267,316	698,257	1,090,158
R-squared	0.701	0.666	0.702	0.760	0.727	0.758

Notes: Columns (1) and (4) run the regression at the city and establishment level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, the productivity growth at the city level pre-BI (1996-1998) interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects. Columns (2) and (5) run the regression run the regression following the standard staggered difference-in-differences strategy reported in equation 10, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, and city (establishment) fixed effects. Columns (3) and (6) run the regression at the city and establishment level restricting the sample to mono-establishment firms. They follow equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects.

TABLE A4. Robustness of effect of ADSL on occupational sorting

	(1)	(2)	(3)	(4)
	Occupational concentration (wage bill HHI)			
VARIABLES	City level		Estab. Level	
T = 0	0.00174	0.000999	0.00153***	0.00149***
	(0.00113)	(0.00113)	(0.000472)	(0.000501)
T = +1	0.00512***	0.00353**	0.00291***	0.00290***
	(0.00163)	(0.00150)	(0.000777)	(0.000802)
T = +2	0.00906***	0.00653***	0.00725***	0.00666***
	(0.00205)	(0.00176)	(0.00106)	(0.00113)
T = +3	0.0122***	0.00856***	0.0119***	0.0109***
	(0.00270)	(0.00243)	(0.00149)	(0.00142)
T = +4	0.0124***	0.00981***	0.0131***	0.0133***
	(0.00331)	(0.00290)	(0.00178)	(0.00176)
Control: pre-BI prod. growth.	✓		✓	
Staggered DiD strategy		✓		✓
Average effect	0.00810***	0.00589***	0.00733***	0.00706***
	(0.00204)	(0.00183)	(0.00102)	(0.00104)
Observations	316,407	140,751	2,936,774	1,607,237
R-squared	0.775	0.759	0.846	0.822

Notes: Columns (1) and (3) run the regression at the city and establishment level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, the productivity growth at the city level pre-BI (1996-1998) interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects. Columns (2) and (4) run the regression run the regression following the standard staggered difference-in-differences strategy reported in equation 10, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, and city (establishment) fixed effects.

TABLE A5. Effect of ADSL on high and low-skill outsourcing

	(1)	(2)	(3)	(4)
	High-skill outsourcing		Low-skill outsourcing	
	Sh. of empl. in HS outs. services	Sh. outs. workers in HS outs. services	Sh. of empl. in LS outs. services	Sh. outs. workers in LS outs. services
T = 0	0.00150***	0.00539***	0.00142	0.00843**
	(0.000518)	(0.00189)	(0.00101)	(0.00359)
T = +1	0.00204***	0.00766**	0.00210	0.00799
	(0.000637)	(0.00293)	(0.00149)	(0.00518)
T = +2	0.00291***	0.0115**	0.00370*	0.0226***
	(0.000816)	(0.00465)	(0.00203)	(0.00596)
T = +3	0.00357***	0.0164***	0.00615**	0.0358***
	(0.00122)	(0.00576)	(0.00267)	(0.00735)
T = +4	0.00394**	0.0134	0.0106***	0.0347***
	(0.00159)	(0.00839)	(0.00334)	(0.00929)
Average effect	0.00279***	0.0109**	0.00479**	0.0219***
	(0.000884)	(0.00438)	(0.00198)	(0.00559)
Observations	423,770	164,880	423,770	188,496
R-squared	0.821	0.727	0.902	0.798

Notes: The regressions are run at the city level following equation 11. All columns control for the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

TABLE A6. Effect of ADSL on share of employment in outsourcing services by sub-category

	(1)	(2)	(3)	(4)	(5)	(6)
	Sh. of empl. in HS outs. Services		Sh. of empl. in LS outs. Services			
	IT services	Consulting, advertising & HR services	Security services	Cleaning services	Driving services	Logistics services
T = 0	0.000451	0.00105***	-0.000145	0.000467	0.00120	-0.000104
	(0.000297)	(0.000395)	(0.000265)	(0.000301)	(0.000914)	(0.000319)
T = +1	0.000615	0.00142***	-0.000371	0.000435	0.00187	0.000169
	(0.000372)	(0.000489)	(0.000387)	(0.000453)	(0.00141)	(0.000479)
T = +2	0.000803	0.00211***	-0.000290	0.000905	0.00246	0.000627
	(0.000513)	(0.000609)	(0.000550)	(0.000565)	(0.00190)	(0.000655)
T = +3	0.000975	0.00259***	-0.000309	0.00232***	0.00371	0.000427
	(0.000684)	(0.000888)	(0.000705)	(0.000692)	(0.00250)	(0.000884)
T = +4	0.00114	0.00280**	-9.28e-06	0.00329***	0.00631**	0.000976
	(0.000902)	(0.00111)	(0.000862)	(0.000857)	(0.00313)	(0.00116)
Average effect	0.000798	0.00200***	-0.000225	0.00148***	0.00311*	0.000419
	(0.000525)	(0.000624)	(0.000522)	(0.000514)	(0.00186)	(0.000671)
Observations	423,770	423,770	423,770	423,770	423,770	423,770
R-squared	0.827	0.814	0.796	0.860	0.909	0.845

Notes: The regressions are run at the city level following equation 11. All columns control for the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

TABLE A7. Effect of ADSL on share of outs. workers in outsourcing services by sub-category

	(1)	(2)	(3)	(4)	(5)	(6)
	Sh. outsourceable workers in HS outs. Services		Sh. outsourceable workers in LS outs. Services			
	IT specialists in IT services	Admin, sales and HR specialists in consulting, advertising & HR services	Security guards in security services	Cleaners in cleaning services	Drivers in driving services	Maintenance and warehouse workers in logistics services
T = 0	0.0121***	0.00481***	0.0290***	0.0373***	0.0103***	-0.00151
	(0.00381)	(0.00182)	(0.00702)	(0.00469)	(0.00383)	(0.00209)
T = +1	0.0211***	0.00690**	0.0505***	0.0652***	0.00837	-0.00241
	(0.00609)	(0.00300)	(0.0108)	(0.00620)	(0.00545)	(0.00356)
T = +2	0.0305***	0.0105**	0.0880***	0.0990***	0.0110	-0.00199
	(0.00839)	(0.00455)	(0.0144)	(0.00855)	(0.00736)	(0.00523)
T = +3	0.0327***	0.0157***	0.133***	0.128***	0.0159*	-0.00516
	(0.0122)	(0.00578)	(0.0197)	(0.0120)	(0.00911)	(0.00711)
T = +4	0.0337*	0.0141*	0.143***	0.158***	0.00406	-0.00760
	(0.0186)	(0.00831)	(0.0284)	(0.0152)	(0.0122)	(0.00950)
Average effect	0.0260***	0.0104**	0.0888***	0.0974***	0.00992	-0.00373
	(0.00909)	(0.00436)	(0.0137)	(0.00855)	(0.00672)	(0.00533)
Observations	56,593	162,706	28,422	89,278	129,649	134,260
R-squared	0.793	0.715	0.823	0.756	0.824	0.803

Notes: The regressions are run at the city level following equation 11. All columns control for the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

TABLE A8. Robustness of effect of ADSL on high-skill outsourcing

	(1)	(2)	(3)	(4)
	High-skill outsourcing			
VARIABLES	Sh. of empl. in HS outs. Services		Sh. outs. workers in HS outs. Services	
T = 0	0.00139***	0.00120**	0.00595***	0.00434**
	(0.000506)	(0.000485)	(0.00218)	(0.00186)
T = +1	0.00164**	0.00182***	0.00891***	0.00622**
	(0.000659)	(0.000657)	(0.00337)	(0.00281)
T = +2	0.00311***	0.00250***	0.0133**	0.0100**
	(0.000832)	(0.000832)	(0.00518)	(0.00435)
T = +3	0.00389***	0.00282**	0.0194***	0.0149**
	(0.00121)	(0.00126)	(0.00621)	(0.00567)
T = +4	0.00425***	0.00311**	0.0190**	0.0144*
	(0.00156)	(0.00156)	(0.00894)	(0.00762)
Control: pre-BI prod. growth	✓		✓	
Staggered DiD strategy		✓		✓
Average effect	0.00286***	0.00229**	0.0133***	0.00997**
	(0.000877)	(0.000897)	(0.00481)	(0.00413)
Observations	316,407	140,751	148,444	63,452
R-squared	0.787	0.798	0.713	0.701

Notes: Columns (1) and (3) run the regression at the city and establishment level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, the productivity growth at the city level pre-BI (1996-1998) interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects. Columns (2) and (4) run the regression run the regression following the standard staggered difference-in-differences strategy reported in equation 10, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, and city (establishment) fixed effects.

TABLE A9. Robustness of effect of ADSL on low-skill outsourcing

	(1)	(2)	(3)	(4)
	Low-skill outsourcing			
VARIABLES	Sh. of empl. in LS outs. Services		Sh. outs. workers in LS outs. Services	
T = 0	0.00165	0.00175*	0.00736*	0.00682*
	(0.00114)	(0.000975)	(0.00381)	(0.00345)
T = +1	0.00249	0.00315**	0.00644	0.00615
	(0.00160)	(0.00157)	(0.00544)	(0.00521)
T = +2	0.00400*	0.00501**	0.0210***	0.0214***
	(0.00206)	(0.00211)	(0.00603)	(0.00609)
T = +3	0.00734***	0.00781***	0.0356***	0.0364***
	(0.00275)	(0.00280)	(0.00748)	(0.00754)
T = +4	0.0112***	0.0126***	0.0330***	0.0374***
	(0.00371)	(0.00356)	(0.00981)	(0.00895)
Control: pre-BI prod. growth	✓		✓	
Staggered DiD strategy		✓		✓
Average effect	0.00534**	0.00606***	0.0207***	0.0216***
	(0.00212)	(0.00210)	(0.00575)	(0.00565)
Observations	316,407	140,751	166,271	69,869
R-squared	0.895	0.884	0.780	0.754

Notes: Columns (1) and (3) run the regression at the city and establishment level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, the productivity growth at the city level pre-BI (1996-1998) interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects. Columns (2) and (4) run the regression run the regression following the standard staggered difference-in-differences strategy reported in equation 10, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, and city (establishment) fixed effects.

TABLE A10. Effect of ADSL on outsourcing outcomes - static regressions

	Outsourcing / wage bill	wage bill HHI	Sh. of empl. in HS outs. services	Sh. outs. workers in HS outs. Services	Sh. of empl. in LS outs. services	Sh. Outs. workers in LS outs. Services
Panel A : city level regressions						
Post ADSL * treated	0.0161**	0.00336**	0.00200***	0.00707***	0.00201	0.0103**
	(0.00743)	(0.00144)	(0.000573)	(0.00250)	(0.00131)	(0.00408)
Observations	233,656	423,770	423,770	164,880	423,770	188,496
R-squared	0.701	0.779	0.821	0.727	0.902	0.798
Panel B : establishment level regressions						
Post ADSL * treated	0.0110	0.00284***	-	-	-	-
	(0.00668)	(0.000609)	-	-	-	-
Observations	1,267,316	3,077,125	-	-	-	-
R-squared	0.760	0.846	-	-	-	-

Notes: The regressions are run at the city and establishment level following a model similar to equation 11, but where instead of including the dynamic post-ADSL effects for every year, we just include a dummy for post-ADSL period interacted with the treatment indicator. All columns control for the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects.

TABLE A11. Descriptive statistics of the share of mobilities going to connected cities

Year	High Skill workers		Low Skill workers	
	Movements to outsourcing	Movements to non-outsourcing	Movements to outsourcing	Movements to non-outsourcing
2000	23%	34%	23%	25%
2001	63%	86%	63%	67%
2002	82%	93%	81%	84%
2003	88%	95%	88%	89%
2004	92%	98%	92%	93%
2005	97%	99%	95%	97%
2006	99%	99%	99%	99%

Notes: Summary statistics computing the share of mobility towards both the outsourcing and non-outsourcing sectors that has for destination a city that is already connected by broadband internet.

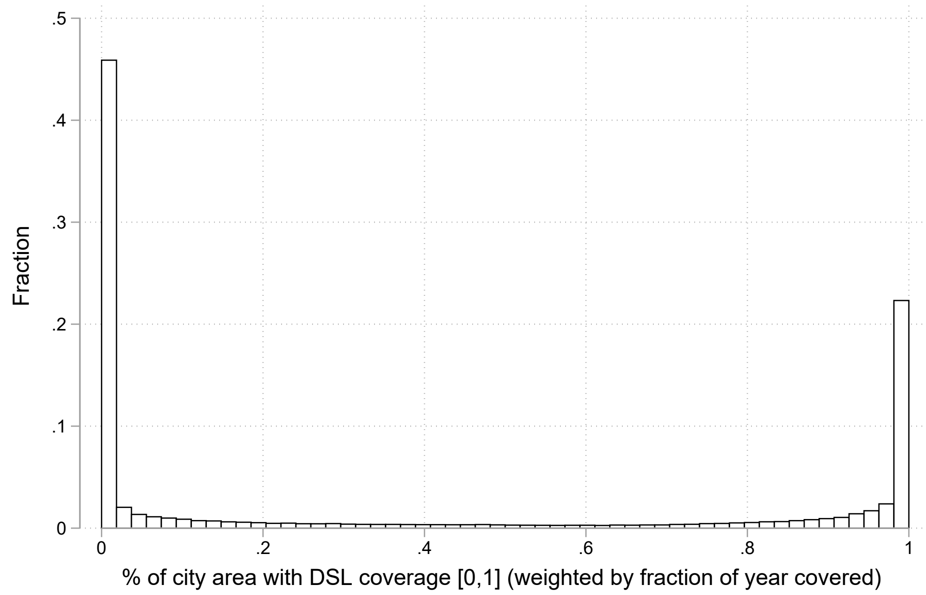
TABLE A12. Average wage and employment in outsourceable occupations across sectors

		outsourcing services	Other services	Manufacturing
		mean/(sd)	mean/(sd)	mean/(sd)
Gross hourly wage workers in HS outsourceable occup. (2010 euros)	overall	30,7	24,8	28,6
		(13,2)	(11,9)	(12,2)
	pre-BI	29,6	22,6	27,1
		(11,7)	(10,7)	(9,2)
	post-BI	31,1	26,0	29,8
		(13,6)	(12,3)	(13,9)
N. of workers in HS outsourceable occup. Per establishment	overall	179,1	42,1	69,5
		(414,6)	(145,4)	(211,7)
	pre-BI	196,7	48,6	65,0
		(369,9)	(154,4)	(188,7)
	post-BI	173,3	38,7	73,0
		(428,2)	(140,4)	(228,0)
Gross hourly wage workers in LS outsourceable occup. (2010 euros)	overall	12,1	13,4	15,3
		(3,4)	(3,9)	(4,8)
	pre-BI	11,6	12,7	14,5
		(3,4)	(3,7)	(3,9)
	post-BI	12,3	13,9	16,1
		(3,4)	(4,0)	(5,2)
N. of workers in LS outsourceable occup. per establishment	overall	154,0	23,7	50,5
		(277,2)	(58,7)	(161,9)
	pre-BI	154,3	27,1	56,7
		(279,0)	(51,2)	(172,8)
	post-BI	153,9	21,6	45,1
		(276,4)	(62,9)	(151,8)

Notes: Summary statistics comparing wages and employment of outsourceable workers (high-skill and low-skill) across different sectors.

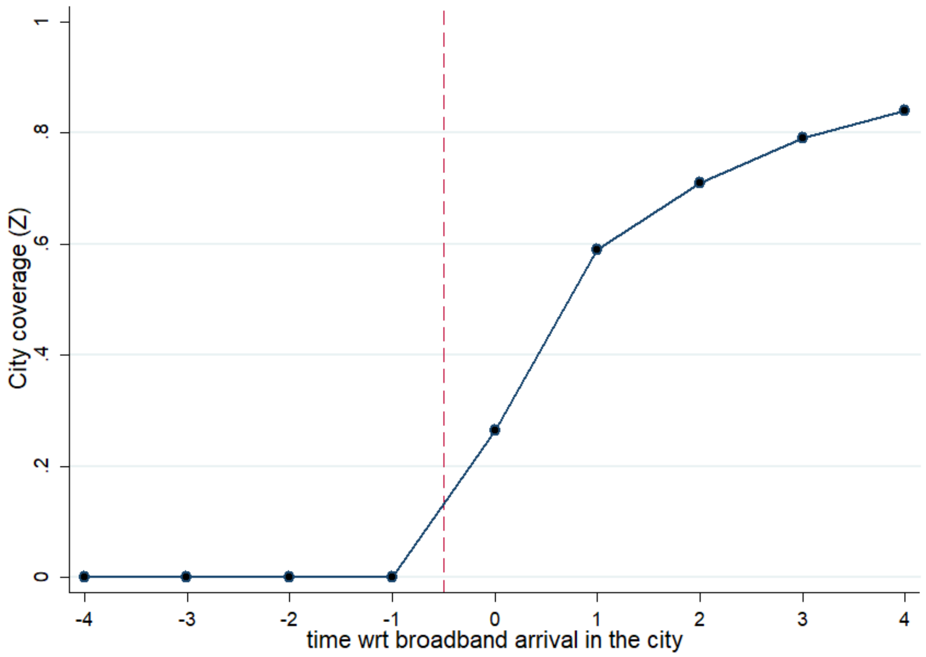
A.2 Figures

FIGURE A1. Distribution of $\widetilde{Z}_{it}$: 1999-2007



Notes: This figure plots the distribution of the continuous measure of local broadband availability (variable $\widetilde{Z}$) as defined in Equation (8). We see that while the measure is continuous and contained between 0 and 1 but presents point of accumulation on 0 and 1.

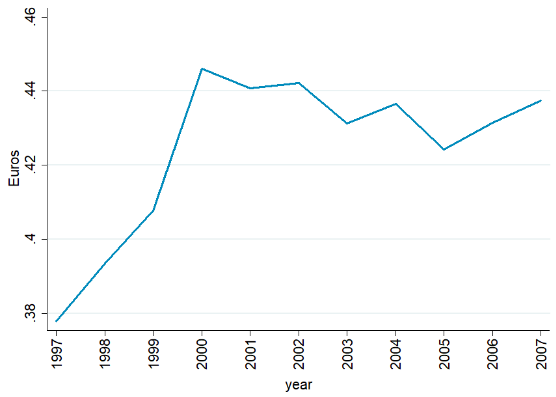
FIGURE A2. Evolution of $\widetilde{Z}_{it}$ before and after the discrete event



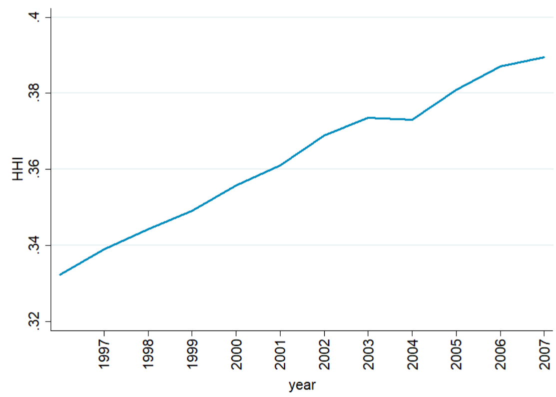
Notes: This figure plots the average of the continuous measure of local broadband availability (variable $\widetilde{Z}$) along the time to event variable, where event is the first year where $\widetilde{Z} > 0$.

FIGURE A3. Overall trends in outsourcing expenditure and occupation concentration

(a) Outsourcing expenditure over wage bill



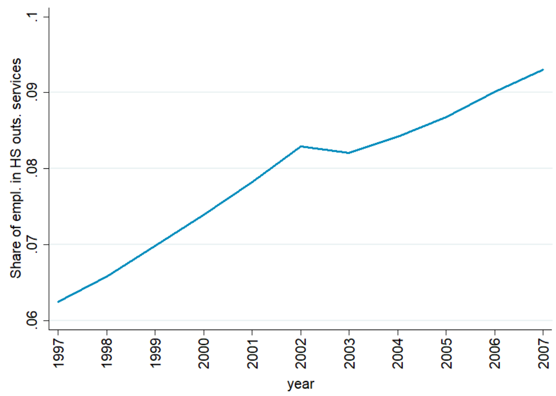
(b) Occup. concentration (HHI over wage bill)



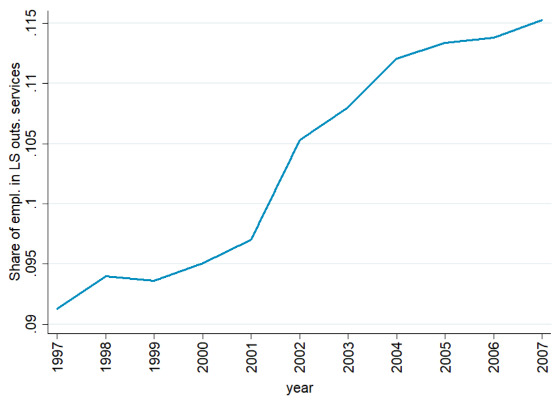
Notes: This Figure shows the evolution over time of the main outcomes of interest in the outsourcing analysis.

FIGURE A4. Overall trends in high and low skill outsourcing

(a) Share of employment in HS outs. services

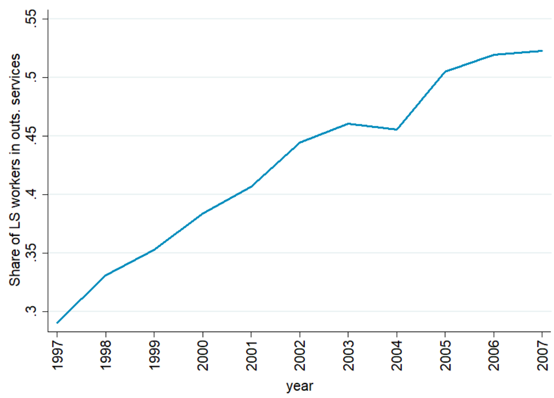


(b) Share of employment in LS outs. services



(c) Share of outs. workers in HS outs. services

(d) Share of outs. workers in LS outs. services



Notes: This Figure shows the evolution over time of the main outcomes of interest in the outsourcing analysis.

B Data Appendix: Administrative Employer-Employee Data

Our main analysis relies on data from the administrative records used by the French government to compute payroll taxes. Our period of analysis spans from 1996 to 2007. The first year is chosen to include a few years prior to the beginning of broadband diffusion, which started in 1999, while the last year corresponds to the final year of broadband expansion. We chose not to include later years because of the effect of the financial crisis. These data are collected yearly by *INSEE* (the French statistics office) and are known as *DADS* ("Déclarations annuelles des données sociales"). The main dataset contains information on all existing work contracts for each establishment in each firm operating in the French territory. The latter allows us to monitor establishments and firms over time but not workers, with the exception of a one-year worker panel dimension available since 2002. This is the main source that we use for the city and firm-level analyses. For the worker-level analysis, we rely on a subsample of this data from the DADS Panel. The latter randomly selects 1/24 of the labor force and follows it across its employment over the entire period. The random selection is achieved through the inclusion of all workers born in October of an even year. The raw data provided to researchers has already undergone substantial verification, and consequently only requires a minimal amount of additional cleaning. For this study, we focus on workers with some degree of attachment to the labor market ("postes non-annexes"), which are defined as contracts involving either more than 120 hours of work or more than 30 days of work, with more than 1.5 hours of work per day, or contracts that paid more than 3 times the monthly minimum wage over the year. We also exclude firms with less than 10 employees, to avoid taking family-run companies into consideration and thus focus on formal businesses. We further exclude some occupations and industries since we are interested only in the private sector. In the following bullet points we specify the excluded occupations by their PCS-2003 classification codes and the excluded industries based on the NAF rev. 1 classification. Given that both of these classifications changed in the middle of our sample (2002), we use official crosswalk tables to identify the same groups between years.

- **Selection of occupations:** We exclude all categories of non-employed people OA-11 (cs 2 [7, 9]) and self-employed farmers (pcs = 1). We further exclude self-employed crafts workers (pcs = 20), liberal professions (pcs = 31), university professors (pcs = 34), school teachers (pcs = 42) and the clergy (pcs = 44).

- **Selection of industries:** We exclude mining and farming ($NAF \in [1, 9]$), utilities ($NAF \in [35, 39]$), the entire public sector ($NAF \in [84, 88]$), and social services ($NAF \geq 90$).

- **High-skill workers**: We define high-skill workers as those belonging to the category including CEOs and the category including executives, managers and engineers (pcs = 2 and pcs = 3).

Once this cleaning is completed, we define the main categories used in the outsourcing analysis as reported in Table B1. For the low skill categorization we follow the categories proposed by Goldschmidt and Schmieder (2017), but we exclude food services because in the PCS classification of occupations it is impossible to distinguish canteen workers from the much larger category of waiters and restaurant workers. The remaining ones are security, cleaning, driving and logistics. For the high-skill categorization we base ourselves on the two largest industry categories that provide professional services to other firms: IT and consulting (which includes strategy consulting, HR and advertising).

TABLE B1. Categorization of outsourceable occupations and outsourcing sectors

High-Skill Outsourcing		
	Outsourcing sectors	sub-category
NAF = 72	IT services	IT
NAF = 74.1	Admin services, management consulting	consulting
NAF = 74.4	Advertising	consulting
NAF = 74.5	HR services	consulting
	Outsourceable occupations	sub-category
PCS = 388	IT engineers	IT
PCS = 478	IT technicians	IT
PCS = 372	HR executives	consulting
PCS = 373	Admin. Executives	consulting
PCS = 461	Admin. Support staff	consulting
PCS = 375	Advertising executives	consulting
PCS = 464a	Advertising and PR support staff	consulting
Low-Skill Outsourcing		
	Outsourcing sectors	sub-category
NAF = 74.6	Security	security
NAF = 74.7	Cleaning	cleaning
NAF = 60.2	urban and road transportation	driving
NAF = 63.1	Maintenance and storage	logistics
NAF = 63.4	Logistics of merchandise transportation	logistics
	Outsourceable occupations	sub-category
PCS = 533, 534	Security guards	security
PCS = 684	Cleaners	cleaning
PCS = 641a	Road drivers	driving
PCS = 643a	Delivery personnel	driving
PCS = 651	Storage machine operator	logistics
PCS = 652	Maintenance worker	logistics
PCS = 653	Warehouse workers	logistics

Notes: List of outsourcing sectors providing services to other firms, and of outsourceable occupations that are employed by them. We broadly categorize them into high- and low-skill services, where the first includes IT and consulting activities, while the second includes security, cleaning, driving and logistics.

C ADSL in France

ADSL (Asymmetric Digital Subscriber Line) is a data communication technology that enables fast data transmission over copper telephone lines: bandwidth and bit rate are said to be asymmetric, meaning that they are greater towards the customer premises (downstream) than the reverse (upstream). Eligibility for ADSL depends on the distance between the final customer and a Local Exchange (LE), since the intensity and the quality of the analogue signal decreases as it is routed over the copper lines. LEs are telephone exchanges owned by the incumbent operator France Télécom into which subscribers' telephone lines connect. Initially dedicated to the telephone network, LEs are essential for internet users who subscribe to ADSL. LEs aggregate local traffic and then direct it via the so-called backbone (i.e. higher levels of the network) towards the World Wide Web. A key feature of ADSL technology is that one can supply high-speed internet by upgrading the LE while relying on the existing (copper) local loop to connect the premises of the final customers. The upgrading involves the installation of equipment inside the LE (a DSLAM) required in order to translate the analogical signal – transmitted via ADSL on the local copper loop – to a numerical signal that can be transmitted to the higher levels of the network. The upgrading of local LEs is the key source of variation that we use in our empirical analysis.

ADSL roll-out in France As evidenced by Malgouyres et al. (2019), the deployment of broadband internet technology beyond France's largest cities was slow at the beginning of the 2000's (see Table C1). The authors show that there were multiple reasons for this staggered deployment. First, France Télécom, the monopolistic telecom supplier, was uncertain regarding the future wholesale price it was going to be able to charge, mainly due to regulatory reasons. Second, at the same time that France Télécom had to invest massively in upgrading its LEs to ADSL, it went through a debt crisis that ended with what was essentially a government bailout in 2002. Urged on by the government – which increased its stake in the firm during the 2002 bailout of the firm – in 2003 France Télécom pledged to cover 90% of the French (mainland) population by the end of 2005, i.e. all LEs with more than 1,000 lines.

Between 2004 and 2007, local governments were involved in broadband internet deployment by subsidising the expansion and favouring competition among providers. Most relevant for broadband expansion was the creation of a contract between local governments, the Plan Département Innovant, whereby France Télécom pledged to

TABLE C1. Year of connection by municipality

Year of Connection	Number connected (in % of total)			
	Cities	Workers	Establishments	Population
2000	2.1	25.0	22.2	18.5
2001	6.6	35.7	34.7	28.9
2002	8.4	19.3	19.6	18.4
2003	12.4	6.7	7.8	9.5
2004	18.4	5.0	5.8	8.4
2005	23.0	4.4	5.4	8.5
2006	18.6	2.2	2.8	5.3
2007	8.8	1.6	1.7	2.5

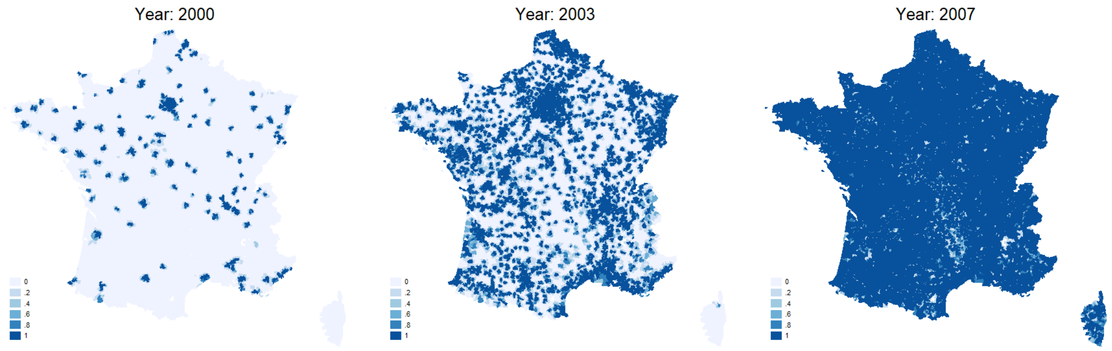
Notes: All values are taken in 1999. The sum of percentages in a column is different from 100 because a small number of cities are not connected to the ADSL in 2007. The number of establishments and workers is based on our final sample (therefore following our cleaning and selection procedures).

equip all LEs in a département with more than 100 connections within a year. The proclaimed target of the plan was to raise coverage to 96% of the French population by the end of 2005 and activate all the remaining LEs by the end of 2006. We account for the role of local government in our empirical analysis by including département-year fixed effects. Overall, the account of the broadband expansion in France over the period suggests that it was gradual due to uncertainty regarding the capacity of France Télécom to undergo the investment until 2002. After 2002, with strong encouragement from the government, France Télécom started covering more secondary areas with a focus on the overall number of lines per LE, with only limited attention paid to local economic potential. Although the coverage was accelerated, it remained gradual due to France Télécom's operational limits and took about two years longer than anticipated in 2003. Because our main effects of interest are identified out of the gradual diffusion of the new technology in different LEs over space and time, addressing the endogeneity of the decision to "treat" one LE before another deserves special consideration. Malgouyres et al. (2019) show that broadband expansion occurred to maximise population coverage with no special consideration for economic potential, a fact that is strongly supported by the statistical analysis of the determinants of broadband coverage that they carried out.

Use of broadband technologies by firms ADSL technology, while progressively replaced by other technologies – notably direct access to the optic fibre or FTTO (fibre to the office) –, is the main way in which firms access the internet. A 2016 survey showed that in that year 73% of SMEs used ADSL technology (Arcep, 2016). The large take-

FIGURE C1. The progressive roll-out of the DSL technology in France—$\widetilde{Z}$

(a) 2000 (b) 2003 (c) 2007



Notes: This figure presents the geographical distribution of the continuous measure of local broadband availability (variable $\widetilde{Z}$).

up reflects the fact that ADSL was a massive improvement in terms of speed (from 56 to 512kbit/s for a transition from a classical to first generation ADSL connection) as well as in terms of connection cost and time. While there is no administrative data on firm-level use of broadband, based on repeated survey data, firms located in cities that received broadband internet earlier experienced higher growth in the proportion of employees that used internet on a regular basis between 1999 and 2004. This statistical association cannot be interpreted causally under the same set of assumptions as our main analysis. It is however strongly suggestive of an impact from broadband availability on broadband adoption.

D BI Expansion and Skill-Biased Technological Change

In this Appendix, we confirm and extend the results of Akerman et al. (2015) showing that broadband internet constitutes a skill-biased technology. In particular, we show that when a city is connected to BI (i) the labor productivity of establishments located in the city increases, (ii) the demand for high-skill workers increases, and (iii) the hourly wage and salary of high-skill workers increase.

D.1 At the city and establishment level

Our identification for the city level analysis follows a stacked difference-in-differences strategy as reported in Equation 11. For the establishment level analysis we follow the same model, but we include establishment fixed effects instead of city fixed effects. The results of the latter can be interpreted as the pure within-firm effect that excludes any changes due to composition. We start by evaluating the impact that BI and the underlying ADSL technology had on firm productivity. We measure labor productivity as the log of value added divided by the total wage bill. Given that the financial data is only available at the company level, we assign productivity to all the establishments of multi-plant firms according to one measured at the overall firm level. In a robustness check, we show the results obtained on a sample that excludes multi-establishment firms. At the city level, we consider the average productivity obtained across the establishments located in the municipality weighted by their size.[35] Secondly - to capture skill-biased technological change - we look at the impact of BI on the share of high-skill workers within cities and establishments.[36]

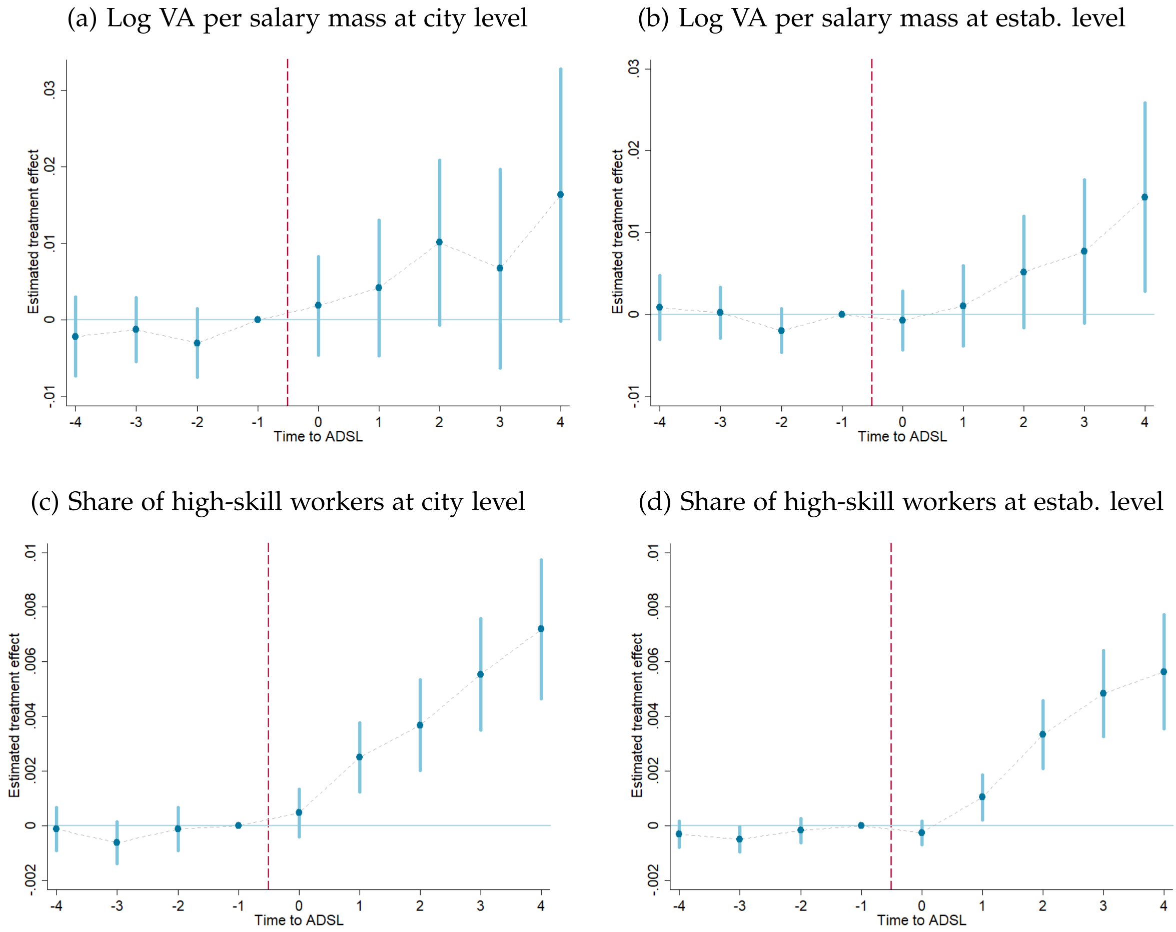
Results obtained from both city and establishment-level regressions are reported in Figure D1 and the corresponding coefficients are given in Table D1. These findings confirm what was expected: the productivity of firms increases when the city in which they are located is connected to BI. The average labor productivity of firms located in the city increases by about 1% over the first five years, and about half of this effect (0.6%) takes place in firms already present in the area before the shock.[37]

[35] See Table I for summary statistics for the main outcome variables in the city level and establishment level samples.

[36] All our measures of employment are expressed in terms of full-time equivalents. High-skill workers are defined based on their occupation, and include executive positions, managers and engineers, which correspond to the highest socio-professional category.

[37] The positive effect of BI on labor productivity is not purely driven by an increase in the skill intensity of the firms located in the city, but goes beyond that. First, by dividing the value added of the firm

FIGURE D1. Firm productivity and broadband access



Notes: This Figure shows regression coefficients and 95% confidence intervals from a dynamic event study where the dependent variable is the log of value added per salary mass within a city or establishment (Panel a and b) or the share of executive workers within a city of establishment (Panel c and d) at t and the specification follows equation 11.

Similarly, before the arrival of BI, the share of high skill workers evolved comparably across cities belonging to different cohorts of ADSL diffusion, conditional on department-specific time trends and the other controls. When cities get access to BI, they experience a general upskilling of their labor force relative to other municipalities. In particular, the share of full-time employment accounted for by the top socio-professional category increases, which is in line with the thesis of skill biased technological change. In terms of magnitude, the share of high-skill workers in a city increases by 0.4 percentage points following the diffusion of BI. This effect can be compared with the baseline average observed in cities at the beginning of the period,

by the wage bill, instead of the firm size, we partially account for the fact that high-skill workers are paid more. Second, if we include the share of high-skill workers as an additional control in the productivity regressions, the coefficients remain widely unchanged (results available upon request).

which was 5.8%: the share of high-skill workers thus increases by about 6.9% after the arrival of BI with respect to the baseline.

TABLE D1. Effect of ADSL on productivity and demand for high-skill workers

	(1)	(2)	(3)	(4)
	Sh. of high skill workers		Log VA / salary mass	
VARIABLES	City level	Estab. Level	City level	Estab. Level
T = 0	0.000653	-2.79e-05	0.00322	-0.000344
	(0.000477)	(0.000240)	(0.00330)	(0.00157)
T = +1	0.00269***	0.00125***	0.00563	0.00152
	(0.000663)	(0.000437)	(0.00452)	(0.00239)
T = +2	0.00386***	0.00352***	0.0115**	0.00565
	(0.000891)	(0.000653)	(0.00561)	(0.00348)
T = +3	0.00573***	0.00504***	0.00806	0.00815*
	(0.00109)	(0.000826)	(0.00669)	(0.00439)
T = +4	0.00738***	0.00583***	0.0177**	0.0148**
	(0.00134)	(0.00108)	(0.00861)	(0.00585)
Average effect	0.00406***	0.00312***	0.00924*	0.00596*
	(0.000822)	(0.000606)	(0.00538)	(0.00330)
Observations	423,770	3,075,954	416,052	2,911,303
R-squared	0.711	0.894	0.622	0.711

Notes: Columns (1) and (3) run the regression at the city level, following equation 11, where controls are the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level. Columns (2) to (4) run the same specification on the outcome computed at the establishment level, replacing city fixed effects by establishment fixed effects.

Such results could arise for two reasons: either because BI fosters the entry of new establishments with a higher average skill level than the incumbents, or because the average establishment already present in the city increases its share of high-skill workers. To capture the extent to which composition effects play a role, we compare our city-level results with similar event studies at the establishment level, which only keep the plants already present in the city before the arrival of ADSL in the sample. The effect on high skill workers within existing establishments is qualitatively similar to the one at the city level. This suggests that the increase in share of skilled workers is not (only) driven by a composition effect but is also a phenomenon taking place within existing firms. The magnitude is however slightly smaller: BI increases the share of high skill workers within existing firms by 0.3 percentage points compared to a baseline average of 10% (3% growth with respect to baseline). Tables D3 and D4 show the robustness of the results to including a control for pre-BI productivity growth at the city level interacted with year dummies, to running a standard dynamic difference-in-differences model as reported in Equation 10, and to keeping only mono-establishment firms in the productivity analysis. Finally, Table D5 shows the static coefficients obtained from a stacked difference-in-differences on the post-BI period.

D.2 At the individual level

In this subsection, we show that the evidence of increasing demand for skilled workers translates into increased wages, for our individual panel. As explained in Section 3, our data allow us to follow part of the workers over time. More precisely, we can follow every worker born in October of an even year (roughly 1/24 of the population) between 1994 and 2010. With these data, we can look at the individual wage effect of BI expansion, i.e., we can consider the change in hourly wage that follows the connection of a worker's city to ADSL. We therefore estimate the following model:

$$\log(w_{i,t}) = \beta \tilde{Z}_{c(i),t} + X\gamma + \psi_{k,t} + \nu_i + \zeta_{s(i)} + \varepsilon_{i,t}, \tag{14}$$

where $w_{i,t}$ is the hourly wage of individual i over year t on average. $\tilde{Z}_{c(i),t}$ is the variable that captures the share of the city $c(i)$, where individual i works, that is connected to BI. To some reasonable extent, $\tilde{Z}$ can be seen as a dummy variable indicating whether the city has been connected to BI prior to year t. X is a vector of time-varying individual characteristics: age, age squared, an indicator of whether the job is part-time (as opposed to full-time) and gender. Finally, $\psi_{k,t}$, ν_i, $\zeta_{s(i)}$ are a set of labor market area k times year t fixed effects, individual fixed effects and sector $s(i)$ fixed effects. ε is an idiosyncratic error that we assume can be correlated within labor market areas but not across. Finally, β captures the effect (in percentage points) of being connected to BI on wage, controlling for observable and time-varying unobservable worker characteristics.

Table D2 presents our results and Table D6 presents the summary statistics of the variables used for the regression. Column (1) includes all workers (around 11 millions) and shows that the coefficient of the dummy variable $C_{c(i),t}$ (first line, labeled "connected") is positive and significant. Its magnitude suggests that the hourly wage permanently increases by 3% on average for all workers once connected to BI. In this specification, we did not include individual fixed effects ν_i but control for initial wage

TABLE D2. Effect of ADSL on individual wage

Sample	(1) All workers	(2)	(3) 3 skills	(4)	(5) 2 skills	(6)
Connected	0.030***	0.006***	0.013***	-0.016***	0.010*	-0.014***
	(0.003)	(0.001)	(0.004)	(0.003)	(0.005)	(0.003)
× High-Skilled			0.042***	0.116***	0.052***	0.116***
			(0.013)	(0.007)	(0.015)	(0.009)
× Int-Skilled			0.004	0.025***		
			(0.004)	(0.003)		
Age	0.032***	0.044***	0.026***	0.041***	0.020***	0.035***
	(0.003)	(0.006)	(0.002)	(0.005)	(0.001)	(0.003)
Age Sq.	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender	0.106***		0.123***		0.125***	
	(0.004)		(0.007)		(0.002)	
Short Time	-0.043***	0.035***	-0.020***	0.035***	0.001	0.059***
	(0.009)	(0.001)	(0.004)	(0.001)	(0.005)	(0.002)
High Skill			0.649***	0.210***	0.644***	0.337**
			(0.005)	(0.005)	(0.008)	(0.007)
Int. skill			0.168***	0.039***		
			(0.003)	(0.002)		
Initial wage (log)	0.346***		0.334***		0.249***	
	(0.025)		(0.016)		(0.020)	
<u>Fixed Effects</u>						
LMA × year	✓	✓	✓	✓	✓	✓
Sector	✓	✓	✓	✓	✓	✓
Individual		✓		✓		✓
Obs.	7,810,286	7,808,176	7,810,286	7,808,176	4,316,357	4,256,281
R Sq.	0.46	0.78	0.62	0.79	0.70	0.85

Notes: This Table shows regression results from an estimation of equation (14). Variable description is given in Table D6 of the Online Appendix A. All workers are included in the regressions, except in column (5) and (6) in which we drop intermediate skill workers. All regressions include a labor market area (*zone d'emploi*) times year fixed effect as well as a sector fixed effect at the 2 digit level. Columns (2), (4) and (6) also include an individual fixed effect. Heteroskedasticity robust standard errors clustered at the labor market area level under parenthesis.

to capture the level of skill of the worker.[38] Including an individual fixed effect would better control for unobserved worker heterogeneity (which includes education) and this is presented in column (2). Our coefficient of interest remains positive and significant but somehow lower (0.6%). Columns (3) and (4) produce the same type of regression as Akerman et al. (2015) where we interact $C_{c(i),t}$ with a dummy variable for each skill level. In line with their results, we see that the effect of BI on wages is significantly larger for high skill-workers than for others. Columns (5) and (6) confirm these results by restricting our analysis to only low and high-skill workers (i.e. excluding intermediate skill workers from the sample).

[38] Initial wage is defined as the logarithm of wage per hour taken in the first year in which the worker appears in the panel, this year is then removed from the regression.

Overall, these results confirm what we reported at the city level: BI is associated with a larger demand for high-skill workers and this translates into higher wages, even when controlling for unobserved heterogeneity and the usual controls. These results also show that the increasing demand for high-skill workers observed at the city and establishment level is not a pure composition effect as, overall, the arrival of BI benefits this class of workers more.

D.3 Additional Tables

TABLE D3. Robustness of effect of ADSL on demand for high-skill workers

	(1)	(2)	(3)	(4)
	Sh. of high skill workers			
VARIABLES	City level		Estab. Level	
T = 0	0.000650	0.000431	1.91e-05	-0.000147
	(0.000464)	(0.000413)	(0.000254)	(0.000253)
T = +1	0.00240***	0.00218***	0.00130***	0.000972**
	(0.000662)	(0.000615)	(0.000472)	(0.000446)
T = +2	0.00367***	0.00316***	0.00360***	0.00298***
	(0.000926)	(0.000844)	(0.000706)	(0.000656)
T = +3	0.00548***	0.00505***	0.00523***	0.00454***
	(0.00116)	(0.00108)	(0.000906)	(0.000852)
T = +4	0.00649***	0.00632***	0.00589***	0.00560***
	(0.00141)	(0.00131)	(0.00117)	(0.00114)
Control: pre-BI prod. growth	✓		✓	
Staggered DiD strategy		✓		✓
Average effect	0.00374***	0.00343***	0.00321***	0.00279***
	(0.000860)	(0.000788)	(0.000657)	(0.000624)
Observations	316,407	140,751	2,935,705	1,606,571
R-squared	0.721	0.731	0.895	0.894

Notes: Columns (1) and (3) run the regression at the city and establishment level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, the productivity growth at the city level pre-BI (1996-1998) interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects. Columns (2) and (4) run the regression run the regression following the standard staggered difference-in-differences strategy reported in equation 10, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, and city (establishment) fixed effects.

TABLE D4. Robustness of effect of ADSL on productivity

	(1)	(2)	(3)	(4)	(5)	(6)
			Log VA / salary mass			
VARIABLES	City level			Estab. Level		
T = 0	0.00653**	0.00243	0.00444	0.000311	0.00126	0.00109
	(0.00306)	(0.00302)	(0.00290)	(0.00154)	(0.00164)	(0.00180)
T = +1	0.00987**	0.00593	0.00515	0.00182	0.00335	0.00214
	(0.00441)	(0.00417)	(0.00436)	(0.00236)	(0.00241)	(0.00274)
T = +2	0.0143**	0.0101**	0.00834	0.00553	0.00753**	0.00625
	(0.00581)	(0.00509)	(0.00521)	(0.00345)	(0.00320)	(0.00398)
T = +3	0.0133*	0.00783	0.00783	0.00897**	0.0105***	0.0111**
	(0.00677)	(0.00600)	(0.00611)	(0.00429)	(0.00395)	(0.00545)
T = +4	0.0243***	0.0167**	0.0139*	0.0166***	0.0184***	0.0220***
	(0.00896)	(0.00744)	(0.00722)	(0.00596)	(0.00485)	(0.00677)
Control: pre-BI prod. growth	✓			✓		
Staggered DiD strategy		✓			✓	
Mono-establishments only			✓			✓
Average effect	0.0137**	0.00860*	0.00795*	0.00664**	0.00820***	0.00851**
	(0.00543)	(0.00476)	(0.00472)	(0.00324)	(0.00302)	(0.00387)
Observations	315,106	138,689	390,877	2,781,076	1,513,306	2,191,875
R-squared	0.620	0.605	0.580	0.710	0.686	0.693

Notes: Columns (1) and (4) run the regression at the city and establishment level, following equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, the productivity growth at the city level pre-BI (1996-1998) interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects. Columns (2) and (5) run the regression run the regression following the standard staggered difference-in-differences strategy reported in equation 10, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, and city (establishment) fixed effects. Columns (3) and (6) run the regression at the city and establishment level restricting the sample to mono-establishment firms. They follow equation 11, where the dependent variables are the weighted average of the firm outcomes at the city level, and controls are the population density in 1999 interacted with year dummies, department x year fixed effects, city (establishment) fixed effects and sample fixed effects.

TABLE D5. Effect of ADSL on skill-biased technical change - static regressions

	Log VA / salary mass	Sh. of high skill workers
Panel A : city level regressions		
Post ADSL * treated	0.00550	0.00204***
	(0.00393)	(0.000571)
Observations	416,052	423,770
R-squared	0.622	0.711
Panel B : establishment level regressions		
Post ADSL * treated	0.000925	0.000942***
	(0.00188)	(0.000334)
Observations	2,911,303	3,075,954
R-squared	0.711	0.894

Notes: The regressions are run at the city and establishment level following a model similar to equation 11, but where instead of including the dynamic post-ADSL effects for every year, we just include a dummy for post-ADSL period interacted with the treatment indicator. All columns control for the population density in 1999 interacted with year dummies, department x year fixed effects, city fixed effects and sample fixed effects.

TABLE D6. Variable description for Table D2

Variable	Description	Mean	p25	p75
Log of wage	log of hourly wage (dependent variable)	2.41	2.10	2.63
Age	Age of the worker	37	28	46
Age Sq.	Age $\times$ Age	1,507	784	2,116
Gender	Gender of the worker	0.63	0	1
Short Time	Dummy for declaring working part time	0.17	0	1
High Skill	Dummy for working in a high skill occupation	0.13	0	1
Int. Skill	Dummy for neither working in high or low skill occupation	0.45	0	1
Initial Wage (log)	Log of hourly wage taken in the first year the worker appear in the data	2.22	1.93	2.40

Notes: Variable description used in the panel data wage regression and basic descriptive statistics.

E Theory Appendix

Note: we drop the index j when the context does not command it.

E.1 Proof of Proposition 1

Proof. First, the CES structure of the production function for a given occupation yields the following elasticities:

$$\begin{aligned} \frac{\partial H_i}{\partial n_i} &= H_i^{1/\sigma_i}\mu_i^{1/\sigma_i}n_i^{-1/\sigma_i} \Longrightarrow \frac{\partial H_i}{\partial n_i}\frac{n_i}{H_i} = H_i^{\frac{1-\sigma_i}{\sigma_i}}\mu_i^{\frac{1}{\sigma_i}}n_i^{\frac{\sigma_i-1}{\sigma_i}} \\ \frac{\partial H_{i,j}}{\partial s_{i,j}} &= H_{i,j}^{1/\sigma_i}(1-\mu_i)^{1/\sigma_i}s_{i,j}^{-1/\sigma_i} \Longrightarrow \frac{\partial H_i}{\partial s_i}\frac{s_i}{H_i} = H_i^{\frac{1-\sigma_i}{\sigma_i}}(1-\mu_i)^{\frac{1}{\sigma_i}}s_i^{\frac{\sigma_i-1}{\sigma_t i}} \end{aligned}$$

Note also that this elasticity:

$$e_i \equiv \frac{\partial H_i}{\partial s_i}\frac{s_i}{H_i} = \frac{(1-\mu_i)^{\frac{1}{\sigma_i}}s_i^{\frac{\sigma_i-1}{\sigma_i}}}{(1-\mu_i)^{\frac{1}{\sigma_i}}s_i^{\frac{\sigma_i-1}{\sigma_i}} + \mu_i^{\frac{1}{\sigma_i}}n_i^{\frac{\sigma_i-1}{\sigma_i}}} \in [0,1]$$

and $\frac{\partial H_i}{\partial n_i}\frac{n_i}{H_i} = 1 - e_i$.

Second, the first-order conditions can be combined to give a relationship between s_i and n_i:

$$s_i = \frac{1-\mu_i}{\mu_i}\left[\frac{\rho+1}{a_i^{\rho}r_i\gamma_i}\right]^{\sigma_i} n_i^{\rho\sigma_i+1} = \lambda_i n_i^{\rho\sigma_i+1},$$

and the cost share of outsourced workers is thus given by:

$$\eta_i^c \equiv \frac{\gamma_i r_i s_i}{\gamma_{i,j} r_i s_i + w(n_i)n_i} = 1 - \frac{1}{1+\gamma_i r_i a^{\rho}\lambda_i n_i^{\rho(\sigma_i-1)}}.$$

As long as $\sigma_i > 1$ and $\rho > 0$, we therefore have:

$$\frac{\partial \eta_i^c}{\partial n_i} > 0.$$

Using the relationship between n_i and s_i and log differentiating H_i, it is straightforward to show that

$$d\log\left(H_i\right) = d\log n_i\left(1 + H_i^{1/\sigma_i - 1}(1-\mu_i)^{1/\sigma_i} s_i^{1-1/\sigma_i}\rho\sigma_i\right) = d\log n_i\left(1 + e_i\rho\sigma_i\right)$$

Next, log-differentiating PY:

$$d\log\theta\frac{\varepsilon-1}{\varepsilon} + \left(\sum_{i'}\alpha_{i'}d\log(H_{i'})\right)\frac{\varepsilon-1}{\varepsilon} = (1/\sigma_i + \rho)\, d\log n_i + d\log H_i,$$

which can be rewritten as:

$$d\log\theta + \left(\sum_{i'}\alpha_{i'}d\log n_{i'}(1 + e_{i'}\rho\sigma_{i'})\right) = \frac{\varepsilon}{\varepsilon-1}\left(1/\sigma_i + \rho + 1 + e_i\rho\sigma_i\right) d\log n_i. \qquad (15)$$

This expression is valid for all i which shows that $d\log n_i$ are either all positive or all negative as $d\log\theta > 0$. To show that they are all positive, we first multiply the above equation by α_i and then sum for all i:

$$d\log\theta = \frac{1}{\varepsilon-1}\left(\sum_{i'}\alpha_{i'}d\log n_{i'}(1 + e_{i'}\rho\sigma_{i'} + \varepsilon(1/\sigma_{i'} + \rho))\right) > 0.$$

This implies that $\frac{d\log(n_i)}{d\log\theta} > 0$ and then $\frac{d\eta_i^c}{d\theta} > 0$. □

E.2 Proof of Proposition 2

Proof. Starting from equation (15) and using the fact that $s_i = \lambda_i n_i^{\rho\sigma_i+1}$, we know that:

$$\frac{1}{(\rho\sigma_i+1)}\frac{d\log(s_i)}{d\log(\theta)}\left(\frac{1}{\sigma_i} + 1 + \rho + \rho\sigma_i e_i\right),$$

is independent of i. Hence, a sufficient condition to have $\frac{d\log(s_1)}{d\log(\theta)} < \frac{d\log(s_2)}{d\log(\theta)}$ is:

$$(\rho\sigma_2+1)\left(\frac{1}{\sigma_1} + 1 + \rho + \rho\sigma_1 e_1\right) = (1+\rho\sigma_2)(2+\rho+\rho e_1) > (\rho+1)\left(\frac{1}{\sigma_2} + 1 + \rho + \rho\sigma_2 e_2\right)$$

Because $e_1 \in (0,1)$, then a larger sufficient condition is:

$$(1+\rho\sigma_2)(2+\rho) > (1+\rho)\rho\sigma_2 + (1+\rho)(1 + \frac{1}{\sigma_2} + \rho)$$

which is true as long as $\sigma_2 > 1 + \rho$.

Similarly, equation (15) can be used to show that:

$$\frac{d\log(n_i)}{d\log(\theta)}\left(\frac{1}{\sigma_i}+1+\rho+\rho\sigma_i e_i\right),$$

is independent of i. This shows that as long as:

$$\sigma_2 e_2 > \frac{1+\rho}{\rho}, \text{ then } \frac{d\log(n_2)}{d\log(\theta)} < \frac{d\log(n_1)}{d\log(\theta)}.$$

This show that:

$$\frac{d\log(s_1/n_1)}{d\log(\theta)} < \frac{d\log(s_2/n_2)}{d\log(\theta)},$$

and thus:

$$\frac{d\log(r_1\gamma_1 s_1/(n_1 w(n_1))}{d\log(\theta)} < \frac{dlog(r_2\gamma_2 s_2/(n_2 w(n_2))}{d\log(\theta)}.$$

Then following an increase in θ, the non-core occupation will experience a relative increase in the share of its labor cost coming from outsourced workers that is larger than what the core occupation experiences. In fact, because we have assumed that $\sigma_1 = 1$, the core occupation do not experience any change in its cost share of outsourced workers which concludes the proof.

□

E.3 Proof of Proposition 3

Proof. The first order conditions can be combined to show that:

$$\frac{n_i w_i}{PY} = \alpha_i \frac{\varepsilon-1}{\varepsilon}\frac{1-e_i}{1+\rho} \text{ and } \frac{r_i s_i}{PY} = \alpha_i \frac{\varepsilon-1}{\varepsilon} e_i$$

so that the revenue share of occupation i is given by:

$$\frac{n_i w_i + r_i s_i}{PY} = \alpha_i \frac{\varepsilon-1}{\varepsilon}\frac{1+\rho e_i}{1+\rho} \in \left[\alpha_i \frac{\varepsilon-1}{\varepsilon(1+\rho)}; \alpha_i \frac{\varepsilon-1}{\varepsilon}\right]$$

This shows that as long as $\alpha_{i+1} < \frac{\alpha_i}{1+\rho}$, the revenue share increases as σ_i decreases. Note that with two occupations 1 and 2, this is true as long as:

$$\rho < \frac{2\alpha_1 - 1}{1-\alpha_1}$$

Because $\lambda_1 < \lambda_2 < \ldots < \lambda_N$, then:[39]

$$n_i w_i + r_i s_i > n_{i+1} w_{i+1} + r_{i+1} s_{i+1} \Longrightarrow n_i > n_{i+1}$$

Adding to the fact that $\frac{d\log(n_1)}{d\log(\theta)} > \frac{d\log(n_2)}{d\log(\theta)}$, this shows that the larger occupation in terms of in-house workers ($n_1 > n_2$) is also the one that will increase the most its number of in-house workers, which results in an increase in the HHI index. □

E.4 The case of a reduction in the cost of outsourcing

In this extension, we consider the case of a reduction in the value of $\gamma_{i,j}$ for a firm j. We assume that the relative decrease is the same for all occupations, i.e. that $d\log(\gamma_{i,j}) = d\log(\gamma)$. As usual, we drop the subscript j for the sake of clarity.

We show that under a large set of assumptions, firms respond to a reduction of the cost of outsourcing γ by increasing their outsourcing intensity which results in an increasing level of concentration of occupation in the firm.

To show this, first note that as long as $\sigma_i > 0$:

$$\frac{d\log(\eta_i^c)}{d\log(\gamma)} < 0 \Longleftrightarrow \frac{d\log(n_i)}{d\log(\gamma)} < 1/\rho$$

The combination of the two first order conditions continue to give the same relationship between n_i and s_i, only this time:

$$d\log(s_i) = -\sigma_i d\log(\gamma) + (\rho\sigma_i + 1) d\log(n_i) \tag{16}$$

Lemma 1. *At least one type of occupation must have $d\log(s_i)/d\log(\gamma) < 0$*

Proof. The full differentiation of $d\log(H_i)$ gives:

$$d\log(H_i) = e_i d\log(s_i) + (1-e_i) d\log(n_i) = d\log(s_i)\frac{1+\rho\sigma_i e_i}{1+\rho\sigma_i} + \frac{(1-e_i)\sigma_i}{1+\rho\sigma_i} d\log(\gamma)$$

Hence, differentiating the first order condition with respect to s_i and summing over

[39] This is because $n_{i+1} w_{i+1} + r_{i+1} s_{i+1} = a^{-\rho} n_{i+1}^{\rho+1} + \lambda_{i+1} n_{i+1}^{\rho\sigma_{i+1}+1} > a^{-\rho} n_{i+1}^{\rho+1} + \lambda_i n_{i+1}^{\rho\sigma_i+1}$

all i after having pre-multiplied by α_i

$$\frac{\varepsilon-1}{\varepsilon}\sum_{j\in\mathbf{N}}\alpha_j d\log(H_j) = d\log(H_i) + d\log(\gamma) + \frac{1}{\sigma_i}d\log(s_i),$$

becomes:

$$-\sum_{i\in\mathbf{N}}\alpha_i\frac{d\log(s_i)}{d\log(\gamma)}\left[\frac{1}{\varepsilon}\frac{1+\rho e_i\sigma_i}{1+\rho\sigma_i}+\frac{1}{\sigma_i}\right] = 1+\frac{1}{\varepsilon}\sum_{i\in\mathbf{N}}\alpha_i\frac{1-e_i}{1+\rho\sigma_i}\sigma_i > 0.$$

Which shows that at least one $\frac{d\log(s_i)}{d\log(\gamma)}$ must be smaller than 0.

□

Coming back to the two type of occupation case where $\mathbf{N} = \{1, 2\}$ and $\sigma_1 = 1$, we know that η_1^c is constant and η_2^c will increase following a drop in γ if $d\log(n_2)/d\log(\gamma) < 1/\rho$. Let's assume that this is not the case, i.e. that $d\log(n_2)/d\log(\gamma) \geq 1/\rho > 0$.

Then $d\log(s_2)/d\log(\gamma) > 1/\rho$ from equation (16). And from the previous lemma, we know that $d\log(s_1)/d\log(\gamma) < 0$.

Using again equation (16), we also have

$$\frac{d\log(n_1)}{d\log(\gamma)} < \frac{\sigma_1}{\rho\sigma_1+1} \leq \frac{1}{\rho},$$

and finally:

$$\frac{d\log(H_1)}{d\log(\gamma)} < \frac{1-e_1}{\rho} < \frac{1}{\rho} \text{ while } \frac{d\log(H_2)}{d\log(\gamma)} > \frac{1}{\rho}$$

Using the differentiated first order condition with respect to the second occupation yields:

$$\frac{\varepsilon-1}{\varepsilon}\left(\alpha_1 d\log(H_1) + (1-\alpha_1)d\log(H_2)\right) = d\log(H_2) + \frac{1}{\sigma_2}d\log(s_2) + d\log(\gamma)$$

whence:

$$\begin{aligned}\frac{\varepsilon-1}{\varepsilon}\left(\frac{\alpha_1}{\rho}+\frac{(1-\alpha_1)d\log(H_2)}{d\log(\gamma)}\right) &> \frac{\varepsilon-1}{\varepsilon}\left(\alpha_1 d\log(H_1)+(1-\alpha_1)d\log(H_2)\right)\\ &\geq \frac{d\log(H_2)}{d\log(\gamma)}+1+\frac{d\log(s_2)}{d\log(\gamma)}\\ \implies \left(\frac{\varepsilon-1}{\varepsilon}(1-\alpha_1)-1\right)\frac{d\log(H_2)}{d\log(\gamma)} &> 1-\frac{\varepsilon-1}{\varepsilon}\frac{\alpha_1}{\rho}+\frac{1}{\sigma_2}\frac{d\log(s_2)}{d\log(\gamma)}\end{aligned}$$

The left-hand side of this last inequality is negative and the right hand side is larger than:

$$\frac{1}{\sigma_2 \rho} + 1 - \frac{\varepsilon - 1}{\varepsilon} \frac{\alpha_1}{\rho},$$

which is positive as long as α_1 is not too large or ρ is not too small. This leads to an impossible statement and hence contradict the assumption that $d\log(n_2)/d\log(\gamma) \geq 1/\rho$.

E.5 Extension to more than 2 occupations

The model can be extended to a fixed number of occupation $N = |\mathbf{N}|$. The proof of Proposition 1 already considers the general case. The proof from Proposition 2 can be extended in the following way.

We know that for every occupation i:

$$\frac{1}{(\rho\sigma_i + 1)} \frac{d\log(s_i)}{d\log(\theta)} \left(\frac{1}{\sigma_i} + 1 + \rho + \rho\sigma_i e_i \right),$$

is independent of i which means that $\frac{d\log(s_i)}{d\log(\theta)} < \frac{d\log(s_{i+1})}{d\log(\theta)}$ is true as long as:

$$(\rho\sigma_{i+1} + 1) \left(\frac{1}{\sigma_i} + 1 + \rho + \rho\sigma_i e_i \right) > (\rho\sigma_i + 1) \left(\frac{1}{\sigma_{i+1}} + 1 + \rho + \rho\sigma_{i+1} e_{i+1} \right)$$

a sufficient condition is that:

$$\frac{1}{1 + \rho\sigma_{i+1}} > 1 - \left(\frac{1}{\sigma_i} - \frac{1}{\sigma_{i+1}} \right)$$

This is true as long as ρ is sufficiently close to 0.

Similarly:

$$\frac{d\log(n_i)}{d\log(\theta)} \left(\frac{1}{\sigma_i} + 1 + \rho\sigma_i e_i \right),$$

is independent of i which means that $\frac{d\log(n_{i+1})}{d\log(\theta)} < \frac{d\log(n_i)}{d\log(\theta)}$ is true as long as:

$$\sigma_{i+1} e_{i+1} > \frac{1}{\rho\sigma_i} + \sigma_i$$

This in particular implies that the σ_i are not too close with each other, i.e. that the different occupations are characterised by very different level of substituability. In such a case, Proposition 3 also applies and the concentration of i-house occupation following an increase in θ becomes larger.

Note that these are sufficient conditions that can be derived without solving numerically the model for $N > 2$. In Appendix E.7, we present some simulation for $N = 4$ and show that the concentration of in-house workers indeed increase following a shock in θ.

E.6 Markups and outsourcing share

The model gives a clear prediction on the sign of the relationship of the inverse in-house labor share and the outsourcing intensity as well as the relationship of sales over total labor cost and outsourcing intensity. To see this formally, we start from the first order conditions which can be combined to show that:

$$\frac{n_i w_i}{PY} = \alpha_i \frac{\varepsilon - 1}{\varepsilon} \frac{1 - e_i}{1 + \rho} \text{ and } \frac{r_i s_i}{PY} = \alpha_i \frac{\varepsilon - 1}{\varepsilon} e_i,$$

where we recall that $e_i = \partial \log(H_i) / \partial \log(s_i)$. The revenue share of occupation i is given by:

$$\frac{n_i w_i + r_i s_i}{PY} = \alpha_i \frac{\varepsilon - 1}{\varepsilon} \frac{1 + \rho e_i}{1 + \rho}$$

We define the markup $\mathcal{M}$ as the ratio of PY over total cost:

$$\mathcal{M} = \frac{PY}{\sum_{i \in \mathbf{N}} s_i \gamma_i r_i + w(n_i) n_i} = \frac{\varepsilon}{\varepsilon - 1} \frac{1 + \rho}{1 + \rho \sum_{i \in \mathbf{N}} \alpha_i e_i}$$

Let us now define $\mathcal{O}$ as the share of total labor cost coming from outsourced workers:

$$\mathcal{O} = \frac{\sum_{i \in \mathbf{N}} s_i \gamma_i r_i}{\sum_{i \in \mathbf{N}} s_i \gamma_i r_i + w(n_i) n_i} = \frac{(1 + \rho) \sum_{i \in \mathbf{N}} \alpha_i e_i}{1 + \rho \sum_{i \in \mathbf{N}} \alpha_i e_i}$$

This implies in particular that:

$$\sum_{i \in \mathbf{N}} \alpha_i e_i = \frac{\mathcal{O}}{1 + \rho - \rho \mathcal{O}}$$

And finally:

$$\mathcal{M} = \frac{\varepsilon(1 + \rho)}{\varepsilon - 1} (1 - \frac{\rho}{1 + \rho} \mathcal{O})$$

Hence the model predicts a *negative* relationship between the ratio of revenue over total labor cost and outsourcing intensity.

Taking back this last equation and multiplying by $1/(1-\mathcal{O})$ yields:

$$\frac{PY}{\sum_{i\in\mathbf{N}} w_i n_i} = \frac{\varepsilon}{\varepsilon-1}\left[\frac{1}{1-\mathcal{O}} + \rho\right]$$

So that the model predicts a *positive* relationship between the inverse in-house labor share and outsourcing intensity.

These relationships are tested empirically in Table A1.

E.7 Numerical examples

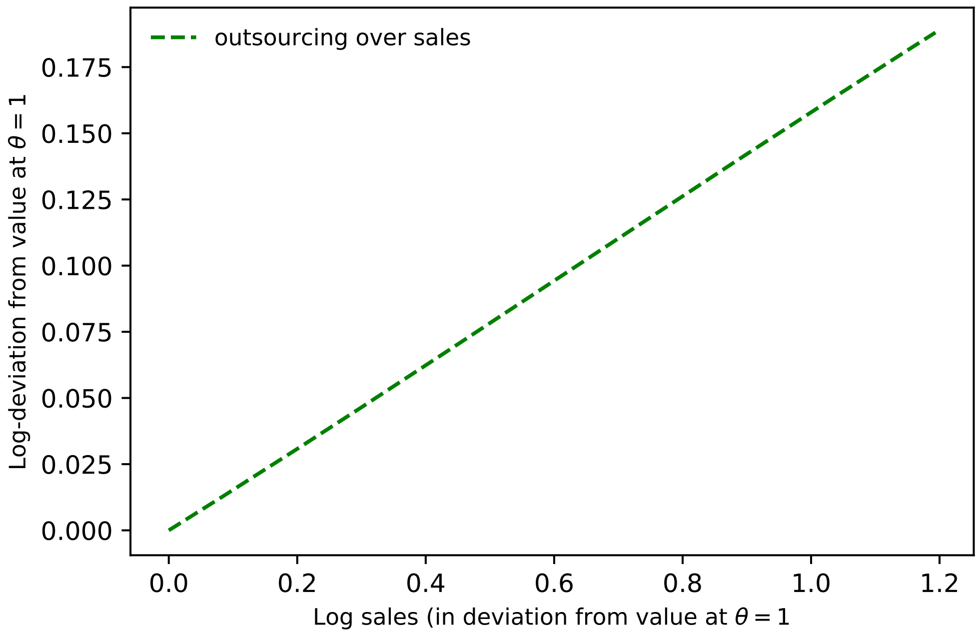
Comparative statics: increase in productivity θ. We consider a specific case with a firm with 4 occupations: 2 high skill ($\alpha_i = 1/3$) and 2 low skill ($\alpha_i = 1/6$), 2 core ($\sigma_i = 0.50$) and 2 non-core occupations ($\sigma_i = 2.5$). In the baseline, we consider that the two dimensions are unrelated. Here, we consider how different variable of interest evolve with respect to θ. We consider support the interval $[1,2]$ as support for θ.[40]

Figure E1 starts by showing the positive relationship between firm size, as measured by sales and outsourcing intensity defined here as the ratio of outsourcing expenditures to sales.

Figure E2 presents a set of results. Panel 2(a) shows how the optimal scale of production and sales evolves as productivity increase. Production Y increases log-linearly with θ, with an elasticity close to 1. Sales, which are proportional to profit in this model, increases also linearly but because the elasticity of demand ε is finite, the revenue / profit function is concave in productivity. Panel 2(b) displays the effect of productivity on the use of in-house and outsourced labor services. Both increase with a roughly constant elasticity (log-linear) but we see that, due to the rising cost of hiring in-house faced by monopsonic employers as they scale-up, they progressively outsource more, resulting in a shrinking in-house to outsourced labor ratio. Panel 2(c) show that this declining ratio is heterogeneous across occupations. It displays the ratio for a core and non-core occupation with the same weight in the Cobb-Douglas

[40]The number of occupations is set to 4 so that $\mathbf{N} = \{1,2,3,4\}$. The vector of parameters regarding occupations in production function are as followed: $\boldsymbol{\sigma} = [0.5, 0.5, 2.5, 2.5]$; ; $\boldsymbol{\gamma r} = [0.25, 0.25, 0.25, 0.25]$; $\boldsymbol{\mu} = [.75, .75, .75, .75]$; $\boldsymbol{\alpha} = [1/3, 1/6, 1/3, 1/6]$. Regarding labor supply, we set: $\boldsymbol{\rho} = [1,1,1,1]$, $\mathbf{a} = [1,1,1,1]$. The other parameters are: $\varepsilon = 5$, $I = 1$ and $N_i = 1, \forall i \in \mathbf{N}$.

FIGURE E1. Outsourcing intensity as a function of sales following an increase in productivity



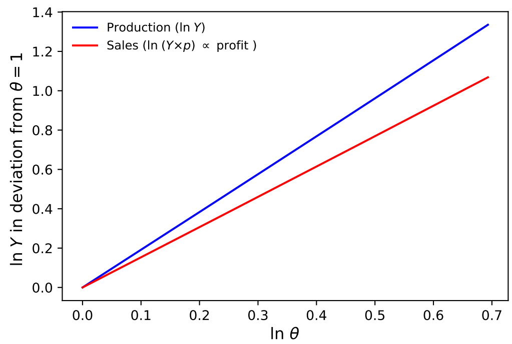
Notes: The figure provide comparative statics with respect to an increase in Hicks neutral productivity shifter θ. We consider support the interval $[1, 2]$ as support for θ.

production function. Panel 2(d) makes the same point but focusing on the level of the cost share represented by in-house labor. We see that both shares are high for core occupation and tend to increase with size while the opposite is true of the two non-core occupations.

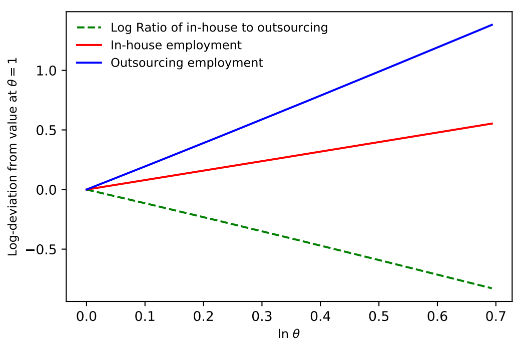
The figure E3 displays four other comparative statics. Panel 3(a) shows the wage in level. Unsurprisingly, high-skill occupations (1 and 3) have the highest wages. We see however that the firm size wage premium is stronger among core occupations independently of skill-level. Panel 3(b) show how log-wage deviates from the initial situation. We see that core and non-core determines almost entirely the magnitude of the elasticity of wage to size. Overall, panels 3(a) and 3(b) are consistent with the empirical existence of a size wage-premium (Oi and Idson, 1999). Moreover, it has been documented that skill-wage premium is stronger in large firms. Through the lenses of our model, this would imply that skills (α_i) tend to be higher in more core occupation (smaller σ_is). Here, we have explicitly made the choice of decorrelating these dimensions, it is plausible however that skill and "core-ness" are positively correlated, in particular if coreness of an occupation is determined by how difficult the tasks it entails are to codify, it seems likely that such tasks might also be requiring high-skill

FIGURE E2. Scale of production, employment and outsourcing as productivity goes up

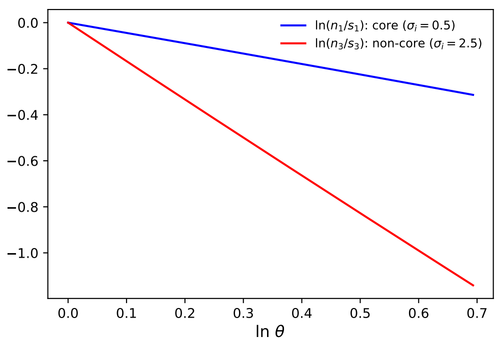
(a) Scale of production



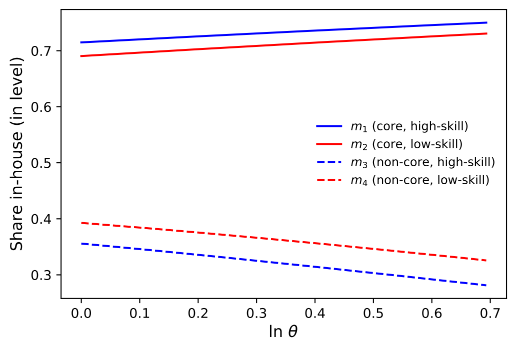
(b) Employment : in-house versus outsourced



(c) In-house ratio in employment, across occupations



(d) In-house cost share, across occupations

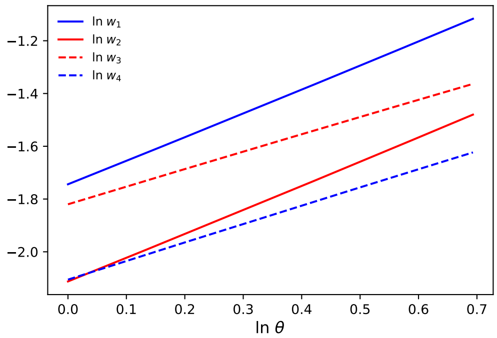


Notes: The figure provide comparative statics with respect to an increase in Hicks neutral productivity shifter θ. We consider support the interval $[1, 2]$ as support for θ.

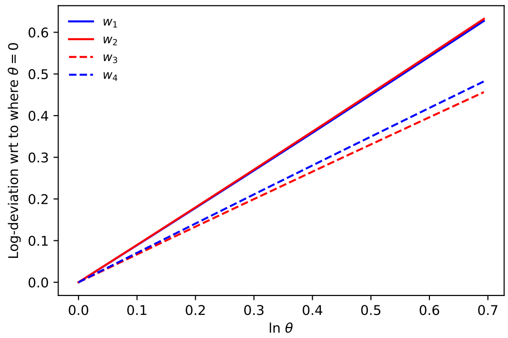
labor and have a high economic return. Correlating these dimensions is straightforward in our model and strengthens the key results displayed below regarding the increase in the share of outsourcing and the increase in occupational specialization. Panel 3(c) compute the HHI index for in-house labor and cost at the firm level across occupations. We see that both employment and cost based HHI increases and that this increase is stronger in terms of costs. This indicates that overall, the firm is concentrating its employment in and spending on in-house labor services on a fewer core occupations. This is a prediction we will be able to test explicitly. Finally, panel 3(d) presents how spending on outsourcing over in-house labor cost (both summed across all occupations) evolves as productivity and scale go up . We do see an increase in this ratio which is somewhat less marked than the equivalent ratio in terms of employment because of the size wage premium associated (see Panel 2(b)) with in-house labor services.

FIGURE E3. Wage premium, outsourcing over in-house cost and index of occupational segregation (HHI)

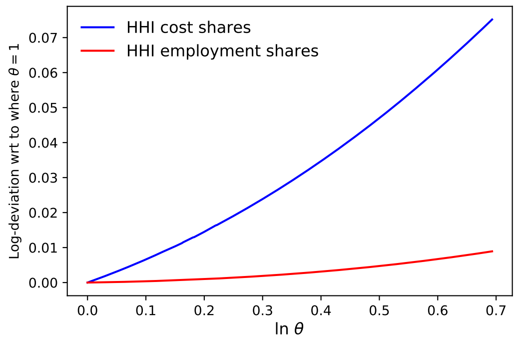
(a) Wage by occupation



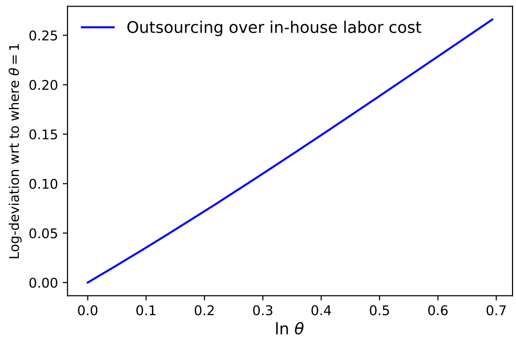
(b) – normalized with respect to $\theta = 1$



(c) HHI in terms of cost shares



(d) Outsourcing over in house labor cost



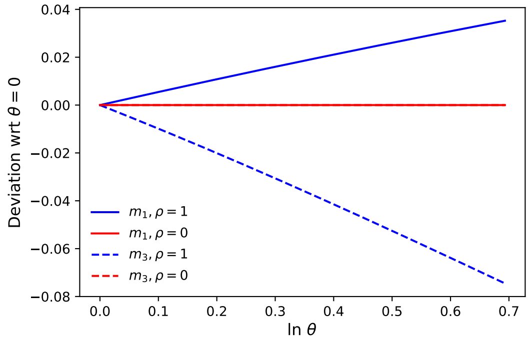
Notes: The figure provide comparative statics with respect to an increase in Hicks neutral productivity shifter θ. We consider support the interval $[1, 2]$ as support for θ.

Additional comparative statics. Figure E4 displays similar comparative statics comparing the baseline case and the case ($\rho_i = 1$) with no market power on the in-house labor market ($\rho_i = 0$). It shows see that the in-house cost share per occupation (m_i) and the the outsourcing over in-house labor cost ratio does not change with productivity when wage are competitively set, highlighting the key role of labor market frictions in explaining our results.

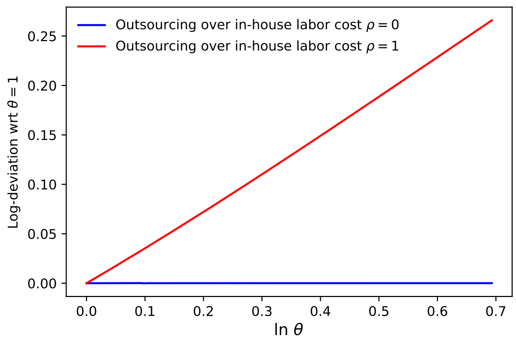
As mentioned above, broadband is also likely to result in a decrease of outsourcing cost, which we capture with a decline in the term γ_{ij}. Decrease in this parameter leads to broadly similar comparative statics as the previous case as display in Figure E5.

FIGURE E4. Increase in θ with $\rho_i = 1$ (baseline) and $\rho_i = 0$ (no market power)

(a) Scale of production



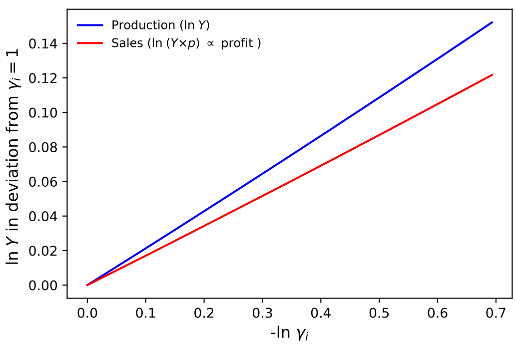
(b) Employment : in-house versus outsourced



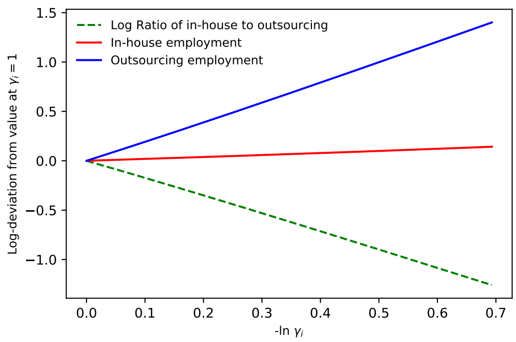
Notes: The figure provide comparative statics with respect to an increase in Hicks neutral productivity shifter θ. We consider support the interval $[1, 2]$ as support for θ.

FIGURE E5. Scale of production, employment and outsourcing as **cost of outsourcing** goes down

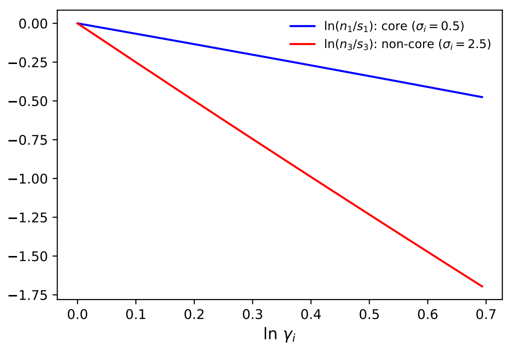
(a) Scale of production



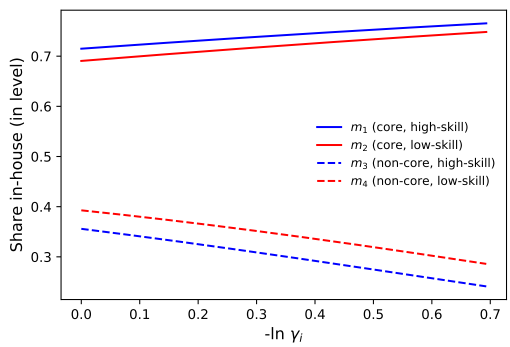
(b) Employment : in-house versus outsourced



(c) In-house ratio in employment, across occupations



(d) In-house cost share, across occupations



Notes: The figure provide comparative statics with respect to an decrease in shifter of outsourcing cost γ_{ij} which occurs uniformly across occupations. We consider $\gamma : 1 \rightarrow 0.5$.