Technological Change and Domestic Outsourcing*

Antonin Bergeaud Clément Malgouyres
Clément Mazet-Sonilhac Sara Signorelli

October 2023

Abstract

Does domestic outsourcing react to technological change? We study the staggered diffusion of broadband internet in France in the 2000s, and show that connected firms increased their outsourcing expenditures while decreasing the diversity of occupations they employ in-house. Meanwhile, employment in non-core occupations became increasingly concentrated in firms specializing in subcontracting services. Finally, we provide evidence that workers in high-skill occupations experienced salary gains from being outsourced, while workers in low-skill occupations lost out. Overall, we show that the deployment of new technologies stimulated domestic outsourcing in this context, with important implications for labor market inequality.

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

*Addresses: Bergeaud: HEC Paris, CEP-LSE and CEPR; Malgouyres: PSE, CREST and Institut des Politiques Publiques; Mazet-Sonilhac: Bocconi University, Finance Dept., Collège de France and IGIER; Signorelli: University of Amsterdam. We would like to thank Philippe Aghion, Adrien Bilal, Johannes Boehm, Mathilde Le Moigne and the seminar participants at College de France, Sciences Po, X-CREST, LAGV, JMA, AFSE, and RCEA for their helpful comments. Any errors are our own. Corresponding author: Antonin Bergeaud: 1 Rue de la Libération, 78350 Jouy-en-Josas, FR. Email: bergeaud@hec.fr

1 Introduction

The implications of international outsourcing (or offshoring), i.e., 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 research 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 remains silent on the factors underlying these trends, particularly whether domestic outsourcing is primarily driven by changes in (labor market) institutions or by the adoption of new technologies.

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, namely broadband internet (BI hereafter), on employment outsourcing by French firms. As explained in Abramovsky and Griffith (2006), general-purpose technologies, including ICT technologies, prompt important organizational changes within firms. In particular, by reducing communication and transaction costs, the diffusion of BI is likely to modify firms' optimal boundaries and affect the terms of their make-or-buy decisions (see, e.g., Lewis and Sappington, 1991; Garicano and Rossi-Hansberg, 2006; Bloom et al., 2014; Aghion et al., 2019b). Therefore, it is natural to assess BI's impact on the allocation of workers across businesses, particularly through an increase in domestic outsourcing.

We start by laying out a simple conceptual framework to discuss the different channels, identified by prior work, through which BI can impact domestic outsourcing. We

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

then use employee- and employer-level administrative data on the universe of French workers and firms from 1996 to 2007 to explore these links empirically. Specifically, we adopt an event study design, exploiting the staggered roll-out of the broadband infrastructure: the deployment of BI in France took nearly a decade to complete, from 1999 to 2007, and was particularly slow in the beginning. While cities connected earlier were typically larger and had higher population density, we exploit quasi-random variations in the timing of broadband connection across cities in the same province (*département* in French) with similar characteristics. Building on recent advances in the two-way fixed effects literature on staggered adoption contexts.[2], we adopt a stacked difference-in-differences approach (as developped in Cengiz et al., 2019).

Our results show that BI increased firms' expenditures in outsourcing services and the occupational segregation of establishments, as captured by a Herfindahl-Hirschman Index (HHI) of concentration. This effect is driven both by changes within existing establishments and changes in establishment composition, through entries and exits in the cities connected to the internet.[3] Interestingly, outsourcing affects not only low-skill workers such as cleaners or drivers, but also skilled professionals. To track these effects, we establish a list of sectors specializing in services to other firms, then we identify the occupations that are likely to be outsourced to such sectors.[4] We show that, following BI expansion, workers employed in "outsourceable occupations" in both the high- and low-skill segments 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

[2]See e.g. Callaway and Sant'Anna (2020); De Chaisemartin and d'Haultfoeuille (2020); Goodman-Bacon (2018); Borusyak et al. (2021).

[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. Especially with regard to occupation, this result is particularly pronounced in France.

[4]A detailed description of these definitions is reported in Appendix B. The definition of low-skill outsourcing follows Goldschmidt and Schmieder (2017) closely 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).

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 earning disparities.

The analysis in this paper contributes to several strands of literature. First, we relate to recent empirical analyses exploiting BI diffusion as a plausibly exogenous technological and informational shock. In France, Malgouyres et al. (2021) show that the BI expansion was associated with an increase in imports among affected firms.[5]

More closely related to our study, Akerman et al. (2015) evaluate the skill bias generated by BI in Norway. They find that broadband availability increases both firm productivity and the high-skill wage premium. We contribute to this literature by diving deeper into the concrete mechanisms through which skill bias occurs. In particular, we show that it arises partly through the reorganization of firms and the reallocation of high and low-skill workers across firms, not just through a shift in the firm-level production function. Given frictional labor markets, the allocation of workers to firms has important implications for the distribution of earnings (Song et al., 2019). Consistent with this point, we show that these reorganizations affect heterogeneous workers in a way that amplifies preexisting wage inequalities.

Second, our paper is linked to the literature on domestic outsourcing.[6] Goldschmidt and Schmieder (2017) show how Germany has experienced an explosion in the domestic outsourcing of low-skill occupations since the early 1990s, which resulted in wage reductions for outsourced workers, mainly driven by the loss of firm-specific

[5]Consistently, Akerman et al. (2018) show that BI narrowed the role of distance in explaining bilateral trade in Norway. In the United Kingdom, DeStefano et al. (2018) find that a similar shock affected firm size but not firm productivity, which suggests that local institutions matter in evaluating the effect of broadband expansion (see also Haller and Lyons, 2015). On the contrary, Bertschek et al. (2013) finds a positive causal effect of BI on innovation and productivity in Germany (see also Fabling and Grimes, 2021; Grimes et al., 2012).

[6]Weil (2014) describes how the nature of work has changed in the 21st century as more large companies have switched to a "fissured workplace" business model.

rents.[7] Bilal and Lhuillier (2021) consider the aggregate implications of domestic outsourcing in terms of wages and productivity. Other work has further highlighted a possible link between technology and outsourcing.[8]

We contribute to this literature in three main aspects. First, we exploit a natural experiment that allows us to test the causality of the relation between a specific technology—BI —and domestic outsourcing. Second, we extend the spectrum of affected workers to high-skill occupations, which enables us to explore the heterogeneity of the wage effect across skill groups.[9] Third, we capture margins of adjustment largely ignored by previous studies. Although most earlier work focuses on individual transitions from in-house to outsourced activities (Dube and Kaplan, 2010), or on plant-level outsourcing events (Goldschmidt and Schmieder, 2017), the rise in domestic outsourcing is likely also driven by newly-created firms using outsourcing services more intensely than firms exiting the market, and by the concentration of new entrants in the service-providing sectors. Our city-level approach allows us to capture such margins, and comparing them with the results obtained from a plant-level analysis on incumbents clarifies the magnitude of the composition effects. In most cases, we find that city-level estimates are larger than plant-level ones, consistent with the notion that firm entry is an amplifier of the effect of BI on domestic outsourcing.

The rest of the paper is organized as follows. Section 2 briefly discusses the mechanisms through which a technological change such as BI can be related to an increase

[7]They focus on four emblematic activities— cleaning, logistics, security, and food service— and introduce an innovative measure to capture the outsourcing of such tasks. Drenik et al. (2020) quantify the resulting loss in wages and show that outsourced workers share about half the share of the firm-specific wage premium that insiders earn.

[8]Using data on US firms, Fort (2017) shows that technology lowers the cost of coordination within a firm and ultimately increases the fragmentation of production. Cortes and Salvatori (2019) show that firms in the UK have become increasingly specialized over the past 20 years, partly driven by an increase in domestic outsourcing for 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 less-educated workers.

[9]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, including some high-skill occupations such as advertisers, accountants, IT specialists, and legal professionals (Ono, 2003; Berlingieri, 2014; Goldschmidt and Schmieder, 2017).

in domestic outsourcing. Section 3 details the data and our empirical methodology, while Section 4 presents results on the effect of BI on outsourcing. Finally, Section 5 looks more closely at the consequence of outsourcing on workers' wage trajectories, and Section 6 concludes the paper.

2 Technology-led domestic outsourcing?

In this section, we discuss the mechanisms through which access to broadband internet can affect a firm's decision to outsource. Abramovsky and Griffith (2006) argue that ICT adoption at the firm level is likely to lower the adjustment and monitoring costs of outsourcing. BI could act in the same way, by facilitating the exchange of information between regular firms and their various subcontractors.[10] Given that outsourcing costs are heterogeneous across firms, for instance because of differences in product complexity (Tadelis, 2002), we expect BI to lead to a greater proportion of firms finding it financially beneficial to outsource certain activities rather than produce them in-house.[11]

Abraham and Taylor (1996) distinguish three additional motives underpinning a firm's decision to outsource its labor: (i) the realization of wage and benefit savings, (ii) access to the specialized skills possessed by subcontractors, and (iii) the management of volatility in output demand. We focus our discussion on the first two motives. The first is connected to a recent literature that emphasizes how labor market frictions lead to some rent-sharing between firms and in-house workers (Card et al., 2018; Lamadon

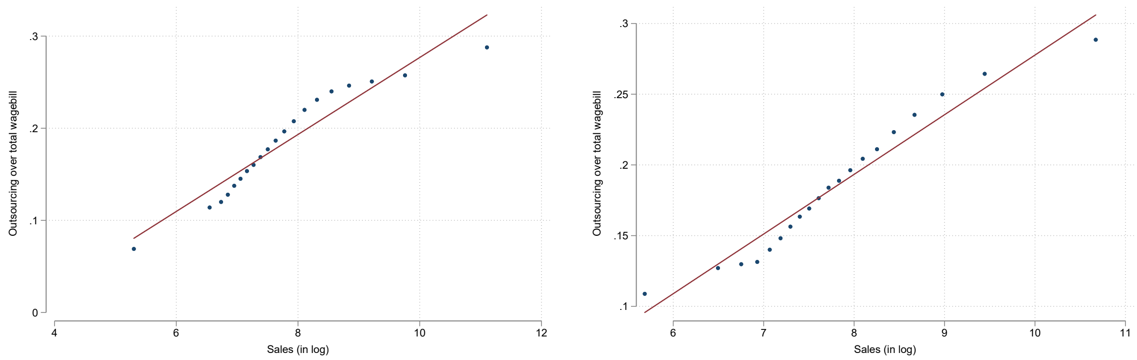
[10]In fact, as we will show in Figure II, BI caused an increase in firm-level IT hardware intensity. Accordingly, the BI shock might be operating through the very mechanism conceptually and descriptively highlighted by Abramovsky and Griffith (2006). See subsection "Broadband availability versus broadband adoption" in Section 3.3 for more details on this result.

[11] It is important to highlight the distinction between temporary agency workers and outsourced workers. Temporary agency workers are employed by a staffing agency and assigned to a client firm, where they perform tasks as in-house employees (e.g., covering the front desk for an individual on sick leave). The utilization of temporary agency workers is subject to regulatory constraints and limitations. In contrast, outsourced workers, which are the focus of this paper, are provided by a specialized firm and subject to minimal control by the client company.

et al., 2019; Kline et al., 2019). This creates an incentive to outsource some occupations to outside contractors, as the market price for these services is independent of the firm's rents.[12] This incentive grows as firms become more productive or enjoy larger rents. In line with Akerman et al. (2015), a first effect of BI is to increase firm-level productivity, triggering more outsourcing as firms expand in imperfect labor markets. The relevance of this motive is consistent with a key empirical regularity: the positive relationship between firm size (measured by revenue) and the propensity to outsource, which is shown in Figure I.

FIGURE I. Correlation between outsourcing intensity and firm sales

(a) Unconditional (b) Absorbing sector fixed effects



Notes: The figure presents the binned scatter plot for the correlation between firm outsourcing intensity and the log of firm sales. The left panel reports the row correlation and the right panel reports the correlation after the 5-digit sector fixed effect is absorbed.

The second motive for outsourcing is linked to subcontractors' ability to provide cheaper quality-adjusted services, which derives from their access to more specialized skills thanks to economies of scale (Bartel et al., 2005). Subcontracting firms offer their services to multiple clients. Accordingly, their scale of production is likely larger than those of their individual clients if those clients were to produce the service in-house. To the extent that BI boosts firms' productivity, it will increase the size of the market for outsourcing services, which in turn allows for greater economies of scales, thus magnifying the incentives to outsource.

[12]Even if there is some rent-sharing with respect to outsourced workers, to the extent that this rent-sharing is weaker than with respect to in-house workers— as shown by Drenik et al. (2020) in the case of temp workers— firms will remain incentivized to outsource part of their production.

Finally, in the case of IT services, there is one additional motive to be noted. BI is likely to boost directly the efficiency of IT subcontractors, driving their marginal revenue productivity further up. For the IT-intensive outsourcing sector, the BI shock corresponds to both an increase in efficiency and a reduction in cost. If barriers to entry are low, new IT service firms will enter the market and put a downward pressure on prices, increasing regular firms' incentives to outsource.[13]

In Appendix E, we sketch a simple partial equilibrium model to highlight a non-trivial mechanism related to the first motive described by Abraham and Taylor (1996): wage and benefit savings in presence of labor market frictions. We show how a positive, Hicks-neutral productivity shock can lead firms to outsource a (non-random) part of their workforce. This effect relies on the assumption that firms possess some wage-setting power when hiring in-house workers, but they are price-takers when they buy outsourcing services.[14] This difference implies that the relative cost of an in-house workforce versus outsourcing increases as the firm's optimal size goes up, following for instance a surge in productivity. This mechanism generates the positive correlation between outsourcing intensity and size displayed in Figure I. In this context, a technological shock such as broadband expansion, which boosts firms' productivity, leads to more outsourcing. Importantly, the model sketched in Appendix E generates two additional testable predictions that we bring to the data. First, BI increases outsourcing intensity, measured as the ratio of outsourcing costs to in-house workers' payroll. Second, under a broad set of conditions, it also leads firms to become more segregated in the type of workers they employ, by hiring a smaller set of occupations

[13] Accordingly, the producer price index (PPI) of the broader sector that includes most outsourcing service firms (unfortunately, PPI values are only available at the 2-digit sector level) over the PPI of the total economy declined by about 0.05 between 2000 and 2007.

[14] This assumption is in line with empirical evidence obtained in the case of temporary workers— a specific instance of domestic outsourcing (Drenik et al., 2020). Firms face imperfectly elastic labor supply curves and will accordingly need to raise wages in order to expand their in-house workforce. Imperfect elasticity in labor supply is generated by idiosyncratic workers' tastes for the amenities offered by firms (such as working conditions, commute, and corporate culture), as is standard in the recent literature on monopsony (Card et al., 2018; Lamadon et al., 2019). Given asymmetric information on workers' heterogeneous preferences for jobs, firms are unable to perfectly discriminate and fully price these amenities into individual-specific wages.

in-house.

To sum up, a succinct review of potential mechanisms reveals that there are several theoretical reasons why BI access could boost outsourcing, including lower transaction costs, how firms expand in presence of imperfect labor markets, and rising productivity among IT subcontractors. We now turn to the empirical analysis to assess the magnitude of this effect.

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, the (*Déclaration annuelle des données sociales* or DADS). This dataset constitutes the primary source used to compute payroll taxes and gives detailed individual information including salary, hours worked, occupation, age, gender, and an identifier showing the employing firm and establishment. It is not possible to follow these workers over time, except for a random sample of 1 out of every 24 workers, which we refer to as the DADS panel. In this data, we also lack information on the level of education attained. Therefore, we always define the skilled nature of jobs based on the occupation codes, which are highly correlated with educational attainment in the French context.

Our second main data source is the administrative financial records curated by the Bank of France and called *Fichier inter-Bancaire des Entreprises* (or FiBEn). FiBEn includes firms with a yearly turnover of 750,000 euros or higher, and has the advantage of recording outsourcing expenditures as a separate item.[15] Contrary to the DADS,

[15]In contrast, in the more widely used *FIchier Complet Unifié de Suse* (FICUS), the value of outsourcing expenditures is included within a broader category of spending, making it impossible to isolate the impact of broadband internet on this particular sub-category.

FiBEn does not provide information at the establishment level, but only at the more aggregate firm level.[16] 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.

We construct several measures as proxies for domestic outsourcing. The most direct one is firm-level outsourcing expenditures, scaled by the in-house wage bill.[17] To gain 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.[18] Finally, we define typical outsourcing sectors and outsourceable occupations as follows, consistently with Goldschmidt and Schmieder (2017):

- **Outsourceable occupations**[19]: Non-core occupations that are likely to be outsourced, off-shored, or subcontracted. We divide them into high-skill outsourceable occupations (IT specialists and consultants) and low-skill outsourceable occupations (cleaners, security workers, drivers, and logistics workers). The full list of occupations and their definitions is detailed in Appendix B.

- **Outsourcing sectors**: Administrative and business support service sectors. High-

[16]Our preferred model for the outsourcing expenditure results keeps multi-establishment firms and assigns outsourcing intensity to the location of the headquarters, since information on the other plant locations is not available.

[17]The available information on outsourcing expenditure does not allow us to distinguish between international outsourcing (offshoring) and domestic outsourcing. We expect domestic outsourcing to represent the first order response as international trade in services remains fairly rare and a lot of the services that we consider (especially in the low-skill outsourcing sector) are inherently non-tradable activities (cleaning and security for instance). Moreover, using survey data from 2016 (INSEE Survey on Global value chains of SME), 7.4% of firms report doing domestic outsourcing versus only 2.4% for offshoring, an odds ratio of about 3.

[18]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 level (also known as the PCS classification), and ψ_{oi} is the share of the wage bill of the establishment attributed to workers employed in each occupation. At the city level c, we compute the average establishment concentration as follows: $HHI_c = \sum_{i=1}^{N} \omega_{ic} HHI_{ic}$, where ω_{ic} is the share of employment in the city accounted for by each establishment i.

[19]Throughout the paper, we make the assumption that one occupation is a fixed bundle of tasks. This is admittedly a simplification, since the literature has shown that the task content of jobs evolves over time, including evolution 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 the administrative data, which does not allow us to observe tasks beyond their occupational dimension.

skill outsourcing sectors: IT and consulting sectors. Low-skill outsourcing sectors: cleaning, security, driving, and logistics services. The full list of sectors and their definitions is detailed in Appendix B.

At the city level, we look at the expansion of the outsourcing sectors by measuring their share of overall employment, and we look at shifts towards these sectors by measuring the share of workers in outsourceable occupations that are employed by them and not elsewhere. At the individual level, we look at the mobility of workers across firms and focus on moves towards the outsourcing sectors.

Data on broadband expansion ADSL (Asymmetric Digital Subscriber Line) is a data communication technology that enables fast data transmission over copper telephone lines. The eligibility for ADSL depends on the distance between the final customer and a Local Exchange unit (LE), since the quality of the signal decreases as it is routed over the copper lines. Our main information of interest documents the date when each LE was upgraded in mainland France.[20] While we do not have data on firm-level use of BI, suggestive evidence suggests that ADSL was widely adopted by firms once it became available to them.[21] Additionally, we 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 census block 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 as $\widetilde{Z}_{it}$, is a time-weighted percentage of the area covered by BI in

[20] See Appendix C for a more detailed account of the ADSL expansion. The data was collected through the official website by Malgouyres et al. (2021) and validated manually.

[21] A 2016 survey by the French regulator shows 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. Additionally, we find that firms increased their investment in IT hardware in response to the the BI shock. See subsection "Broadband availability versus broadband adoption" in Section 3.3 for more details on this result.

city i. It is formally defined as:

$$\widetilde{Z}_{it} = \sum_{b \in i} D_{b,t} \frac{\mathcal{A}_{b,t}}{\sum\limits_{b' \in i} \mathcal{A}_{b',t}}, \tag{1}$$

where $b \in i$ denotes the census blocks included in city i and $D_{b,t}$ is the share of the days of year t with BI access in b. Finally, $\mathcal{A}_{b,t}$ denotes the area covered by census block b. The variable $\widetilde{Z}_{it}$ is continuous with support between 0 and 1.[22] In most of our study, we discretize $\tilde{Z}$ by defining the year of treatment as the first year where $\tilde{Z}$ becomes positive. We denote as $C_{i,t}$ the corresponding discrete (binary) variable, and we refer to it as the "year of connection." This transformation 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 an event study design. Figure A3 in Appendix A shows the evolution of the average $\tilde{Z}$ within cities after the year of first connection. On average, the first year is characterized by a degree of connection of 30%, which jumps to 60% the following year and quickly reaches 80%.

Table I reports the summary statistics for the main outcomes of interest. Panel A describes the city-level data while Panel B describes the establishment level data. On average, firms spend on outsourcing 30% of the amount they spend on in-house workers. The high- and low-skill outsourcing sectors account for 2% and 11% of employment, respectively, and 8% (25%) of employment in high-skill (low-skill) outsourceable occupations is located within their respective outsourcing sector.

[22] $\widetilde{Z}_{it}$ will be equal to one if all of its areas had BI access for the entire year. It will be equal to 1/2 if the entire city had access to broadband over half the year t. In practice, however, it is very strongly concentrated at 0 and 1, with very few intermediate observations; see Figure A2 in Appendix A.

TABLE I. Summary Statistics

	mean	(sd)
Panel A: City level		
Share high skill workers	0.07	(0.07)
Occupational concentration (HHI)	0.38	(0.16)
Value added over wage bill (2010 euros)	2.08	(1.84)
Outsourcing exp. over wage bill (2010 euros)	0.31	(0.58)
Share of empl. in high-skill outs. services	0.02	(0.09)
Sh. outsourceable workers in HS outs. Services	0.08	(0.22)
Share of empl. in low-skill outs. services	0.11	(0.24)
Sh. outsourceable workers in LS outs. Services	0.25	(0.38)
Average establishment size	77	(263)
N. of establishments in city	13.2	(59.7)
Panel B: Establishment level		
Share high skill workers	0.11	(0.17)
Occupational concentration (HHI)	0.38	(0.19)
Value added over wage bill (2010 euros)	2.19	(5.78)
Outsourcing exp. over wage bill (2010 euros)	0.40	(1.13)

Notes : This table presents the summary statistics of the main variables used in the regression. Averages and standard deviations are computed over the full period from 1997 to 2007. Observations: 172,132 for Panel A and 2,439,340 for Panel B

3.2 The diffusion of Broadband Internet in France

As evidenced by Malgouyres et al. (2021), the deployment of BI beyond France's largest cities began very slowly at the dawn of the 2000s, due to multiple reasons linked to the country's historical telecom supplier.[23] It then accelerated between 2004 and 2007, thanks to the involvement of regional governments who favored competition among providers. Table II 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.

[23]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 in 2002 with what was essentially a government bailout. Urged on by the government—which increased its stake in the firm during the bailout —in 2003, France Télécom pledged to cover 90% of the French (mainland) population by the end of 2005.

TABLE II. Year of connection by city

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).

Given that our main effects of interest are identified out of the gradual diffusion of BI across cities, addressing the endogeneity of the decision to "treat" one LE before another deserves special consideration. Malgouyres et al. (2021) describe how broadband expansion occurred to maximize population coverage with no special consideration for economic potential (see Appendix C for more detail). Our identification strategy relies on the assumption that, within a given department,[24] 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.[25] In the next section, we detail how we test the plausibility of this assumption.

[24]In France, a department *(département)* is a type of administrative division. There are 94 department located in mainland France, with an average population of 700,000 in 2019. On average, each department encompasses around 300 cities, which will serve as the primary unit of analysis in this paper. The map at the top of Figure A1 provides a visual representation of the size of both departments and cities in France.

[25]See Malgouyres et al. (2021) for an in-depth discussion of the determinants of the geography of broadband expansion.

3.3 Empirical Strategy

To identify the causal effect of BI, we exploit the fact that the dissemination of the ADSL technology across cities was staggered over a period of nearly 10 years, from 1999 to 2007. While the timing of diffusion is likely to be correlated with the size and density of the city, 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:

$$Y_{c,t} = \sum_{\tau=-k}^{k'} \alpha_\tau \mathbb{1}\left\{t = t_c + \tau\right\} + \boldsymbol{\gamma_t X_{c,t}} + \nu_c + \varepsilon_{c,t} \tag{2}$$

where Y is the outcome of interest, t_c is the year of BI arrival in 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 fixed effects. 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 BI, compared to cities within the same department and with similar densities that have yet to be connected to BI. 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 BI expansion. To avoid capturing endogenous changes in establishment location as a response to BI diffusion, we fix the city as the one where each establishment appears for the first time in our data, and we keep it constant throughout the period.

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 were treated by the end of the period, we adopt the stacked difference-in-differences approach employed by Vannutelli (2020), which is also similar to Cengiz et al. (2019) and Deshpande and Li (2019).[26] The latter consists of 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 the control cohorts are only considered for the period preceding their respective treatment. The cohort receiving BI in 2007 serves as a 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:

$$Y_{c,s,t} = \alpha_0 Treat_{c,s} + \sum_{\tau=-k}^{k'} \alpha_\tau D^\tau \times Treat_{c,s} + \sum_{\tau=-k}^{k'} \beta_\tau D^\tau + \boldsymbol{\gamma_t X_{c,t}} + \nu_c + \pi_s + \varepsilon_{c,s,t} \quad (3)$$

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 $\mathbb{1}\{t = t_c + \tau\}$ in equation (2)) and α_τ identify the pre- and post-treatment dynamic effects. We also add fixed effects for each of the samples stacked (π_s). Standard errors are clustered at the department level, which also accounts 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

[26] 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) in a variety of corporate finance applications. Gardner (2022) shows that this approach estimates a weighted average of the cohort-period specific average treatment effect on the treated where the weights are a positive function of the number of treated units and the variance of treatment within each stacked event.

two-way fixed effects model presented in equation (2), for comparison.[27] Finally, in the main graphs of the paper we also show the robustness of our results to controlling for the propensity of cities to adopt BI early based on observable baseline characteristics, which we describe in detail below.

Tests of the identification assumptions First, our approach assumes that the timing of BI arrival within a department is as good as random, conditional on trends explained by city density, and on time and city fixed effects. The latter might be violated if BI is deployed first in cities that were already on a positive productivity trend or that are larger in terms of economic activity (conditional on city density), in cities with a specific industry composition, or in cities going through changes in the political party in power. To test the validity of this assumption, we run alternative specifications adding controls for each one of these elements, constructed as follows:

1. **Productivity trends**: growth in labor productivity between 1996 and 1998, measured as average value added per worker in the city, interacted with year fixed effects.[28]

2. **Size of economic activity**: number of establishments in the city in 1997, interacted with year fixed effects.

3. **Industry composition**: shares of employment in 11 broad economic sectors in 1997, interacted with year fixed effects.

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

[27] In the standard TWFE model it is necessary to exclude two dummies from the regression, respectively for $\tau = -1$ and $\tau = -6$, to avoid multi-collinearity and to identify the fully dynamic underlying data generating process in this staggered design with year and individual fixed effects (Borusyak et al., 2021). However, the stacked approach in our main analysis does not require this restriction.

[28] The value-added information is sourced from the FICUS administrative dataset, which encompasses all firms (except financial services firms) and does not enforce a minimum yearly turnover of 750,000 euros, unlike FiBEn. Although this limitation affects the number of firms (approximately 30% of the FiBEN sample), FiBEn's coverage is generally considered extensive, representing about 70% of the total workforce and value-added data.

election of 1995 and a dummy for change in the political majority between 1995 and 2002, both interacted with a full set of year fixed effects.[29]

Results are presented in Appendix A.3. Given that adding multiple controls quickly runs into the course of dimensionality, we also compute a propensity score for early treatment based on city baseline characteristics, which we interact with year fixed effects and which we include as controls.[30] Results are presented alongside the main coefficients in Section 4 to highlight their robustness. In addition, we use the propensity score to perform placebo tests, where late receivers of broadband (2004 to 2007) are split into high and low propensity to be early receivers (HPLR and LPLR). We then attribute a fictitious treatment date to each city among the late receivers by simply taking the actual date minus 4 years. Finally, we run a regular event-study regression comparing HPLR to LPLR on a sample period, where none of them is actually treated (1997 to 2003) and where therefore we expect no effect of BI. These placebo tests are presented in Appendix A.3.

Our approach further assumes that cities located in the same department and receiving BI later in time are not affected by neighboring areas adopting the technology earlier. This hypothesis would be violated if the arrival of BI were associated with important changes in workers' mobility across cities. To test the plausibility of this assumption, we show that our results are robust to: i) running the analysis at the broader commuting zone (CZ) level instead of the city level,[31] ii) including commut-

[29]These variables are proxies 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 ran in the 1995 and 2002 presidential elections as 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, and 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, and Corinne Lepage.

[30]We define early adopters as the cities receiving broadband in the first 3 cohorts (1999 to 2001). We then estimate a linear probability model where the predictors are the number of establishments in the city, the sectoral composition, the average level of productivity of all firms (measured before 1999), and the productivity growth observed between 1996 and 1998.

[31]Commuting Zones (*"Zones d'emploi"*) or CZs are smaller entities than departments but larger than

ing zone × year fixed effects instead of department × year fixed effects,[32] and iii) adding controls for the BI coverage of other cities within the department to the main specification. All these robustness checks are presented in Appendix A.3. In addition, our individual-level results show that the additional worker mobility generated by BI across cities is small enough to be considered negligible from the point of view of the labor markets of neighboring cities (results shown in Table III and discussed in section 4.2).

Finally, given that our main analysis applies a binary treatment indicator based on the year BI first arrives in the city, one might wonder whether the dynamic effects recovered are driven by the effect of the treatment becoming stronger over time or by the BI coverage of the city becoming larger over time (Figure A2 in Appendix A shows that coverage increases in the first two years following the first BI connection). We show that the dynamic effects are not confounded by changes in coverage, by providing results using the continuous variable $\tilde{Z}_{it}$ instead of the binary treatment. All of these robustness checks are presented in Appendix A.3.

Broadband availability versus broadband adoption. Our paper focuses on a reduced-form estimation in the sense that it captures the effect of local broadband availability on firm behavior and not the effect of firm-level adoption of the technology. Accordingly, our estimates reflect the combination of two effects: the effect of broadband availability on its adoption by firms and the effect of firm-level adoption on firm behavior. Given that we do not have data on firm-level adoption, we cannot disentangle them. However, we view the reduced-form coefficients as policy-relevant, since it is straightforward to increase local broadband availability through public policy, while

cities. They correspond to labor market areas. There are 297 CZs in France, which on average contain 117 cities. See Figure 1(b) for an illustration.

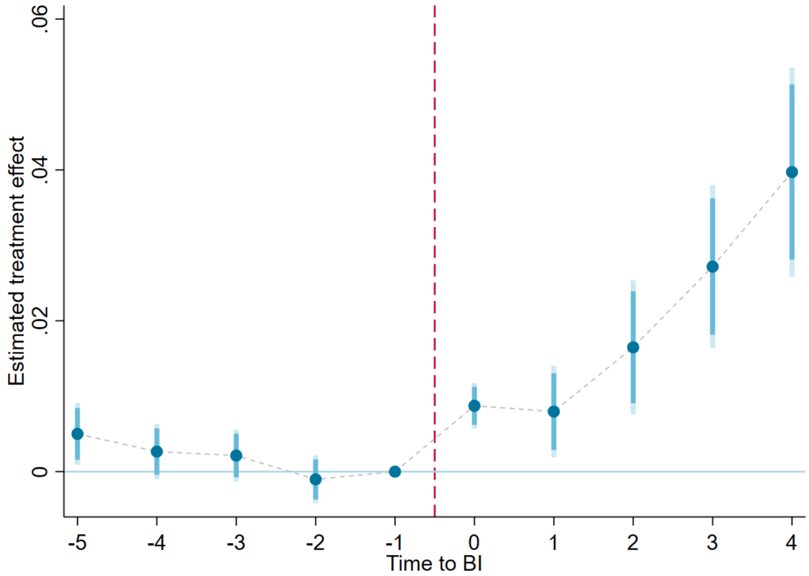
[32] In the presence of strong geographic spill-overs, restricting the control group within a commuting zone would magnify the bias and thus is expected to generate different coefficients.

manipulating adoption involves more intricate interventions.[33] Moreover, building on newly available data for a subset of the sample period (1998-2007), we are able to construct an indicator of IT hardware intensity defined as the (book) value of IT capital over total capital at the firm-level.[34] We average this variable at the city level and we assess, using the baseline specification described in Equation (3), whether it is positively affected by BI availability. As shown in Figure II, we find that this measure increases significantly following BI expansion. The increase is about 3 percentage points four years after connection (from an average pre-broadband level of 7%). Naturally, this is not a direct measure of internet adoption but rather a measure of investment in equipment likely to be complementary to internet usage. While we do not view this exercise as a direct first-stage analysis, we consider this result as supporting the notion that our reduced-form empirical approach effectively captures changes in the behavior of firms with respect to internet use and IT technology more broadly. Moreover, to the extent that ICT investment is associated with a higher propensity to outsource (Abramovsky and Griffith, 2006), this result implies that broadband expansion is likely to boost outsourcing through this channel.

[33]See e.g. Andrews et al. (2018) for a cross-country empirical investigation of the numerous determinants of ICT adoption by firms and their complex interaction.

[34]Recently, Lashkari et al. (2022) use this dataset to study the link between IT hardware investment and returns to scale.

FIGURE II. IT hardware capital and BI



Notes: This Figure shows regression coefficients and 90% and 95% confidence intervals from a dynamic event study (see equation (3) and text for a description of the stacked event-study approach) where the dependent variable is the average ratio of the stock of capital in IT hardware over total capital by firms located in a given city in year t. Following definitions from the BIC-RN dataset, this ratio is defined as (LB-LC)/(BJ-BK) for each firm and then averaged at the city level. Pre-broadband expansion average of the dependent variable equals 7%.

4 Empirical Evidence on BI and Outsourcing

4.1 At the city and establishment level

We start by looking at the causal effect of BI on outsourcing expenditures and occupational segregation across firms. Figure A4 in Appendix A summarizes the evolution of these two outcomes over the period of interest. The outsourcing expenditures over wages bill grew very steeply at the end of the 1990s, going from 0.38 in 1997 to 0.44 in 2000 (a growth of 16%), then stabilized at this high plateau.[35] The level of 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 (3) on these outcomes. Figure

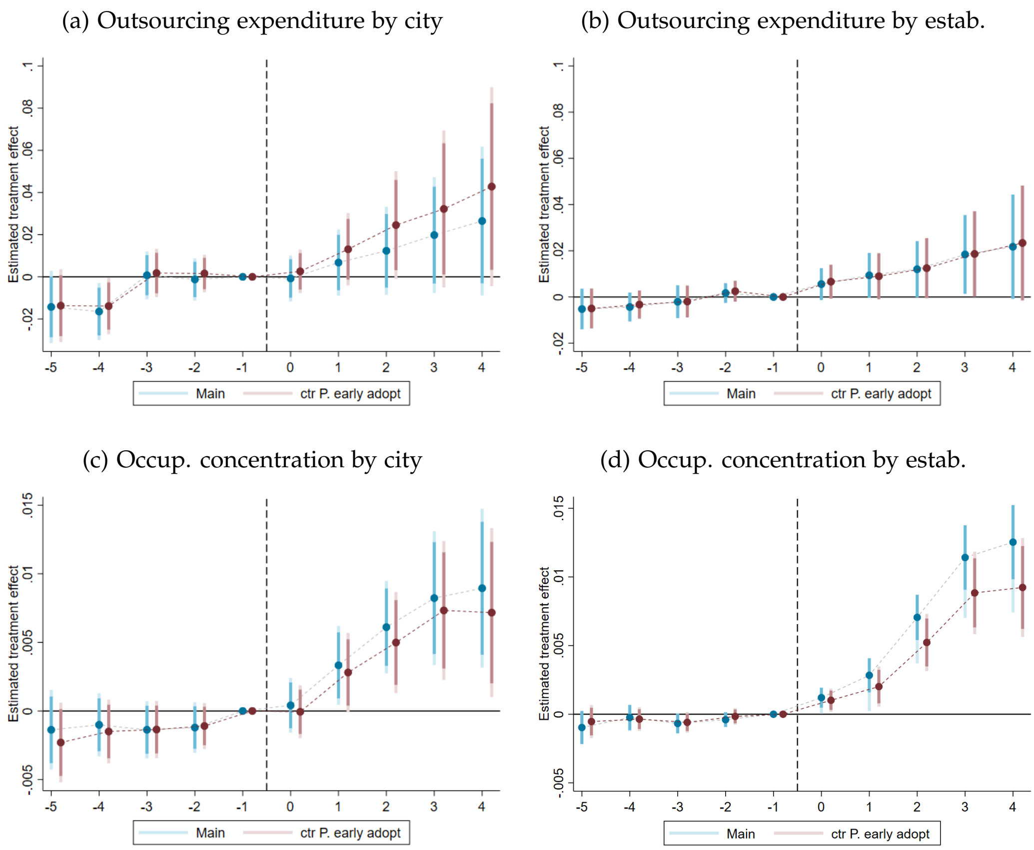
[35]Given that outsourcing can be a direct effect of productivity shocks, as described in our conceptual framework, we can imagine that the growth observed before the arrival of BI might be due to previous technologies, such as the diffusion of robots and computers. Additionally, we also observe a significant drop in prices in the services to firms sector during this period, which suggests that the volume of sales has actually increased over the period.

III presents the preferred event study graphs at the city level and at the establishment level (in blue), along with the robustness tests controlling for the propensity to adopt earlier interacted with year fixed effects (in red). Cities 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 specialized and employed fewer types of occupations in-house. At the same time, establishments increasingly bought these services from other firms. This has resulted in workers becoming increasingly segregated into firms that primarily hire their type of specialists. The establishment level results on outsourcing expenditures follow similar patterns but are smaller in magnitude. Hence, some of the effect measured may operate through composition: newly created firms spend more on outsourcing relative to disappearing firms. On the other hand, the establishment-level results on occupational segregation are of a similar magnitude to the city-level ones. Results remain unchanged when we control for the propensity of cities to adopt BI early, supporting the interpretation that the effect is driven by BI itself and not by differential trends explained by pre-existing characteristics.

Table A1 in Appendix A quantifies the effect at the city and establishment levels by presenting the dynamic post-BI coefficients.[36] Five years after the arrival of BI in the city, we observe an average increase in the outsourcing expenditures over wage bill of about 0.016, which corresponds to a growth of 6% with respect to the baseline levels, but it is marginally not significant. Occupational concentration increases by about 0.006 (1.6% growth relative to baseline) and the effect is highly significant. The effects at the establishment level are smaller but more precise on outsourcing expenditures: the latter increases by 0.014 (+4%), while occupation concentration increases by 0.007 (+1.9%) within existing establishments and the post-treatment coefficients

[36]The number of observations is smaller in the regressions on outsourcing expenditures because the sample is restricted to firms with more than 750,000 euros of turnover, and because the data only reports the headquarters location for multi-establishment firms.

FIGURE III. OUTSOURCING INTENSITY AND OCCUPATIONAL CONCENTRATION WITHIN ESTABLISHMENTS



Notes: This figure shows regression coefficients and the 90% and 95% confidence intervals from a stacked event study design. The city-level specification follows Equation (3), while the establishment-level specification follows the same logic but replaces the city fixed effects with firm fixed effects. The blue lines present our baseline model, while the red lines present the model controlling for the propensity score of early adoption interacted with the year fixed effects.

remain strongly significant.

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.[37] Figure A5 in Appendix A describes the evolution of these outcomes over time. Employment in high-skill outsourcing services increased

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

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. Specifically, 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%). In what follows, we test whether these substantial changes were, at least in part, facilitated by the diffusion of the ADSL technology across France.

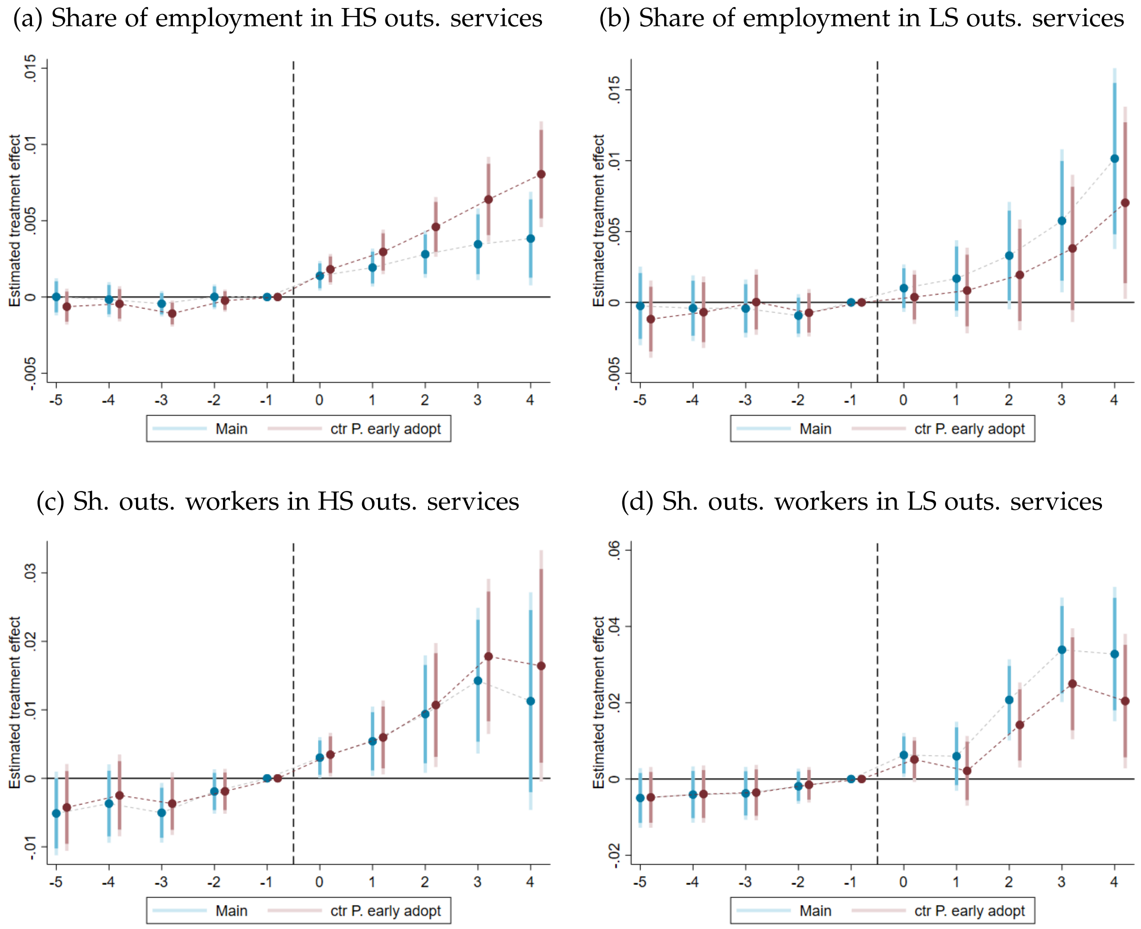
Figure IV reports the event study graphs of the two main outcomes separately for high- and low-skill outsourcing, again showing the main model and the robustness to controlling for the probability of early adoption. Table A2 in Appendix A shows the corresponding regression coefficients, while Tables A3 and A4 show the coefficients obtained from breaking down each industry group even further into its sub-categories, underlining that similar patterns are observed for most of the sub-groups.[38]

Once again, the pre-trends are parallel and the effect materializes after the BI connection, regardless of whether we control for the propensity score. While the magnitude of the coefficients might appear small, the effect is actually far from trivial when compared to the baseline (pre-treatment) values.[39] The average share of city employment accounted for by high-skill outsourcing services was about 1.7% at the beginning of

[38]The regressions on the share of outsourceable occupations within outsourcing sectors have smaller and varying number of observations because they are only defined in cities where there is at least one firm active in outsourcing services, contrary to the regressions on the share of employment in outsourcing services, which is defined everywhere.

[39]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).

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



Notes: This figure shows regression coefficients and 90% and 95% confidence intervals from a stacked 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 4(a) and 4(b)), and the share of outsourceable high-skill (low-skill) workers employed in their respective services in a city at t (Figures 4(c) and 4(d)). The model is presented in Equation (3). The blue lines present our baseline model, while the red lines present the model controlling for the propensity score of early adoption interacted with year fixed effects.

the period, such that the average impact of BI in the five years after its arrival amounts to a growth rate of 16%. Low-skill outsourcing services accounted for about 10% of total city employment in 1997, and the BI connection led to a growth rate of 4.6%. When it comes to the concentration of workers in outsourceable occupations within their respective service sectors, the baseline value is 6.5% for high-skill employees, and the effect of BI amounts to 1 additional percentage point (growth of 16%). For low-skill employees, the baseline value is 18% and the effect of BI 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, especially for non-core activities situated both at the low-skill and high-skill ends of the spectrum.

Robustness Tests Figures A7 and A8 in the Appendix show the robustness tests that we performed on the outsourcing outcomes at the city level, and Figures A10 and A11 in the Appendix show the same robustness tests on the outcomes at the firm level. In Figures A7 and A10 we test whether the results hold when adding additional controls: i) pre-BI productivity growth in the city interacted with year dummies, ii) sectoral composition in each city prior to 1999 interacted with year dummies, iii) share of left-wing voters in the city in 1995 and an indicator of cities switching the political majority between 1995 and 2002, both interacted with year dummies, iv) city size prior to 1999, measured as the number of establishments active in the city, interacted with year dummies, and v) all of the previous controls added together. The fact that the event study graphs remain highly similar for most outcomes suggests that differences in the timing of BI diffusion, across cities of similar density within the same department are not correlated with preexisting differences in city-level productivity growth, with differences in sectoral composition, with differences and changes in political affiliation, or with differences in city size. Figures A8 and A11 in the Appendix show the results obtained i) after controlling for the BI coverage of other cities in the same department - to get a sense of the possible spill-over effects of BI arrival on control cities; ii) after introducing commuting zone $\times$ year fixed effects instead of department $\times$ year, which should magnify the bias in the presence of strong geographic spill-overs; iii) from using the continuous measure of coverage $\tilde{Z}_{it}$ instead of the binary treatment based on the first year of BI arrival; and iv) from running a standard dynamic two-way fixed effects model, as reported in Equation (2). Once again, most of the outcomes remain unchanged, except that outsourcing expenditures becomes flat in the standard staggered event study model. This change suggests that dynamic effects in outsourcing expenditures bias downward the estimates in a model that uses post-treatment observations as controls.

Thirdly, Figures A9 and A12 present the results obtained from placebo tests, where late receivers are split into two groups according to their propensity of early adoption, and the pseudo-treatment is evaluated for the period preceding their actual BI connection. The graphs show flat and non-significant differences across the two groups except for one outcome: the share of employment in low-skill outsourcing services. This presents a positive trend among cities with higher likelihood to be treated early. We conclude that the causal impact of BI on this particular outcome necessitates a cautious interpretation. However, the lack of trends observed in the share of low-skill workers employed in low-skill services confirms that BI did generate a shift in distribution of outsourceable occupations across sectors, taking the overall size as given.

Fourthly, Table A7 reports the coefficients obtained from static regressions where the treatment status is interacted with the post-BI period, and Table A8 shows the coefficients from a standard staggered regression run at the commuting zone level, where the post-BI dummies are interacted with the continuous measure of BI coverage in the area. The magnitude of the static coefficients is slightly smaller than the average of the dynamic coefficients but remains in line with the main results. The effect measured at the employment zone level is smaller and less significant, which is explained by the fact that there is much less variation in the timing of BI arrival at this more aggregate level. Nonetheless, the coefficients qualitatively confirm those reported in the main analysis.

Finally, to further validate the stacked design, we estimate an alternative model using "pure controls" and we perform additional pre-trends tests as suggested by Borusyak et al. (2021).[40] For the first, we restrict the estimation period to only include years up to 2004 and then use cities treated after 2004 as pure controls, using both the stacked design and the standard TWFE design. Results are presented in Appendix figure

[40]We present these tests only for the city-level estimates for the sake of conciseness.

A13 and show very similar coefficients to our baseline estimates. For the second, we re-estimate both the dynamic stacked model and the dynamic TWFE model on the untreated sample - i.e. on the years in which a given city is not yet connected to BI -, to further validate the robustness of the common trend assumption. Because of multicollinearity, in the TWFE model we normalize both $t-4$ and $t-1$ to 0, while we only normalize $t-1$ to 0 in the stack model. Results are presented in Appendix figure A14. We cannot reject the fact that the pre-trend coefficients are jointly equal to 0.

Interpretation of effects To get a more concrete sense of the magnitude of these results, we can wonder how outsourcing would have evolved in the absence of BI.

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 of BI. We follow the procedure adopted in Malgouyres et al. (2021) and we compute the predicted outcome as the actual outcome minus the dynamic effects predicted by our semi-dynamic specification reported in Table A1.

In this exercise, we partition cities into groups that receive broadband in the same year. We refer to these groups as cohorts and denote the first year of BI accessibility as t_0. The contribution of each cohort t_0 to the average effect at time t is the product of two terms. The first term is the weight of the cohort the year prior to its first connection (denoted $w_{t0,t0-1}$), which is computed as its share of total employment in that year. The second term is the average effect to which the cohort is exposed at time t. This effect will depend on the time since treatment $d = t - t_0$. We denote this effect as $\widehat{\alpha}_{t-t_0}$, which corresponds to the coefficient from our estimate of equation (3). We can then obtain the average effect of broadband internet expansion for each year t by summing across cohorts:

$$\overline{a}_t = \sum_{t_0=1999}^{2007} w_{t0,t0-1} \hat{\alpha}_{t-t_0}.$$

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, summed to the predicted effect of BI: $y_t = y_t(0) + \bar{a}_t$. We then obtain the trends in the main outcomes in the absence of our estimated BI effect by inverting this relationship: $y_t(0) = y_t - \bar{a}_t$. This exercise is not a proper counterfactual analysis since it ignores 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.

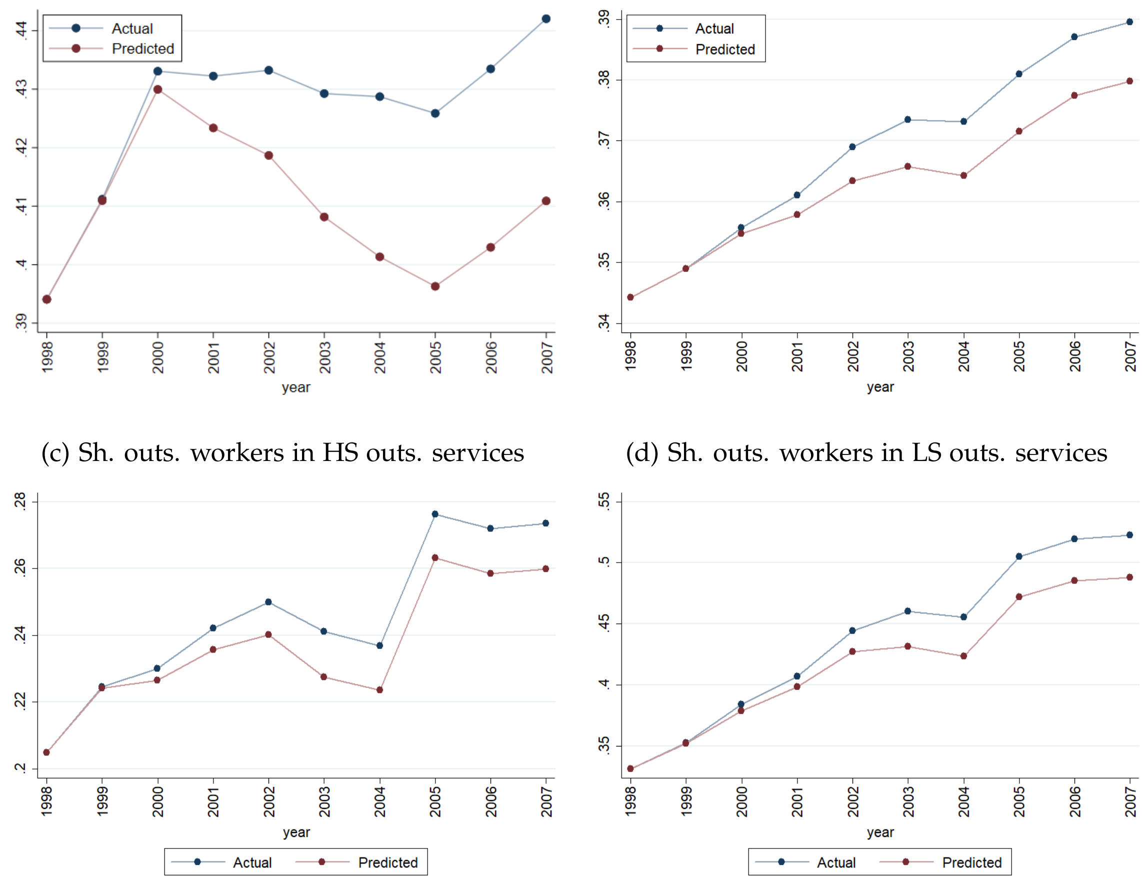
Figure V shows that the outsourcing expenditures over wage bill decreases after the beginning of the 2000s if we subtract the estimated BI effect, and by the end of the period it is 9% smaller than what is actually observed. The occupational segregation across firms and the share of high- and low-skill outsourceable occupations in outsourcing sectors increase less steeply over the period when we subtract the estimated BI effect, reaching in 2007 levels that are 2.5%, 4.9%, and 6.6% smaller, respectively, than what is observed in the data.

4.2 At the individual level

In this section, we utilize our individual panel data to track workers over time and examine their job mobility choices. Specifically, we investigate the impact of BI expansion on the likelihood that workers change jobs. We take into account three types of job mobility: any establishment change, a transition from an establishment outside the outsourcing sector to one within the outsourcing sector, and a shift between two establishments outside the outsourcing sector.

In contrast to the previous model, workers are able to relocate from their current city over the years, which could result in them being treated multiple times. To prevent this, we assign a unique city to each individual i, representing the location where they first experience broadband internet access (during a year denoted as $t_{0,i}$). We also use this year to define the skill of the workers based on their occupation at $t_{0,i}$.

FIGURE V. COMPUTATION OF THE PREDICTED AGGREGATE TRENDS AFTER SUBTRACTING OUR ESTIMATED BI EFFECT



(a) Outsourcing expenditures

(b) Occupation concentration

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

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

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 for each cohort of firms (i.e., all firms for which broadband expansion occurred in the same year), measured in the year of broadband expansion. The weights are normalized so that they sum to one.

We then estimate the following linear probability model:

$$Move_{i,t} = \alpha \mathbb{1}\left\{t \geq t_{c(t_{0,i})}\right\} + \beta Z_{i,t} + \gamma_t \boldsymbol{X}_{\boldsymbol{c(t_{0,i})}} + \varepsilon_{i,t}, \tag{4}$$

where $c(t_{0,i})$ denotes the city in which the individual i worked at $t_{0,i}$ and $X_{c(t_{0,i})}$ is a vector of characteristics specific to the city that includes department fixed effects and the logarithm of density (both interacted with year fixed effects γ_t). Finally, $Z_{i,t}$ denotes individual characteristics: age, age squared, sex, and a dummy for working part-time (as opposed to full-time). The dependent variable $Move_{i,t}$ is a binary variable that equals 1 if the worker i has moved in year t and 0 otherwise. ε is an idiosyncratic error that we assume can be correlated within departments but not across them. α thus captures the effect of being connected to BI on the probability of moving, controlling for observable and time-unvarying unobservable worker characteristics. We augment this model by adding successively city fixed effects $c(t_0, i)$ and individual workers fixed effects. In these more saturated models, the identification is fully driven by the timing of BI expansion.

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
Panel A						
Connected	0.046***	0.006***	0.040***	0.052***	0.002***	0.050***
	(0.010)	(0.002)	(0.008)	(0.003)	(0.000)	(0.003)
Relative	**0.141**	**0.176**	**0.137**	**0.214**	**0.320**	**0.211**
Panel B						
Connected	0.009*	0.002*	0.007	0.023***	0.002***	0.021***
	(0.005)	(0.001)	(0.004)	(0.003)	(0.000)	(0.002)
Relative	**0.028**	**0.075**	**0.022**	**0.095**	**0.243**	**0.090**
Panel C						
Connected	0.011**	0.003**	0.008***	0.015***	0.001***	0.013***
	(0.005)	(0.001)	(0.004)	(0.003)	(0.000)	(0.002)
Relative	**0.034**	**0.080**	**0.029**	**0.060**	**0.177**	**0.056**
Average pre-treatment	0.325	0.032	0.293	0.242	0.007	0.235
Obs.	571,720	571,720	571,720	2,666,453	2,666,453	2,666,453

Notes: This Table presents the point estimate and standard errors of coefficient α in equation (4) as the ratio of α divided by the average value of the dependent variable pre-treatment (**Relative**). The dependent variable is a binary variable equal to 1 if worker i has changed establishment between t and $t+2$. Panel A includes department-year fixed effects, the logarithm of the density of the city interacted with a time dummy and individual controls: gender, age, age squared and a full-time dummy. Panel B adds a city fixed effect and panel C a city and individual fixed effects. 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). 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 department level and robust to heteroskedasticity. Time period: 1995-2008. ***, ** and * respectively denote significance at the 1, 5 and 10% level.

Table III presents our results. We report coefficient α from equation (4) for different types of mobility. Columns (1) to (3) refer only to workers in high-skilled occupations. In Column (1), the dependent variable is equal to 1 in the case of 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 workers in low-skill occupations, and for the corresponding outsourcing sectors (cleaning, driving, security, etc.). Panel A does not include additional fixed effects on top of what is presented in Equation (4), Panel B includes city $c(t_0, i)$ fixed effects, and Panel C adds individual fixed effects. Because the average value of the dependent variable is very different across skills and types of mobility, in the table we also report the baseline average and the computed relative effect of BI in% terms (the line *Relative* in each panel), defined as the ratio of α over the pre-treatment average.

These results show that being connected to BI is associated with a greater propensity to move, both for high-skill and low-skill workers, with a larger relative effect for mobility to an outsourcing sector. Relatively to the pre-treatment average, and conditional on all covariates, being connected to BI induces a 0.141 percentage-point increase in the propensity to switch firms for workers in high-skill occupations, and a 0.214 percentage-point increase for workers in low-skill occupations. This relative change is larger for moves towards outsourcing firms (Columns 2 and 5), whether or not additional fixed effects are added.

Part of this mobility includes moves across cities, which could suggest that control cities in the specifications presented in Section 4.1 might be partly affected by the treatment through an increase in labor mobility. We argue that this effect is one order of magnitude smaller than the main outcomes of interest at the city level, and thus can be considered negligible. New entrant workers in low-skill occupations represent on average 0.5% of the overall employment of receiving cities and 1.3% of the outsourceable employment. New entrant workers in high-skill occupations represent on average 3.7% of the overall employment of receiving cities and 11% of outsourceable employment. In addition, the majority of the inter-city moves of both high-skill and low-skill workers are directed towards cities that are already connected to BI starting from 2001, and thus do not affect the controls (see Appendix Table A5).

Note that this model is different from the one used in Section 4.1 and therefore the identification discussed previously does not apply in the same way here. The advantage of using individual data is that it allows us to focus on the intensive margins of the effect and to look in more detail at the different types of mobility and how they are impacted by the arrival of BI. In Appendix Figure A6 we aggregate the data at the city level, using $c(t_0, i)$ for all individuals i, and we calculate the share of workers in high-skill and low-skill occupations that move to the outsourcing sector every year. We can then estimate the same model as in Section 4. To account for the fact that this model is aggregated from individual level data, we weight the city by population. These results support the causal interpretation of our individual-level regressions, namely that BI increases the mobility of workers from their current firms to firms specialized in outsourcing services.

5 Wage dynamics around individual transitions to outsourcing

Our findings suggest that BI expansion increases the mobility of workers across establishments, with a particularly strong effect on the mobility of workers in outsourceable occupations towards firms in the outsourcing sector. This reallocation of workers across firms might be another mechanism through which technological change affects the wages of low- and high-skill workers differently, on top of the skill-biased effect on the production function shown by Akerman et al. (2015) in Norway. In our context, we find broadly comparable skill-biased effects of broadband expansion on wages to those shown by Akerman et al. (2015)—results are presented in Appendix D. In this section, in order to assess more directly the relationship between mobility to outsourcing on wages, we carry out a series of simple event studies examining wage dynamics around *individual outsourcing events*, again distinguishing between high- and low-skill occupations. 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 both high- and low-skill workers. We thus look in greater detail at the wage dynamics of workers around these events, which helps indicate the working conditions in the outsourcing sectors compared to those elsewhere.

Table A6 in Appendix A describes the average wage and employment of outsourceable occupations across sectors. IT specialists and consultants earn a gross hourly wage of about 30€ in their outsourcing services, which is slightly higher than their counterparts in other services and in manufacturing (25€ and 28€, respectively). Security workers, cleaners, drivers, and logistic workers, on the other hand, earn less per hour 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 estimation the association between outsourcing and 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 experienced one and only one outsourcing event over the period of observation, which allows us to define our event study based on the timing of the move. We then look at the evolution of their hourly wages by estimating the following dynamic model:

$$\log(w_{it}^{ow}) = \sum_{\tau=-v}^{v'} \alpha_\tau \mathbb{1}\left\{t = t_i + \tau\right\} + X\gamma + \psi_{d,t} + \nu_i + \varepsilon_{i,t} \tag{5}$$

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. $\psi_{d,t}$ and ν_i are a set of department d times year t fixed effects, and individual fixed effects.[41] ε is an idiosyncratic error that we assume can be correlated within departments but

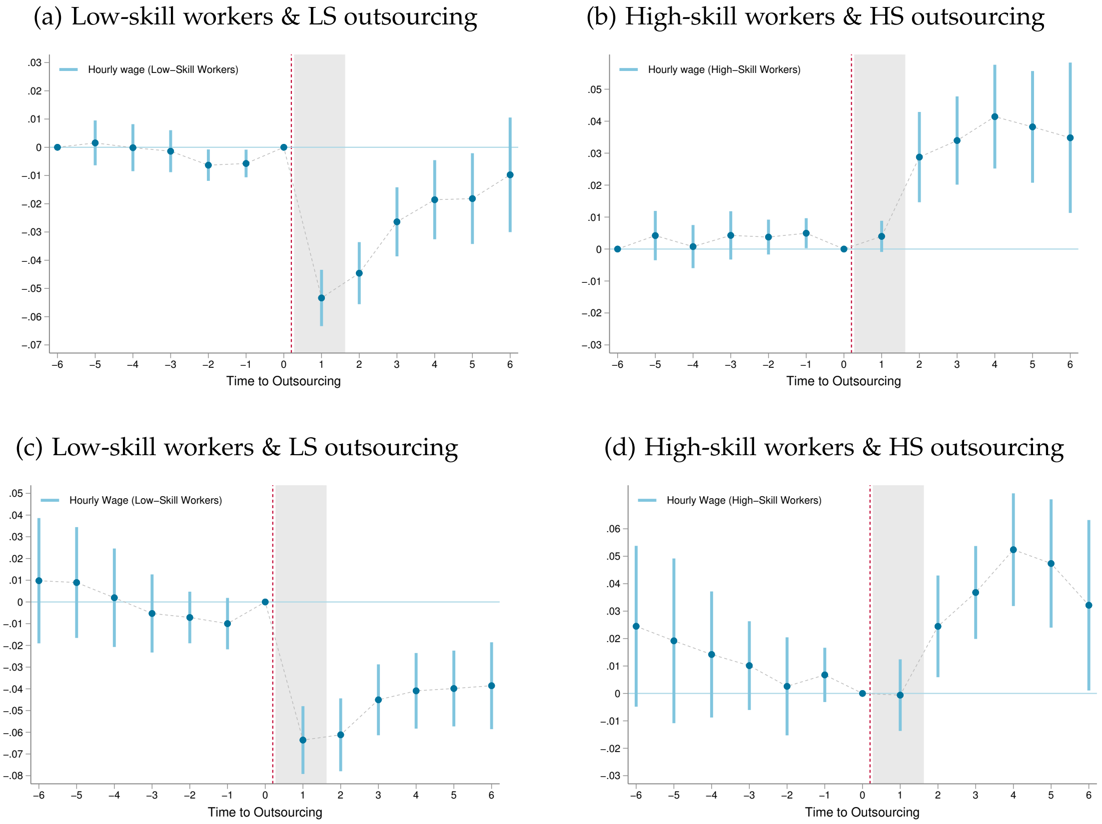
[41] Adding sector $s(i)$ fixed effects does not affect our results.

not across them. We estimate this regression separately for low-skill workers and high-skill workers, with respective samples of 7,995 and 10,404 different workers. We consider $v = v' = 6$, which means that we follow workers over 13 years. Finally, given that in this setting, and similar to the model presented in Equation (2), 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 is exogenous from the point of view of workers' wage trajectories, after controlling for observable characteristics and fixed effects. Still, the absence of pre-trends is indicative of the comparability of wages before workers' mobility to a service firm and accordingly, we believe the changes in wages around the events are interesting from a descriptive standpoint. Results are presented in Figures 6(a) and 6(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 documented by Goldschmidt and Schmieder (2017), who explain this phenomenon by the fact that firms in the outsourcing sectors benefit from lower rents, on average, than other companies. This translates into a lower wage premium for their employees.[42] The average decline 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 wages 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 communication technologies improve. By regrouping these occupations, specialized firms can serve

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

FIGURE VI. Wages of outsourced workers before and after the outsourcing event



Notes: This figure shows regression coefficients and confidence intervals at 95% from a dynamic 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. It also includes several control variables: age, age squared, a part-time dummy, and a set of year times department and individual fixed effects. The shaded area denotes the year of the mobility, which is associated with noisy measures of work duration. The top panels are restricted to workers with one and only one instance of mobility from a non-outsourcing firm to an outsourcing firm. The bottom panels add control groups composed of workers who move from any firm to a non-outsourcing firm. The left-hand side panels are restricted to low-skill workers (39,659 workers in Panel (a) and 173,788 in Panel (c)), while the right-hand side is restricted 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 department level.

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.[43]

To disentangle the specific effect of moving towards the outsourcing sector relative to other types of mobility, in Figures 6(c) and 6(d) we reproduce the same exercise but using the set of workers who transition to a different firm that does not belong to one of the outsourcing sectors as a control group. In this case, the coefficients should be interpreted as the difference in hourly wages compared to workers of a 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-skill 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 widespread diffusion of high-speed internet has fostered many changes in how businesses operate, some of which differ from those observed in previous waves of technological change. In this paper, we examine the role that broadband internet played in incentivizing French firms to outsource some non-core activities, both in the high-skill and low-skill segment, and study its effects on affected workers. We leverage the staggered roll-out of broadband internet across the French territory to adopt an event study design and estimate the causal effect associated with the diffusion of this new technology. This approach compares similarly dense cities within a given department, which gained 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

[43]Higher wages for high-skill workers could also indicate compensation for greater job security risks. Companies may regularly require cleaning or driving services, but only hire consultants as needed, leading to higher wages for the latter due to their inconsistent work hours.

and high-skill occupations. Finally, we provide evidence that high-skill workers experience wage gains through outsourcing, while low-skill workers experience wage losses. These findings are suggestive that the skill-biased impact of broadband internet technology is partly driven by its role in facilitating domestic outsourcing of both low- and high-skilled workers, which has important implications in terms of labor market inequality.

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.

_ , _ , _ , Peter J. Klenow, and Huiyu Li, "Good Rents versus Bad Rents: R&D Misallocation and Growth," 2023. mimeo Stanford.

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.

Andrews, Dan, Giuseppe Nicoletti, and Christina Timiliotis, "Going digital: What determines technology diffusion among firms?," Technical Report 2018.

Arcep, "Etude sur les équipements et usages des PME et ETI," 2016. Arcep.

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

Bartel, Ann P, Saul Lach, and Nachum Sicherman, "Outsourcing and technological change," 2005.

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.

Bertschek, Irene, Daniel Cerquera, and Gordon J Klein, "More bits–more bucks? Measuring the impact of broadband internet on firm performance," *Information Economics and Policy*, 2013, *25* (3), 190–203.

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.

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

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.

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, Pascuel Plotkin, and Benjamin Schoefer, "Paying outsourced labor: Direct evidence from linked temp agency-worker-client data," *The Review of Economics and Statistics*, 2020, pp. 1–28.

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.

Fabling, Richard and Arthur Grimes, "Picking up speed: Does ultrafast broadband increase firm productivity?," *Information Economics and Policy*, 2021, *57*, 100937.

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

Gardner, John, "Two-stage differences in differences," *arXiv preprint arXiv:2207.05943*, 2022.

Garicano, Luis 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.

Grimes, Arthur, Cleo Ren, and Philip Stevens, "The need for speed: impacts of internet connectivity on firm productivity," *Journal of Productivity Analysis*, 2012, *37* (2), 187–201.

Haller, Stefanie A and Sean Lyons, "Broadband adoption and firm productivity: Evidence from Irish manufacturing firms," *Telecommunications Policy*, 2015, *39* (1), 1–13.

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.

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.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard, "Information technology and returns to scale," *Available at SSRN 3458604*, 2022.

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," *Journal of International Economics*, 2021, *133*, 103520.

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.

Tadelis, Steven, "Complexity, flexibility, and the make-or-buy decision," *American Economic Review*, 2002, *92* (2), 433–437.

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 simple model

A Additional Results

A.1 Tables

TABLE A1. Effect of ADSL on outsourcing expenditure and occupational sorting

	(1)	(2)	(3)	(4)
	Outsourcing / wage bill		Occupational concentration (wage bill HHI)	
VARIABLES	City level	Estab. Level	City level	Estab. Level
T = 0	0.00304	0.00712	0.00128	0.00157***
	(0.00601)	(0.00446)	(0.00118)	(0.000471)
T = +1	0.0102	0.0102*	0.00418**	0.00318***
	(0.00855)	(0.00604)	(0.00162)	(0.000797)
T = +2	0.0157	0.0123	0.00693***	0.00736***
	(0.0111)	(0.00767)	(0.00196)	(0.00104)
T = +3	0.0233	0.0184*	0.00906***	0.0117***
	(0.0144)	(0.0107)	(0.00273)	(0.00145)
T = +4	0.0296	0.0210	0.00981***	0.0128***
	(0.0185)	(0.0141)	(0.00321)	(0.00168)
Average effect	0.0164	0.0138*	0.00625***	0.00733***
	(0.0109)	(0.00804)	(0.00202)	(0.000988)
Baseline mean	0.27	0.36	0.37	0.37
	(0.57)	(1.09)	(0.157)	(0.180)
Observations	293,463	1,335,364	423,770	3,077,125
R-squared	0.682	0.787	0.779	0.846

Notes: ***, ** and * respectively denote significance at the 1, 5 and 10% level. Columns (1) and (3) run the regression at the city level, following equation 3, 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 × 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 A2. 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)
Baseline mean	0.017	0.065	0.103	0.184
	(0.087)	(0.193)	(0.240)	(0.336)
Observations	423,770	164,880	423,770	188,496
R-squared	0.821	0.727	0.902	0.798

Notes: ***, ** and * respectively denote significance at the 1, 5 and 10% level. The regressions are run at the city level following equation 3. All columns control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

TABLE A3. 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)
Baseline mean	0.004	0.013	0.004	0.010	0.082	0.007
	(0.041)	(0.076)	(0.044)	(0.073)	(0.222)	(0.056)
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: ***, ** and * respectively denote significance at the 1, 5 and 10% level. The regressions are run at the city level following equation 3. All columns control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

TABLE A4. 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)
Baseline mean	0.088	0.047	0.143	0.094	0.193	0.033
	(0.244)	(0.161)	(0.328)	(0.272)	(0.360)	(0.152)
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: ***, ** and * respectively denote significance at the 1, 5 and 10% level. The regressions are run at the city level following equation 3. All columns control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

TABLE A5. Descriptive statistics of the share of mobilities going to connected cities

	High Skill workers		Low Skill workers	
Year	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 A6. Average wage and employment in outsourceable occupations across sectors

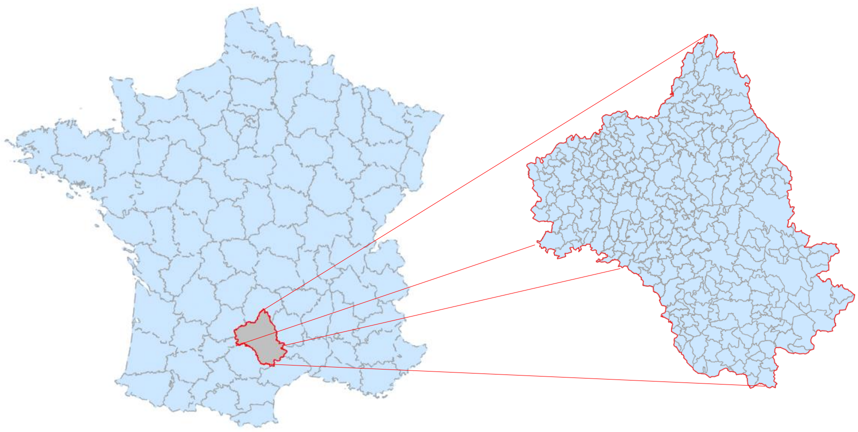
		outsourcing services	Other services	Manufacturing
		mean/(sd)	mean/(sd)	mean/(sd)
Gross hourly wage workers in HS outsourceable occup. (2010 euros)	overall	30.7	24.8	28.6
		(13.2)	(11.9)	(12.2)
	pre-BI	29.6	22.6	27.1
		(11.7)	(10.7)	(9.2)
	post-BI	31.1	26.0	29.8
		(13.6)	(12.3)	(13.9)
N. of workers in 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.

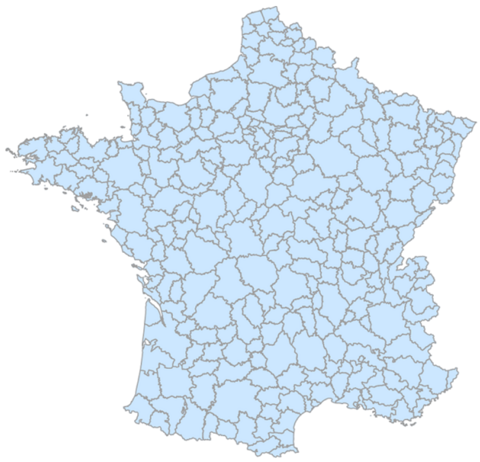
A.2 Figures

FIGURE A1. Departments, commuting zones and cities in France

(a) Departments and cities

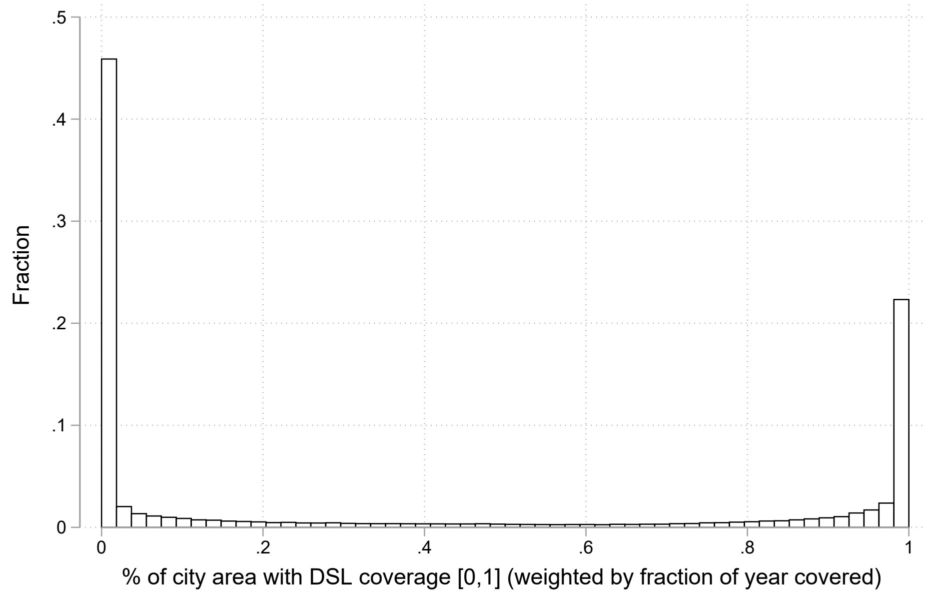


(b) Commuting Zones



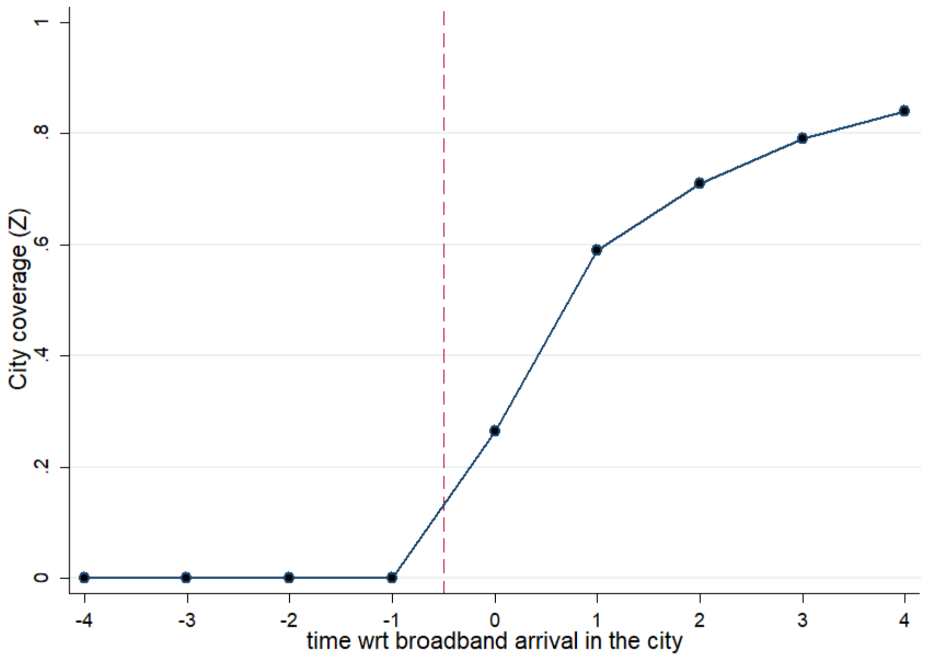
Notes: The first map represents the distribution of departments ("départements") in France and takes the example of Aveyron (department number 12) to show the distribution in cities. The second map represents the distribution of Commuting Zones ("Zones d'emploi").

FIGURE A2. 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 (1). We see that while the measure is continuous and contained between 0 and 1 but presents point of accumulation on 0 and 1.

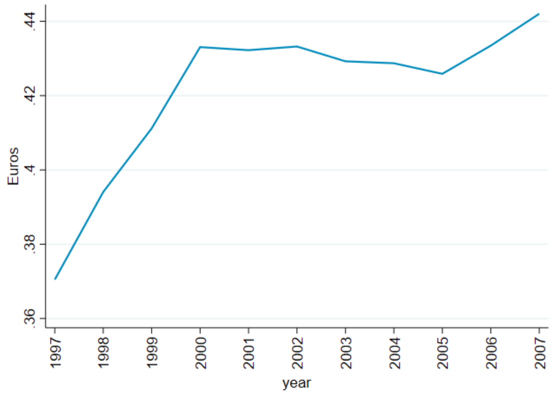
FIGURE A3. Evolution of $\widetilde{Z}_{it}$ before and after the discrete event



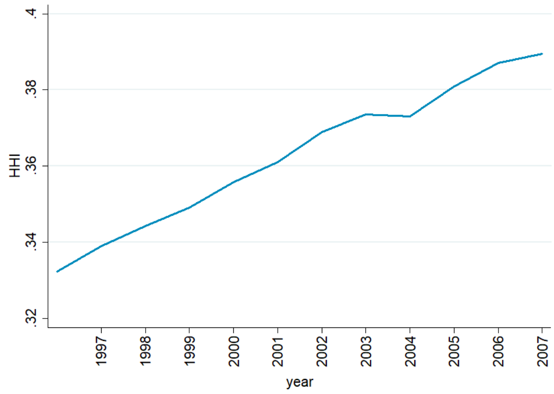
Notes: This figure plots the average of the continuous measure of local broadband availability (variable $\widetilde{Z}$) along the time to event variable, where event is the first year where $\widetilde{Z} > 0$.

FIGURE A4. Overall trends in outsourcing expenditure and occupation concentration

(a) Outsourcing expenditure over wage bill



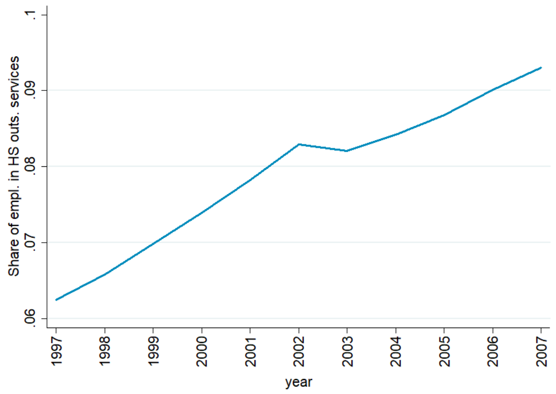
(b) Occup. concentration (HHI over wage bill)



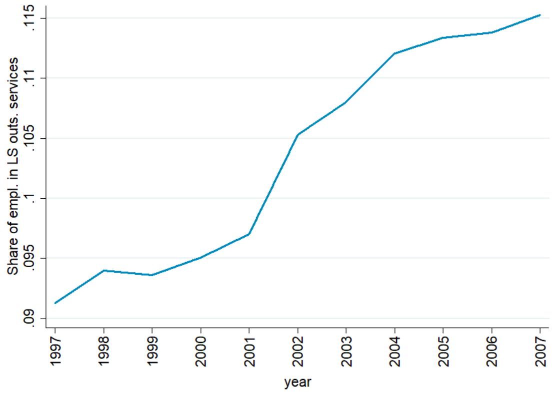
Notes: This Figure shows the evolution over time of the main outcomes of interest in the outsourcing analysis.

FIGURE A5. Overall trends in high and low skill outsourcing

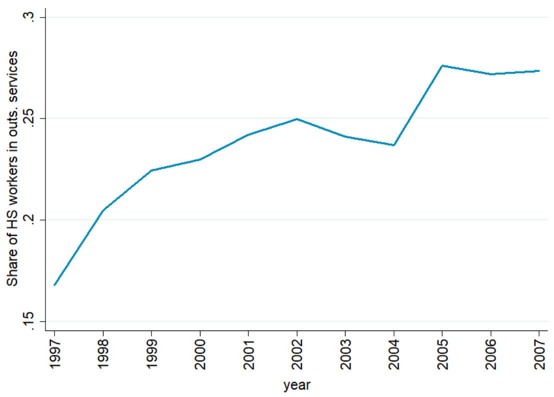
(a) Share of employment in HS outs. services



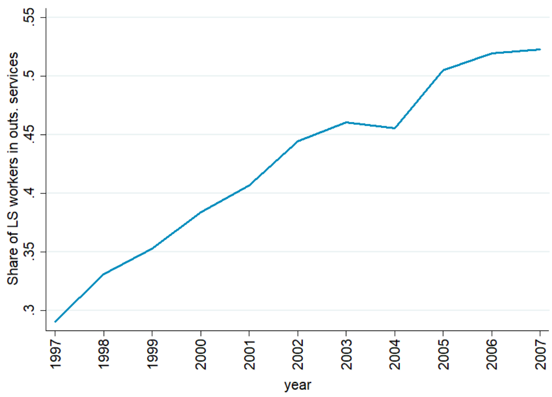
(b) Share of employment in LS outs. services



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



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

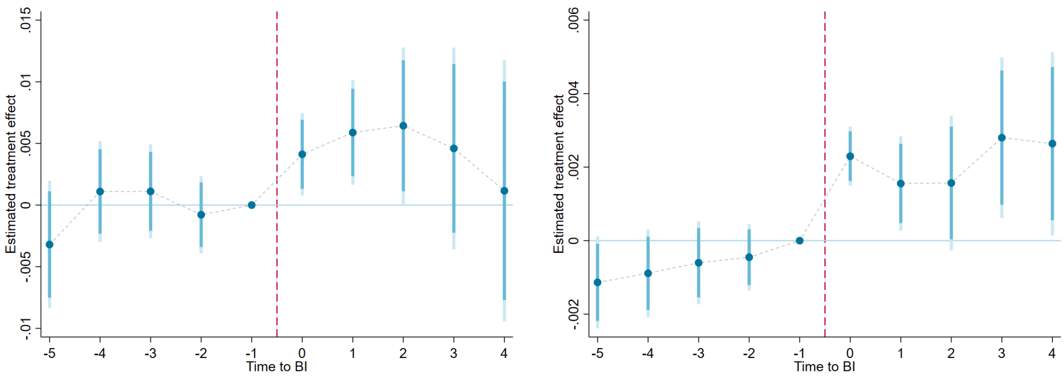


Notes: This Figure shows the evolution over time of the main outcomes of interest in the outsourcing analysis.

FIGURE A6. Share of workers moving to outsourcing firms at the city level

(a) High Skill Workers

(b) Low Skill Workers



Notes: This Figure shows regression coefficients and 90% and 95% confidence intervals from a dynamic event study where the dependent variable is the share of workers moving to an outsourcing firms within a city (respectively for high skill and low skill segments) at *t* and the specification follows equation (3). Observations are weighted by population in 1990.

A.3 Identification

TABLE A7. 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	0.00867	0.00336**	0.00200***	0.00707***	0.00201	0.0103**
	(0.00744)	(0.00144)	(0.000573)	(0.00250)	(0.00131)	(0.00408)
Observations	293,463	423,770	423,770	164,880	423,770	188,496
R-squared	0.682	0.779	0.821	0.727	0.902	0.798
Panel B : establishment level regressions						
Post ADSL * treated	0.00925*	0.00284***	-	-	-	-
	(0.00519)	(0.000609)	-	-	-	-
Observations	1,335,364	3,077,125	-	-	-	-
R-squared	0.787	0.846	-	-	-	-

Notes: The regressions are run at the city and establishment level following a model similar to equation 3, but where instead of including the dynamic post-BI effects for every year, we just include a dummy for post-BI period interacted with the treatment indicator. All columns control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. ***, ** and * respectively denote significance at the 1, 5 and 10% level.

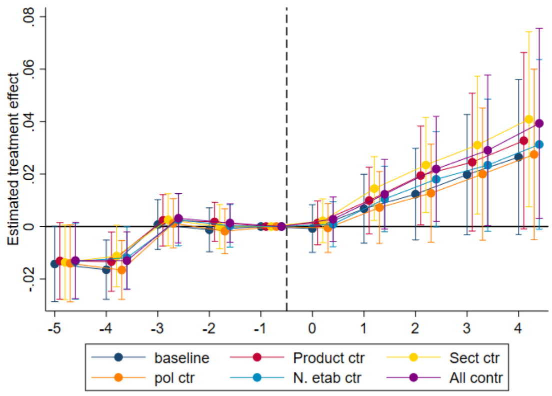
TABLE A8. Effect of ADSL on outsourcing outcomes - regressions at the labor market area level

	(1) Outsourcing / wage bill	(2) Occup. concentration (wage bill HHI)	(3) Sh. of empl. in HS outs. services	(4) Sh. Outs. workers in HS outs. services	(5) Sh. of empl. in LS outs. services	(6) Sh. outs. workers in LS outs. services
T = 0 x Zemp coverage	0.00183	4.02e-05	0.000825	0.00283	0.00381	0.0261
	(0.0350)	(0.00395)	(0.00188)	(0.0239)	(0.00330)	(0.0273)
T = +1 x Zemp coverage	0.0304	-0.00154	0.00102	0.0115	-0.00128	0.000544
	(0.0213)	(00297)	(0.00154)	(0.0160)	(0.00235)	(0.0235)
T = +2 x Zemp coverage	0.0557***	0.00315	0.00237	0.00846	0.000151	0.0461
	(0.0210)	(0.00387)	(0.00246)	(0.0209)	(0.00366)	(0.0345)
T = +3 x Zemp coverage	0.0487**	0.00923**	0.00534*	0.0358	-0.000332	0.141***
	(0.0223)	(0.00463)	(0.00302)	(0.0277)	(0.00467)	(0.0461)
T = +4 x Zemp coverage	0.0453	0.0161***	0.00788**	0.0468	-0.000565	0.199***
	(0.0297)	(0.00527)	(0.00370)	(0.0344)	(0.00653)	(0.0490)
Average effect	0.0364*	0.0054	0.00349	0.0211	0.000356	0.0825***
	(0.0202)	(0.00363)	(0.0022)	(0.0190)	(0.00356)	(0.0291)
Observations	2,880	2,783	2,783	2,783	2,783	2,783
R-squared	0.815	0.891	0.974	0.809	0.952	0.810

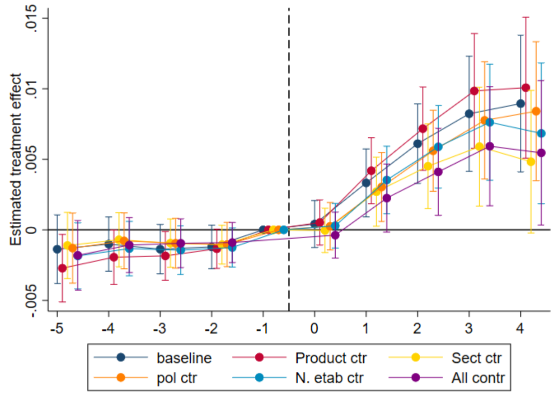
Notes: The regressions are run at the labor market area (*Zone d'emploi*) level. Given that there is little variation in the timing of first BI appearance in labor market areas within the same department, we take advantage of the continuous measure of BI coverage: we estimate a standard staggered model similar to equation 2, but where we interact the dynamic post-BI dummies for every year with the share of the labor market area that is covered in that period. All columns control for the population density in 1999 interacted with year dummies, department × year fixed effects, and labor market area fixed effects. ***, ** and * respectively denote significance at the 1, 5 and 10% level.

FIGURE A7. Robustness of city-level results: adding additional controls

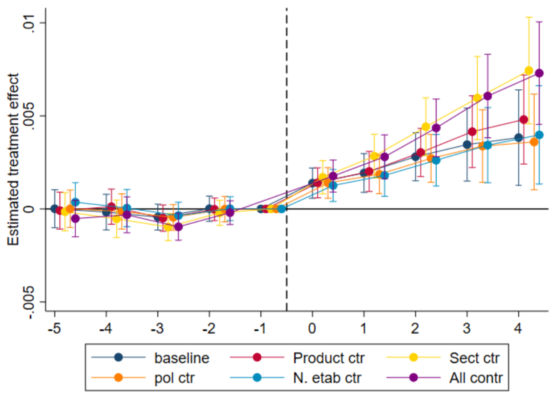
(a) Outsourcing/wage bill



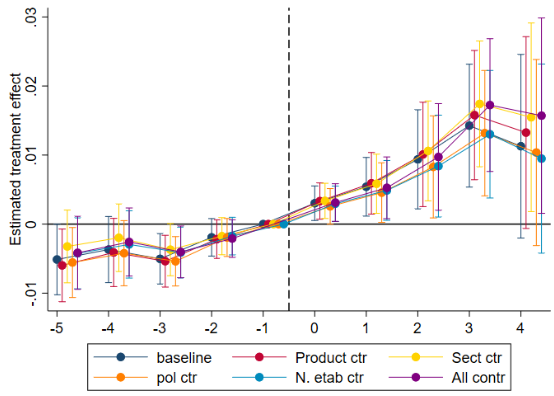
(b) Occup. concentration (HHI)



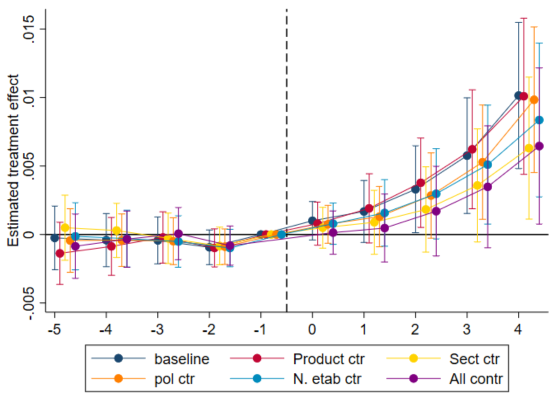
(c) Sh. of empl. in HS outs. services



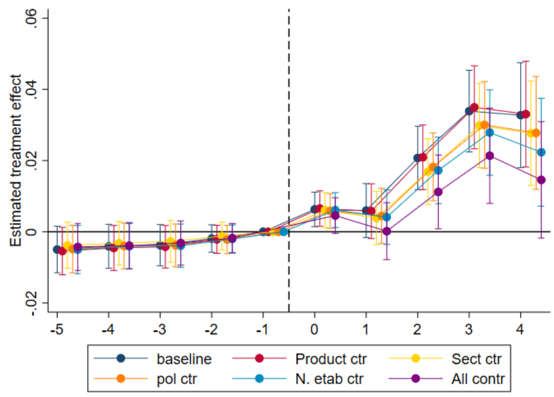
(d) Sh. outs. workers in HS outs. services



(e) Sh. of empl. in LS outs. services



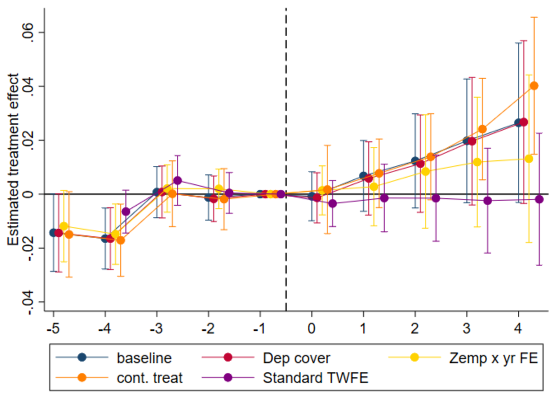
(f) Sh. outs. workers in LS outs. services



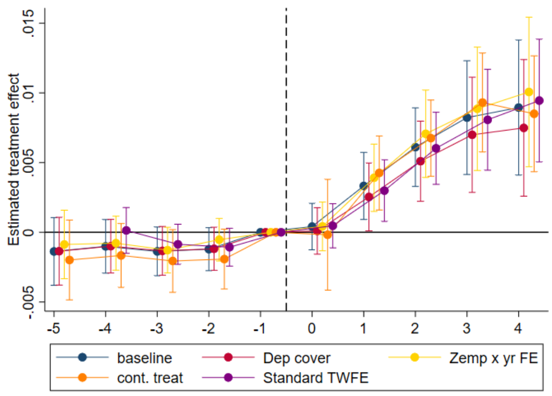
Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating equation (3) on city-level data with different sets of controls. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level. The navy blue line shows the baseline model that we use in the main analysis for comparison. The red line adds controls for the productivity growth observed in each city between 1996 and 1998, interacted with year dummies. The yellow line adds controls for the sectoral composition in each city prior to 1999 interacted with year fixed effects. The orange line adds controls for the share of left-wing votes in the presidential election of 1995 interacted with year dummies and a dummy for wether there was a change in majority between 1995 and 2002, also interacted with year dummies. The light blue line controls for city size prior to 1999, measured as the number of establishments active in the city, interacted with year fixed effects. Finally, the purple line adds all the controls at once.

FIGURE A8. Robustness of city-level results: other robustness tests

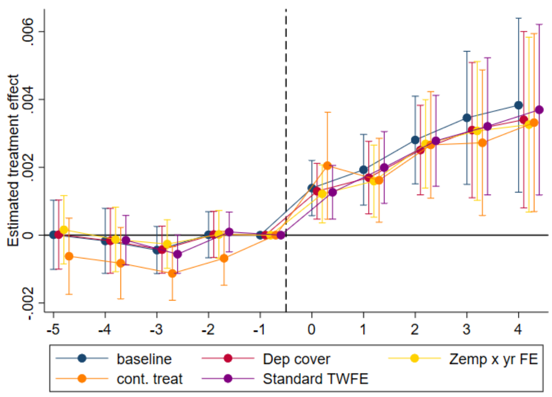
(a) Outsourcing/wage bill



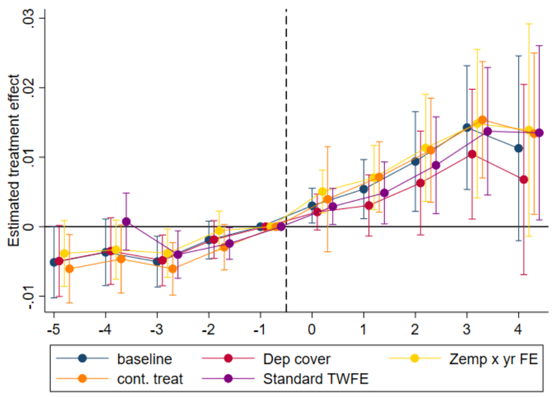
(b) Occup. concentration (HHI)



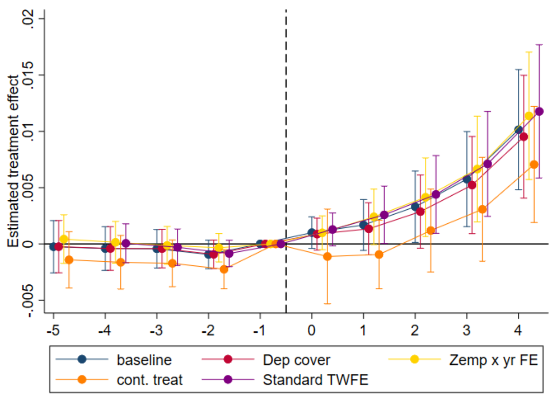
(c) Sh. of empl. in HS outs. services



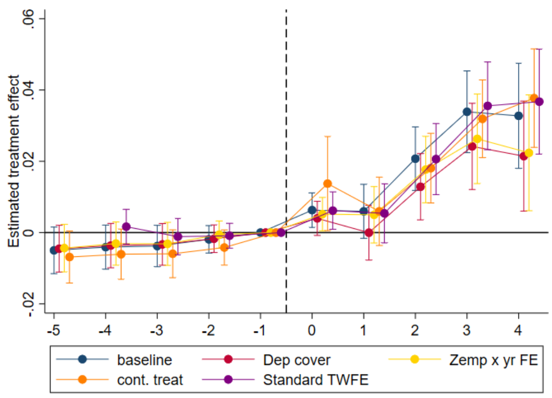
(d) Sh. outs. workers in HS outs. services



(e) Sh. of empl. in LS outs. services



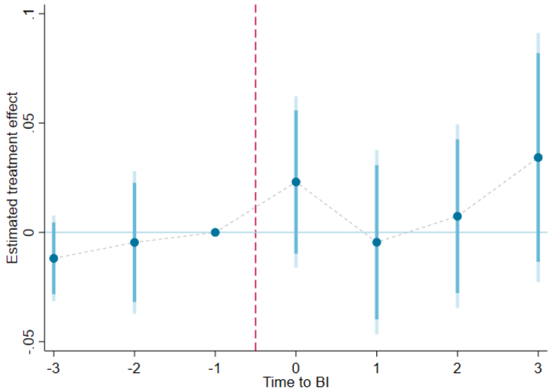
(f) Sh. outs. workers in LS outs. services



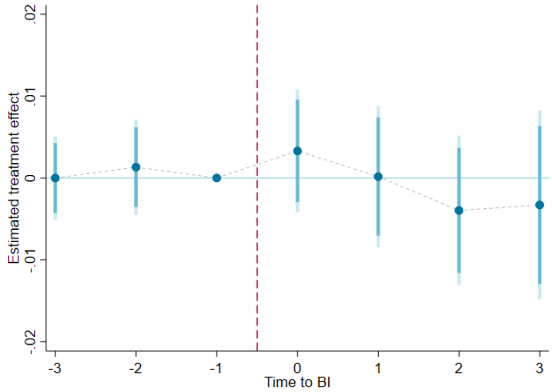
Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating slightly different models. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level. The blue line shows the baseline model that we use in the main analysis for comparison. The red line adds a control for the BI coverage observed in the other cities within the same department. The yellow line shows the results obtained after replacing the department × year fixed effects with commuting zone × year fixed effects. The orange line shows the results obtained while using the continuous measure of treatment $\tilde{Z}_i t$ instead of the binary indicator. The purple line shows the results obtained from running a standard staggered difference-in-differences model as in equation 2.

FIGURE A9. Robustness of city-level results: Placebo tests

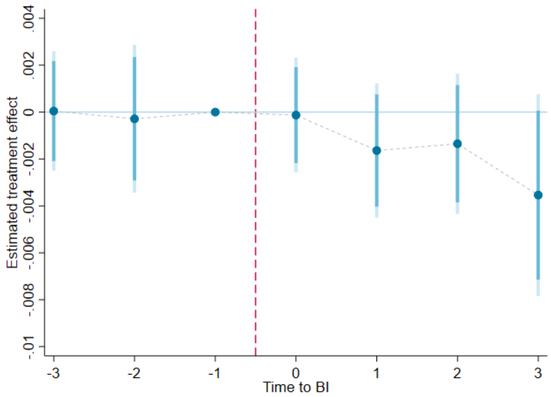
(a) Outsourcing/wage bill



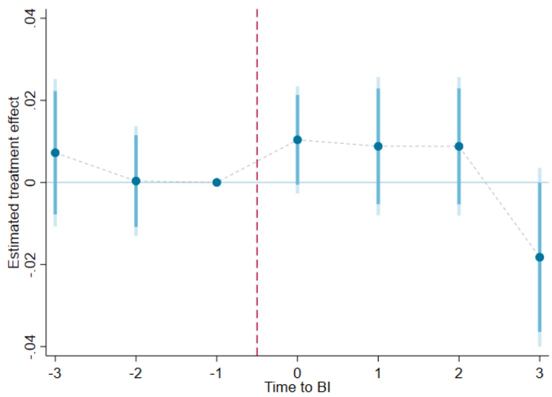
(b) Occup. concentration (HHI)



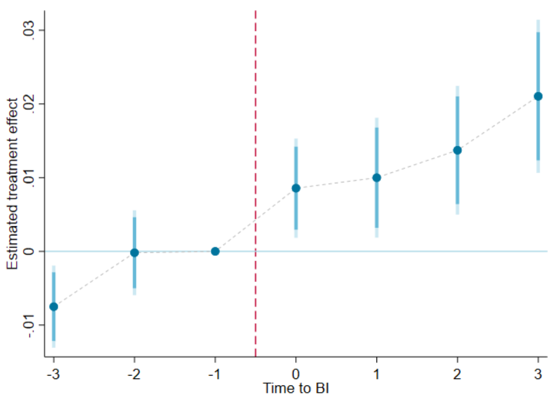
(c) Sh. of empl. in HS outs. services



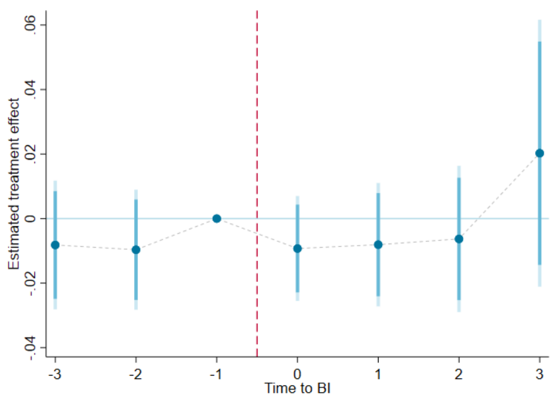
(d) Sh. outs. workers in HS outs. services



(e) Sh. of empl. in LS outs. services



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

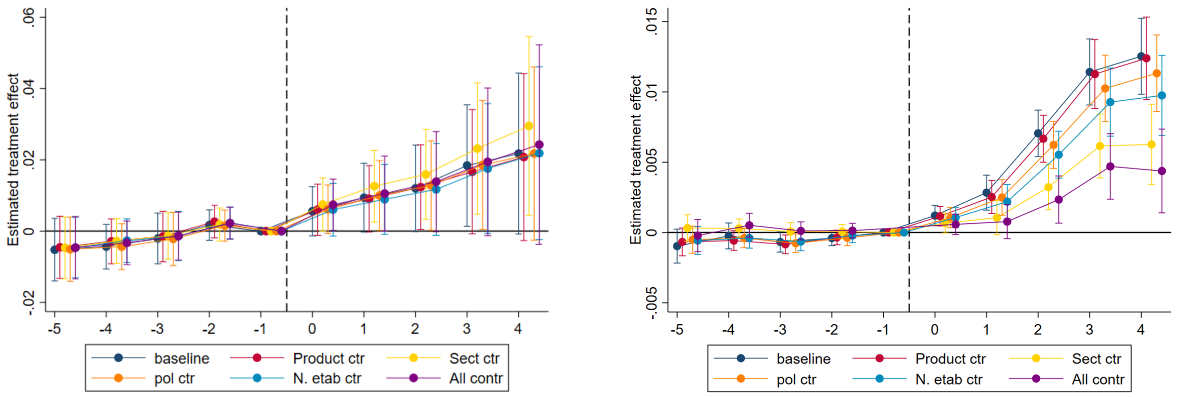


Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from placebo regressions. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level. Late receivers of broadband (2004 to 2007) are split into high (HPLR) and low propensity (LPLR) to be early receivers based on the propensity score. The HPLR are assigned to the pseudo-treatment year computed as the actual year of treatment - 4. The graph presents the pseudo-treatment effect observed over the period where none of the cities in the sample is actually treated (1997 to 2003).

FIGURE A10. Robustness of establishment-level results: adding additional controls

(a) Outsourcing/wage bill

(b) Occup. concentration (HHI)

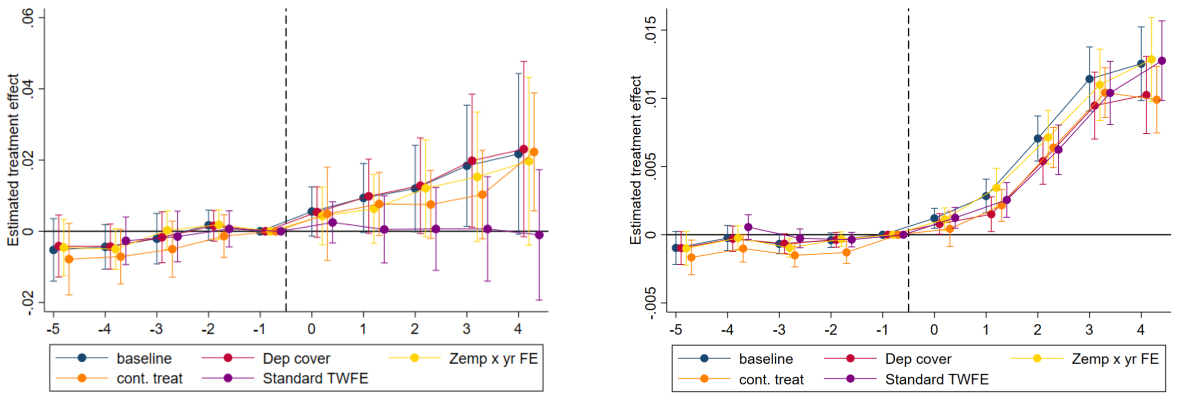


Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating equation (3) on establishment-level data with different sets of controls. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, establishment fixed effects and sample fixed effects. Standard errors are clustered at the department level.The navy blue line shows the baseline model that we use in the main analysis for comparison. The red line adds controls for the productivity growth observed in each city between 1996 and 1998, interacted with year dummies. The yellow line adds controls for the sectoral composition in each city prior to 1999 interacted with year fixed effects. The orange line adds controls for the share of left-wing votes in the presidential election of 1995 interacted with year dummies and a dummy for wether there was a change in majority between 1995 and 2002, also interacted with year dummies. The light blue line controls for city size prior to 1999, measured as the number of establishments active in the city, interacted with year fixed effects. Finally, the purple line adds all the controls at once.

FIGURE A11. Robustness of establishment-level results: other robustness tests

(a) Outsourcing/wage bill

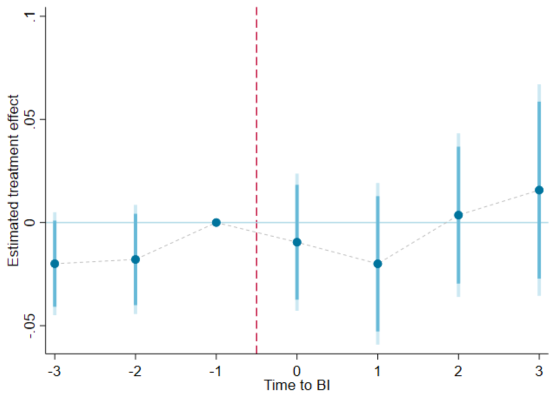
(b) Occup. concentration (HHI)



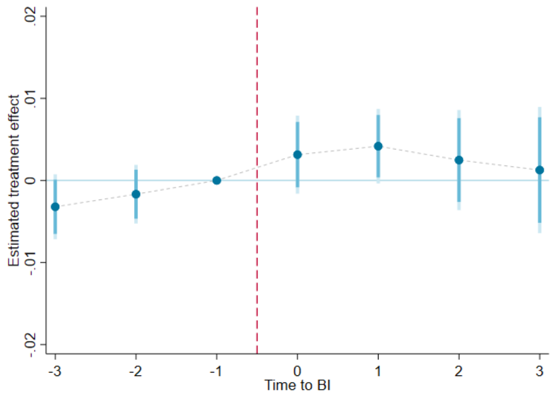
Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating slightly different models. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, establishment fixed effects and sample fixed effects. Standard errors are clustered at the department level. The blue line shows the baseline model that we use in the main analysis for comparison. The red line adds a control for the BI coverage observed in the other cities within the same department. The yellow line shows the results obtained after replacing the department × year fixed effects with commuting zone × year fixed effects. The orange line shows the results obtained while using the continuous measure of treatment $\tilde{Z}_i t$ instead of the binary indicator. The purple line shows the results obtained from running a standard staggered difference-in-differences model as in equation 2.

FIGURE A12. Robustness of establishment-level results: Placebo tests

(a) Outsourcing/wage bill

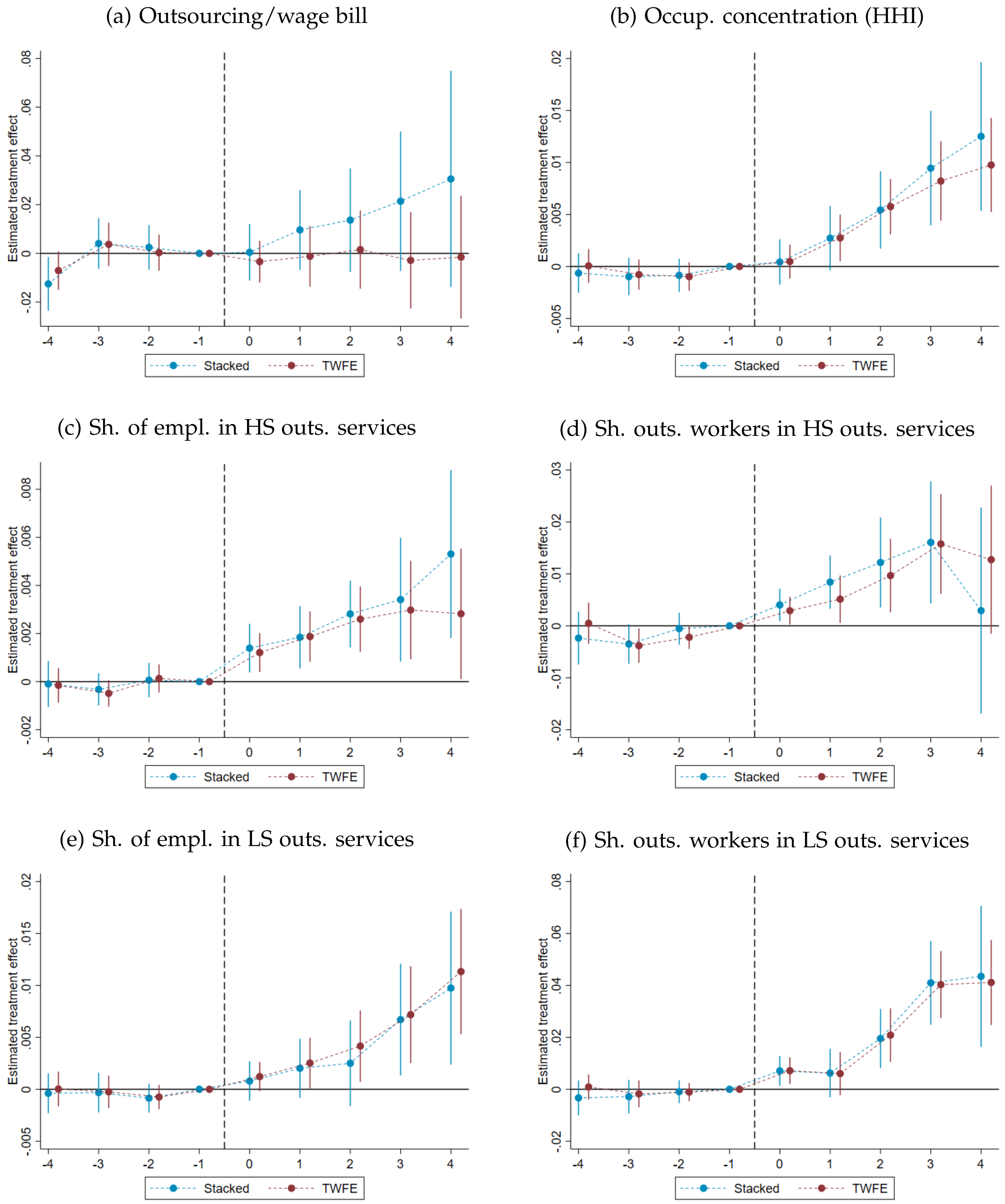


(b) Occup. concentration (HHI)



Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating slightly different models. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, establishment fixed effects and sample fixed effects. Standard errors are clustered at the department level. Late receivers of broadband (2004 to 2007) are split into high (HPLR) and low propensity (LPLR) to be early receivers based on the propensity score. The HPLR are assigned to the pseudo-treatment year computed as the actual year of treatment - 4. The graph presents the pseudo-treatment effect observed over the period where none of the cities in the sample is actually treated (1997 to 2003).

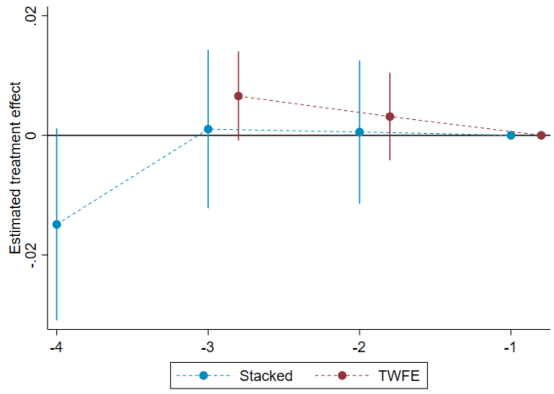
FIGURE A13. Event study graphs at the city level using cohorts treated in 2005-2007 as pure controls



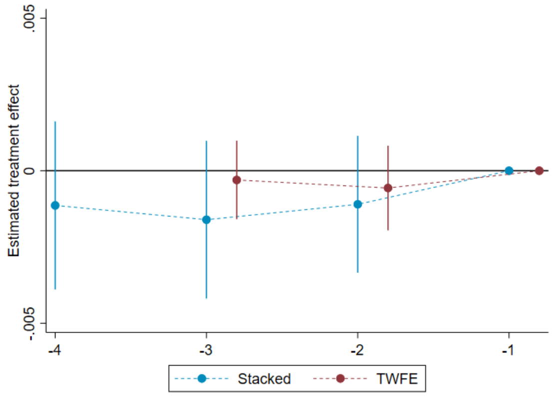
Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating equations (2) (TWFE) and (3) (Stacked) on city-level data for the years 1997 to 2004. As such, the cohorts treated between 2005 and 2007 serve as pure controls since they are never treated over this period. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

FIGURE A14. Additional pre-trend tests for the city level results

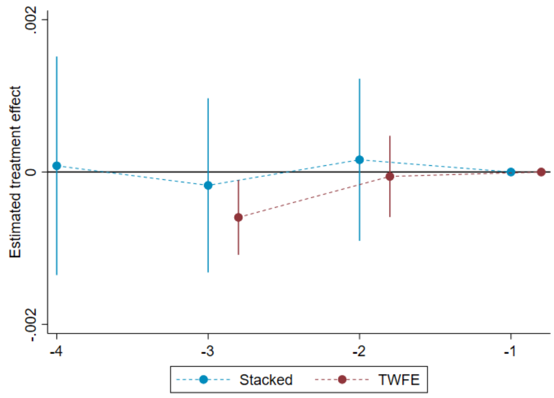
(a) Outsourcing/wage bill



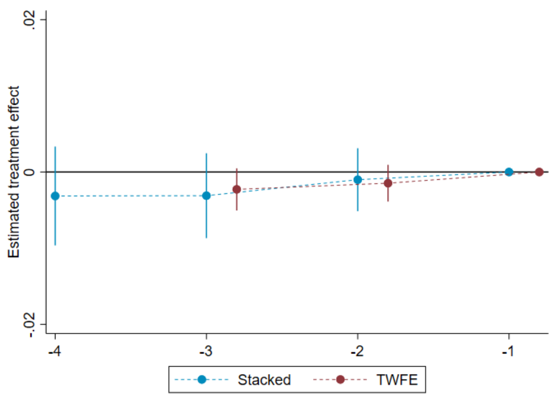
(b) Occup. concentration (HHI)



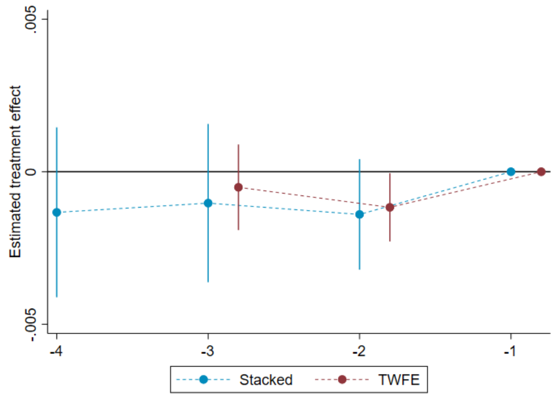
(c) Sh. of empl. in HS outs. services



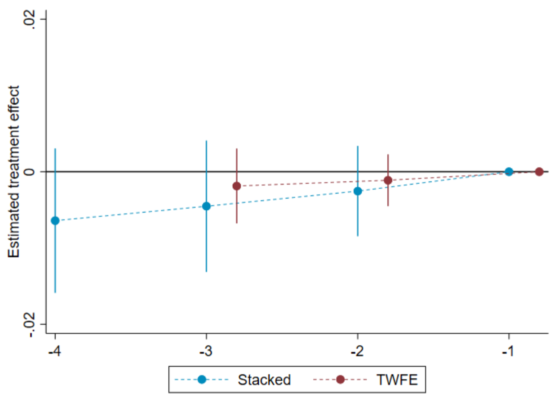
(d) Sh. outs. workers in HS outs. services



(e) Sh. of empl. in LS outs. services



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



Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating equations (2) (TWFE) and (3) (Stacked) on city-level data for the period preceding treatment (the data is cut at T-1 relative to treatment). The stacked model allows to estimate one additional pre-period dummy relative to the TWFE model. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level.

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).

As mentioned in the body of the text, the DADS do not contain information on education. Therefore, high and low-skill are based on intuitive interpretation of the occupational code. Nevertheless, in order to validate the final classification, we use additional survey data in order to assess whether the occupation codes that we use to define skill levels are strongly correlated or not with the level of education attained. To show this we computed some statistics based on the French Labor Force Survey, which includes both dimensions. In Table B2 we see that overall in the French labor market 19% of workers have less than a high-school degree, 58% have a high-school diploma, and 23% have more than a high-school degree. Among executives (used to measure skill-biased technical change in Appendix D) and among high-skill outsourceable occupations, only 3-4% have less than a high-school degree and 60% or more have more than a high-school degree. On the contrary, among low-skill outsourceable occupations 32% have less than a high-school degree and only 8% have more. Overall it appears that the occupations we chose, while conceptually distinct from formal education, correlate strongly with form education attainment.

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.

TABLE B2. Distribution of education levels across occupation categories

	Less than high-school degree	High-school degree	More than high-school degree
Across all occupations	19%	58%	23%
Executives	4%	30%	66%
High-skill outsourceable occup.	3%	37%	60%
Low-skill outsourceable occup.	32%	60%	8%

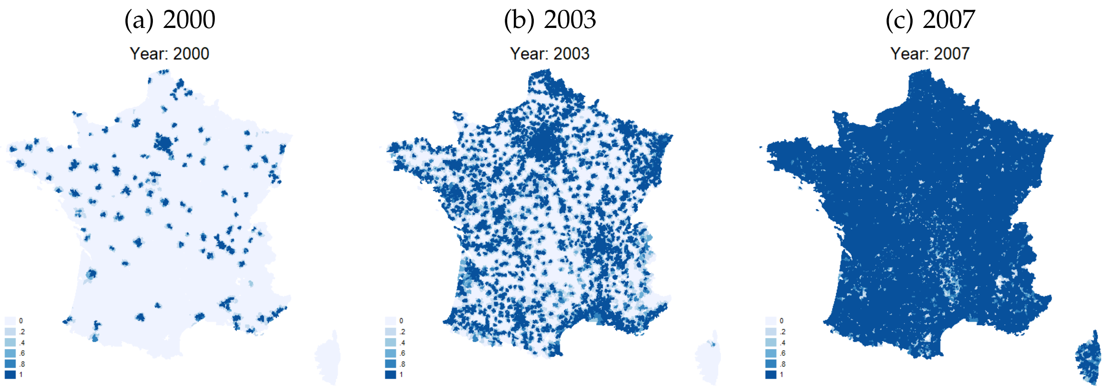
C ADSL in France

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. (2021), 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 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. (2021) 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.

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



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

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-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 3. 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. 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. At the city level, we consider the average productivity obtained across the local establishments, weighted by their size.[44] 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.[45]

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, whether measured by value adder per wagebill or using an estimate of the TFP.[46]. 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%)

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

[45] 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.

[46] We construct the TFP at the aggregate firm level by following the method of Aghion et al. (2023) and then we assign it to establishments using employment weights.

takes place in firms already present in the area before the shock.[47]

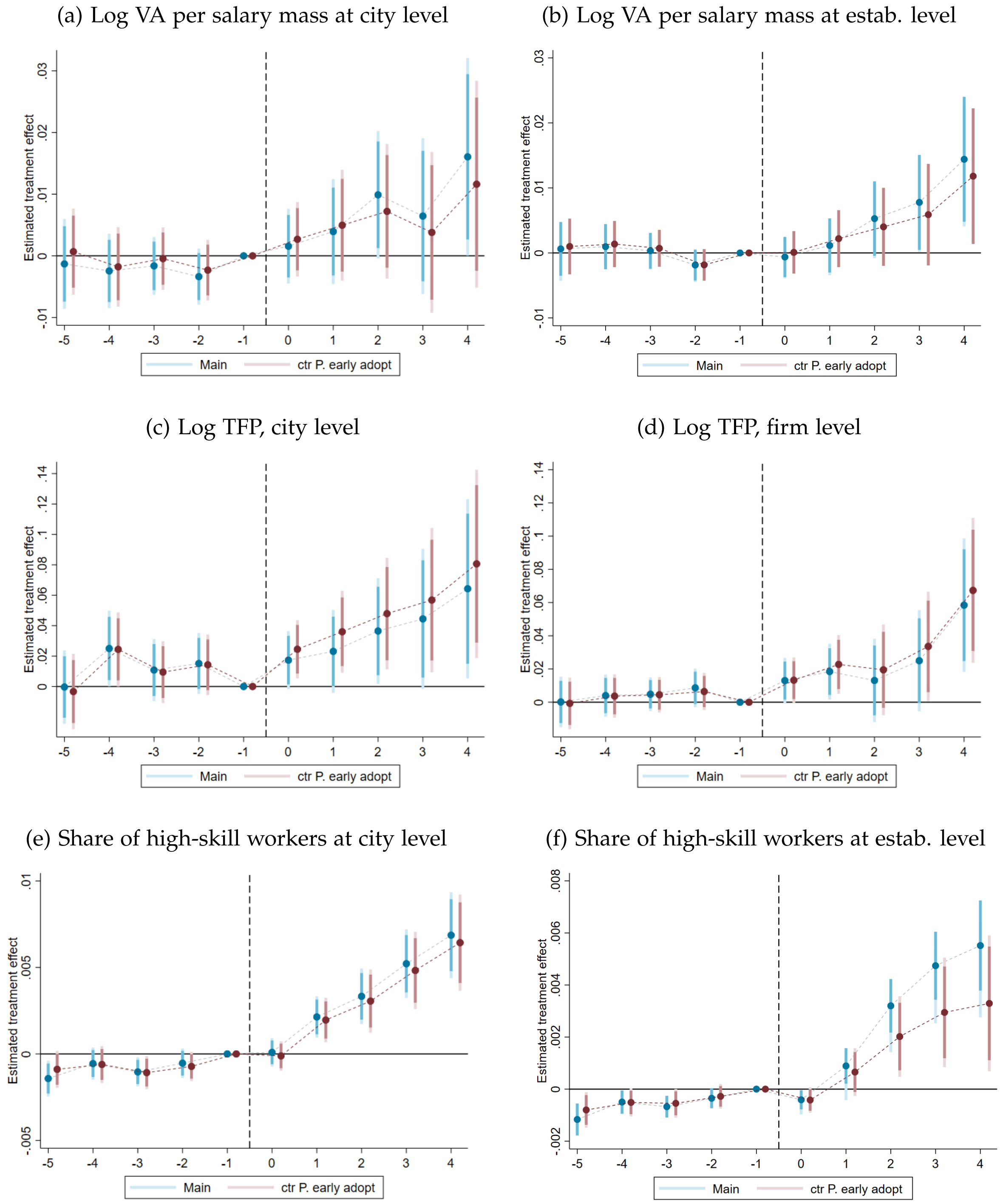
Similarly, before the arrival of BI, the share of high skill workers evolved comparably across cities belonging to different cohorts of ADSL diffusion, conditional on department-specific time trends and the other controls. When cities get access to BI, they experience a general upskilling of their labor force relative to other cities. 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, 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.

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).

Figures D2 and D3 show the same robustness tests that we perform on the outsourcing outcomes. In Figure D2 we test that results hold when adding additional controls: i) pre-BI productivity growth in the city interacted with year dummies, ii) pre-BI sectoral composition in the city interacted with year dummies, iii) share of left-wing voters in the city in 1995 and an indicator of cities switching political majority between 1995 and 2002, both interacted with year dummies, iv) pre-BI city size, measured by the number of establishments in the city, interacted with year fixed effects, v) all the

[47] 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 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).

FIGURE D1. Firm productivity and broadband access



Notes: This Figure shows regression coefficients and 90% 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), TFP (Panel c and d) or the share of executive workers within a city of establishment (Panel e and f) at *t* and the specification follows equation 3. The blue lines present our baseline model, while the red lines present the model controlling for the propensity score of early adoption interacted with year fixed effects. We construct the TFP at the aggregate firm level by following the method of Aghion et al. (2023) and then we assign it to establishments using employment weights.

controls added together. Figure D3 shows the results obtained i) when controlling for the BI coverage of other cities in the same department - to get a sense of possible spill-

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)
Baseline mean	0.058	0.103	0.68	0.644
	(0.060)	(0.150)	(0.327)	(0.463)
Observations	423,770	3,075,954	416,052	2,911,303
R-squared	0.711	0.894	0.622	0.711

Notes: ***, ** and * respectively denote significance at the 1, 5 and 10% level. Columns (1) and (3) run the regression at the city level, following equation 3, where controls are the population density in 1999 interacted with year dummies, department × 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.

over effects of BI arrival on control cities - , ii) when controlling for fixed effects at the commuting zone × year level instead of department × year, iii) when introducing the continuous measure of treatment $\tilde{Z}_{it}$ instead of the binary treatment, iv) when running a standard dynamic two-way fixed effects model as reported in equation (2). Figure D4 shows the placebo tests consisting in splitting late receivers into two groups according to their propensity of early adoption, and evaluating the pseudo-treatment on the period preceding their actual BI connection. The graphs mostly show flat and non-significant differences across the two groups, and if anything trends in the opposite direction than our actual results, which comfort our assumption that probability of early adoption is all but inflating our main coefficients. Table D3 shows the static coefficients obtained from a stacked difference-in-differences on the post-BI period, and Table D4 shows the coefficients from a standard staggered regression run at the labor market area (*Zone d'emploi*) level where the post BI dummies are interacted with the continuous measure of BI coverage in the area. Once again, most of the outcomes remain unchanged, except for outsourcing expenditure that becomes flat in the standard staggered event study model.

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:

$$\log(w_{i,t}) = \beta \tilde{Z}_{c(i),t} + X\gamma + \psi_{d,t} + \nu_i + \zeta_{s(i)} + \varepsilon_{i,t}, \tag{6}$$

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, and an indicator of whether the job is part-time (as opposed to full-time). Finally, $\psi_{d,t}$, ν_i, $\zeta_{s(i)}$ are a set of department d 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 departments but not across. Finally, β captures the effect (in percentage points) of being connected to BI on wage, controlling for observable and time-unvarying unobservable worker characteristics.

Table D2 presents our results and Table D5 presents the summary statistics of the variables used for the regression. Column (1) includes all workers (around 8 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

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All workers		3 skills		2 skills	
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: ***, ** and * respectively denote significance at the 1, 5 and 10% level. This Table shows regression results from an estimation of equation (6). Variable description is given in Table D5 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 département 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 département level under parenthesis.

to capture the level of skill of the worker.[48] 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).

[48]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 and Figures

TABLE D3. 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: ***, ** and * respectively denote significance at the 1, 5 and 10% level. The regressions are run at the city and establishment level following a model similar to equation 3, 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 × year fixed effects, city fixed effects and sample fixed effects.

TABLE D4. Effect of ADSL on skill-biased technical change - regressions at labor market area level

	(1) Sh. of high skill workers	(2) Log VA / salary mass
T = 0 x Zemp coverage	0.000876	0.0656
	(0.00244)	(0.0574)
T = +1 x Zemp coverage	0.00453***	0.0706
	(0.00136)	(0.0512)
T = +2 x Zemp coverage	0.0104***	0.0935*
	(0.00171)	(0.0512)
T = +3 x Zemp coverage	0.0156***	0.115
	(0.00217)	(0.0716)
T = +4 x Zemp coverage	0.0206***	0.136
	(0.00286)	(0.0944)
Average effect	0.0104***	0.0960*
	(0.00172)	(0.0563)
Observations	2,783	2,783
R-squared	0.989	0.781

Notes: ***, ** and * respectively denote significance at the 1, 5 and 10% level. The regressions are run at the labor market area (*Zone d'emploi*) level. Given that there is little variation in the timing of first BI appearance in labor market areas within the same department, we take advantage of the continuous measure of BI coverage: we estimate a standard staggered model similar to equation 2, but where we interact the dynamic post-BI dummies for every year with the share of the labor market area that is covered in that period. All columns control for the population density in 1999 interacted with year dummies, department × year fixed effects, and labor market area fixed effects.

TABLE D5. 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 × 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.

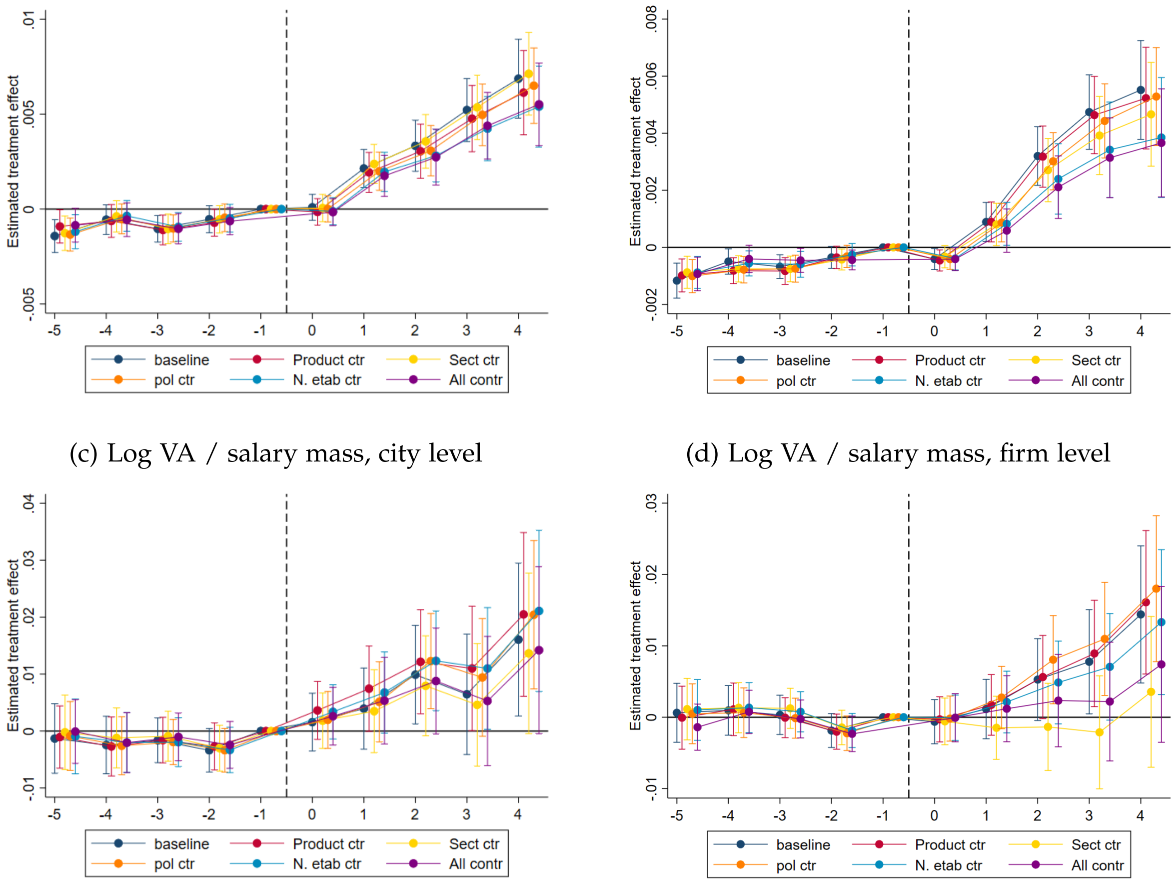
FIGURE D2. Robustness tests on SBTC outcomes: adding additional controls

(a) Sh. of high-skill workers, city level

(b) Sh. of high-skill workers, firm level

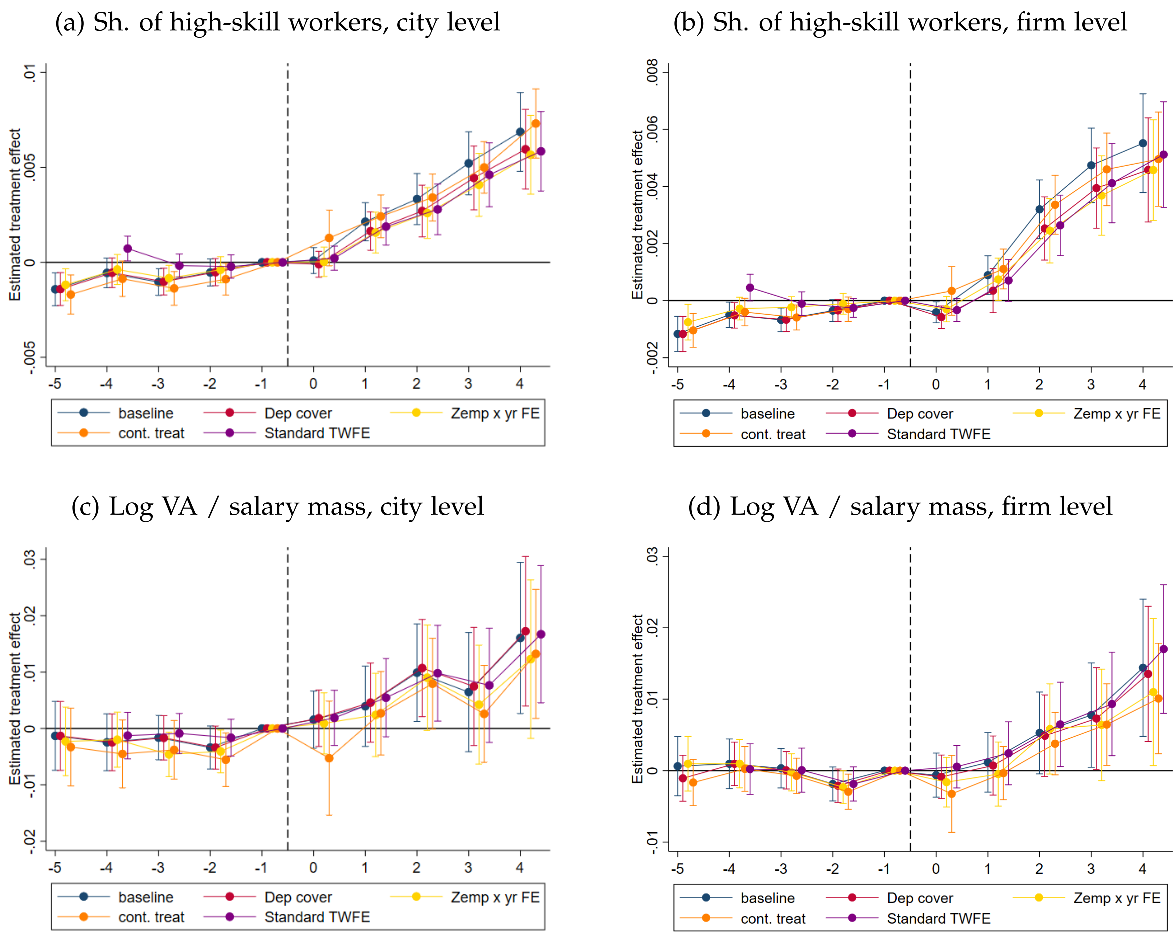
(c) Log VA / salary mass, city level

(d) Log VA / salary mass, firm level



Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating equation (3) on city- and firm-level data with different sets of controls. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects or firm fixed effects depending on the level of aggregation, and sample fixed effects. Standard errors are clustered at the department level. The navy blue line shows the baseline model that we use in the main analysis for comparison. The red line adds controls for the productivity growth observed in each city between 1996 and 1998, interacted with year dummies. The yellow line adds controls for the sectoral composition in each city prior to 1999 interacted with year fixed effects. The orange line adds controls for the share of left-wing votes in the presidential election of 1995 interacted with year dummies and a dummy for wether there was a change in majority between 1995 and 2002, also interacted with year dummies. The light blue line controls for city size prior to 1999, measured as the number of establishments active in the city, interacted with year fixed effects. Finally, the purple line adds all the controls at once.

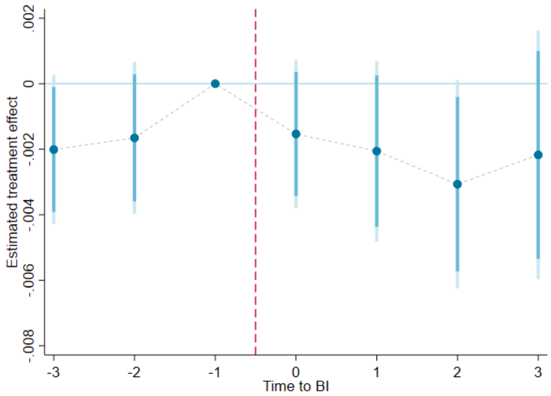
FIGURE D3. Robustness tests on SBTC outcomes: other robustness tests



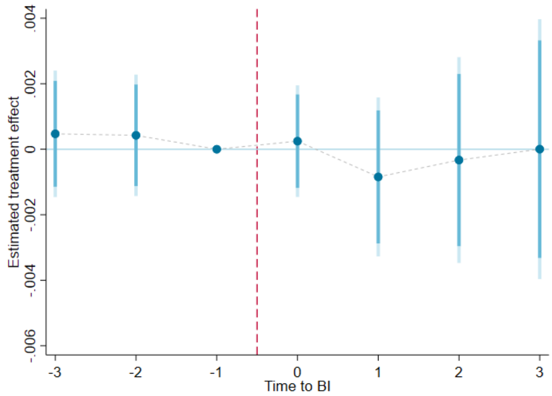
Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from estimating slightly different models. All models control for the population density in 1999 interacted with year dummies, department $\times$ year fixed effects, city fixed effects or firm fixed effects depending on the level of aggregation, and sample fixed effects. Standard errors are clustered at the department level. The blue line shows the baseline model that we use in the main analysis for comparison. The red line adds a control for the BI coverage observed in the other cities within the same department. The yellow line shows the results obtained from running a standard staggered difference-in-differences model as in equation 2. Finally, the orange line, which is only present for the outsourcing expenditure outcome, tests the robustness of this outcome from excluding multi-establishment firms, for which we cannot properly allocate performance across plants.

FIGURE D4. Robustness tests on SBTC outcomes: Placebo tests

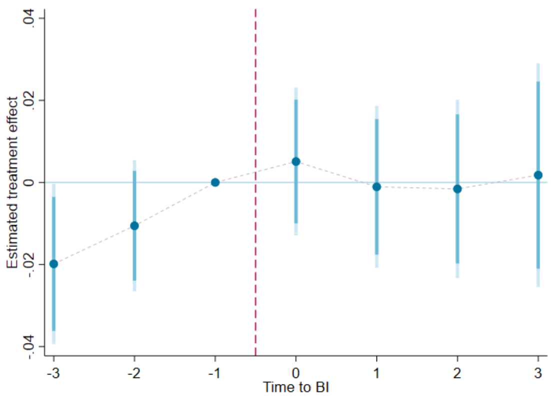
(a) Sh. of high-skill workers, city level



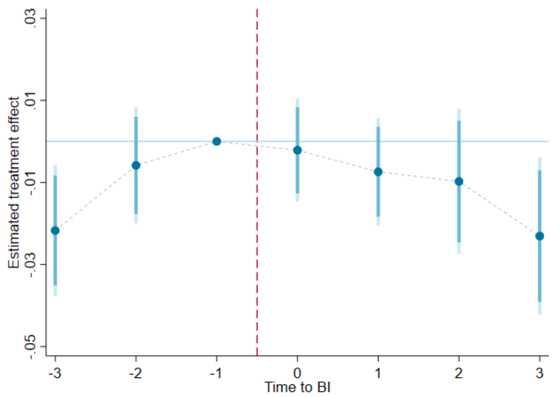
(b) Sh. of high-skill workers, firm level



(c) Log VA / salary mass, city level



(d) Log VA / salary mass, firm level



Notes: This Figure shows the point-estimate and 90% confidence intervals of the event study obtained from placebo regressions. All models control for the population density in 1999 interacted with year dummies, department × year fixed effects, city fixed effects and sample fixed effects. Standard errors are clustered at the department level. Late receivers of broadband (2004 to 2007) are split into high (HPLR) and low propensity (LPLR) to be early receivers based on the propensity score. The HPLR are assigned to the pseudo-treatment year computed as the actual year of treatment - 4. The graph presents the pseudo-treatment effect observed over the period where none of the cities in the sample is actually treated (1997 to 2003).

E A simple illustrative model

In this Appendix, we show how a simple model can generate the prediction that a global shock such as BI, which affect firms productivity and outsourcing cost, can lead firms to increase their outsourcing of some specific occupations.

Production technology: aggregation across occupations. Firms combine several occupations or tasks to produce output using a Cobb-Douglas function with constant returns to scale.[49] We denote output of j as a function of each occupation output $H_{i,j}$ as:

$$Y_j = \theta_j \prod_{i \in \mathbf{N}} H_{i,j}^{\alpha_i} \tag{7}$$

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:[50]

$$H_i = \left(\mu_i^{\frac{1}{\sigma_i}} n_i^{\frac{\sigma_i - 1}{\sigma_i}} + (1 - \mu_i)^{\frac{1}{\sigma_i}} s_i^{\frac{\sigma_i - 1}{\sigma_i}} \right)^{\frac{\sigma_i}{\sigma_i - 1}} \tag{8}$$

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.

[49] As in the empirical analysis, we make the assumption that one occupation is a fixed bundle of tasks.

[50] 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.

The model encapsulates the idea outlined in Section 2 that firms differentiate occupations based on how core they are. This is captures by 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.[51]

This idea is summarized by equation (9) which gives the labor supply curve of a firm j hiring a (in-house) worker in occupation i. Namely:

[51] 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.

$$n_{i,j} = N_i \frac{w_{i,j}^{1/\rho_i}}{\sum_j w_{i,j}{}^{1/\rho_i}} = a_i w_{i,j}^{\frac{1}{\rho_i}}, \tag{9}$$

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.[52] 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:

$$R(Y) = Yp = Y^{\frac{\varepsilon-1}{\varepsilon}} I^{\frac{1}{\varepsilon}} \tag{10}$$

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

$$C(\{n_i, s_i\}_{i\in\mathbf{N}}) = \sum_{i\in\mathbf{N}} n_{i,j} w_i(n_{i,j}) + \sum_{i\in\mathbf{N}} s_{i,j} \cdot \gamma_{i,j} r_i \tag{11}$$

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.

$$\max_{\{n_{i,j}, s_{i,j}\}_{i\in\mathbf{N}}} \pi_{i,j} = Y(\{n_{i,j}, s_{i,j}\}_{i\in\mathbf{N}})^{\frac{\varepsilon-1}{\varepsilon}} I^{\frac{1}{\varepsilon}} - \left(\sum_{i\in\mathbf{N}} n_{i,j} w_i(n_{i,j}) + \sum_{i\in\mathbf{N}} s_{i,j} \cdot \gamma_{i,j} r_i\right) \tag{12}$$

It is fairly straightforward to show that the problem defined in Equation (12) admits a unique positive solution $\{n^*_{i,j}, s^*_{i,j}\}_{i\in\mathbf{N}}$.[53]

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

[52] 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.

[53] 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.

E.1 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:

$$s_{i,j} = \lambda_{i,j} n_{i,j}^{\rho\sigma_i+1}, \text{ where } \lambda_{i,j} = \frac{1-\mu_i}{\mu_i}\left[\frac{\rho+1}{a_i^{\rho} r_i \gamma_{i,j}}\right]^{\sigma_i}. \tag{13}$$

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:

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

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

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

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

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

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

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:

$$d\log\theta + \left(\sum_{i'} \alpha_{i'} d\log n_{i'} (1 + e_{i'}\rho\sigma_{i'})\right) = \frac{\varepsilon}{\varepsilon - 1} (1/\sigma_i + \rho + 1 + e_i\rho\sigma_i)\, d\log n_i. \quad (14)$$

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 $\frac{d\eta_i^c}{d\theta} > 0$. □

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. Starting from equation (14) 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 (14) 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. □

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. 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:[54]

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

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. □

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.3, 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.

E.2 The case of a reduction in the cost of outsourcing

So far we have considered the effect of an increase in θ. As we have explained in Section 2, BI is also likely to have reduced the cost of outsourcing for firms. 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.

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

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 $\sigma_i > 0$:

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

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

$$d\log(s_i) = -\sigma_i d\log(\gamma) + (\rho\sigma_i + 1)d\log(n_i) \tag{15}$$

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 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 (15). And from the previous lemma, we know that $d\log(s_1)/d\log(\gamma) < 0$.

Using again equation (15), we also have

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

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:

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

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

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

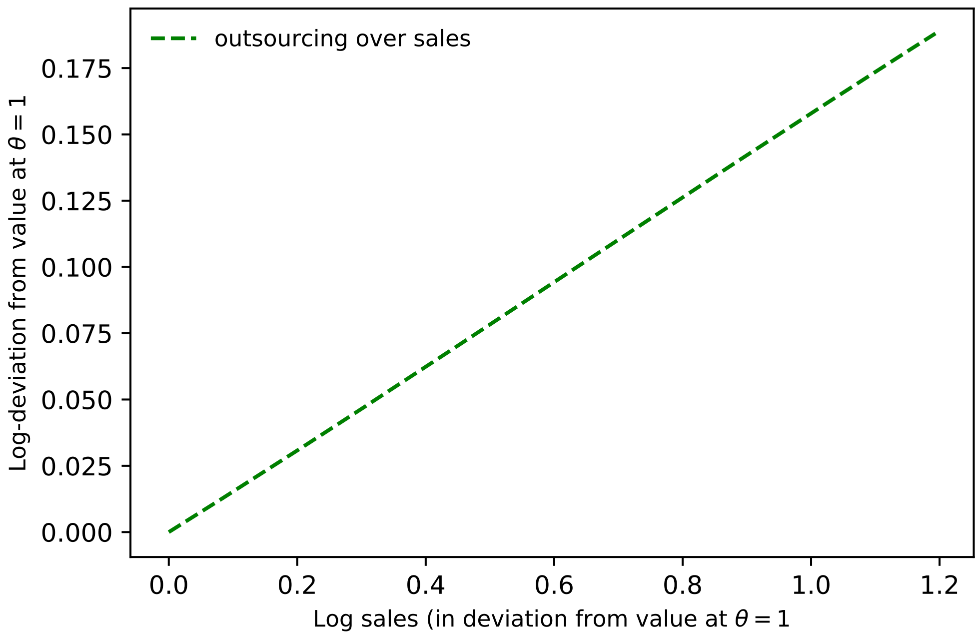
which is positive as long as α_1 is not too large or ρ is not too small. This leads to an impossible statement and hence contradict the assumption that $d\log(n_2)/d\log(\gamma) \ge 1/\rho$.

E.3 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 the interval $[1,2]$ as the support for θ.[55]

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 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.

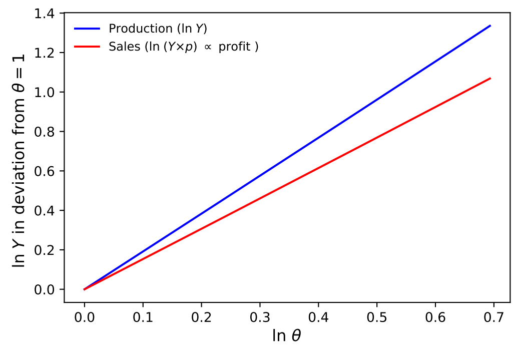
Figure E2 presents a set of results. Panel 2(a) shows how the optimal scale of production and sales evolves as productivity increase. Production Y increases log-linearly with θ, with an elasticity close to 1. Sales, which are proportional to profit in this model, increases also linearly but because the 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

[55]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 $N_i = 1, \forall i \in \mathbf{N}$.

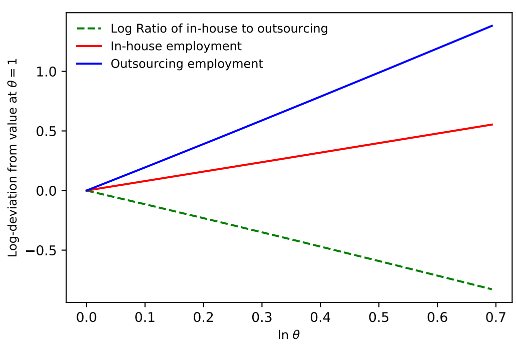
2(c) show that this declining ratio is heterogeneous across occupations. It displays the ratio for a core and non-core occupation with the same weight in the Cobb-Douglas production function. Panel 2(d) makes the same point but focusing on the level of the cost share represented by in-house labor. We 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.

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

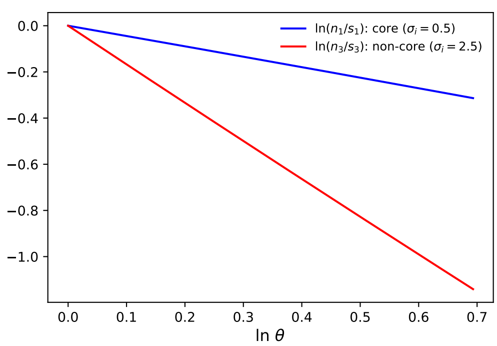
(a) Scale of production



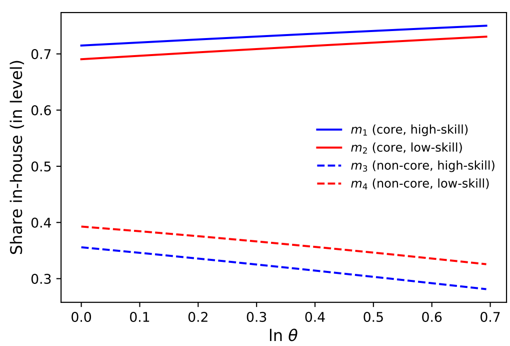
(b) Employment : in-house versus outsourced



(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.

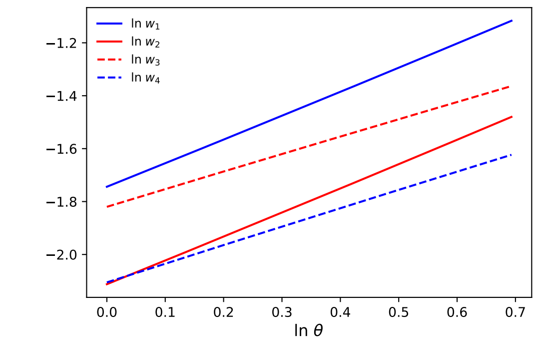
The figure E3 displays four other comparative statics. Panel 3(a) shows the wage in level. Unsurprisingly, high-skill occupations (1 and 3) have the highest wages. We see however that the firm size wage premium is stronger among core occupations independently of skill-level. Panel 3(b) show how log-wage 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 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.

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 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.

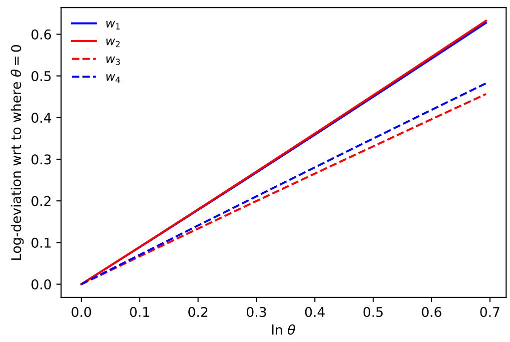
As mentioned above, broadband is also likely to result in a decrease of outsourcing cost, which we capture with a decline in the term γ_{ij}. Decrease in this parameter leads to broadly similar comparative statics as the previous case as displayed in Figure E5.

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

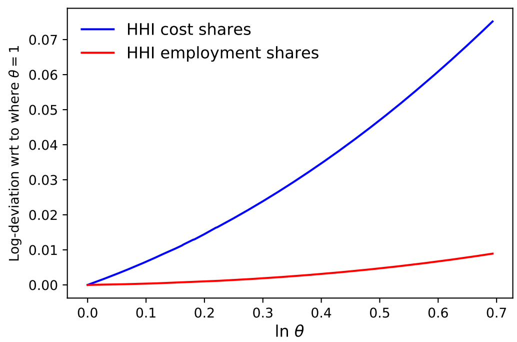
(a) Wage by occupation



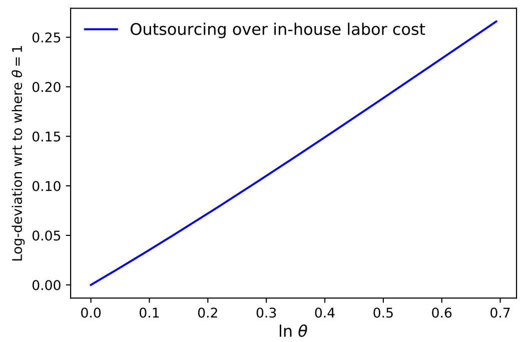
(b) – normalized with respect to $\theta = 1$



(c) HHI in terms of cost shares



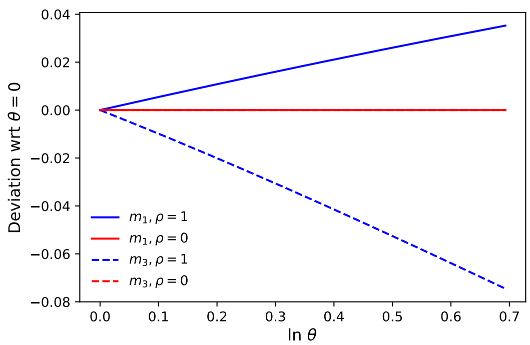
(d) Outsourcing over in house labor cost



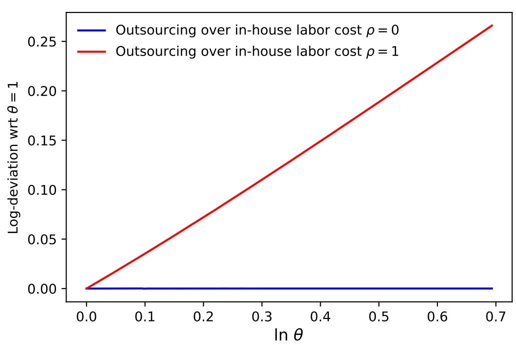
Notes: The figure provide comparative statics with respect to an increase in Hicks neutral productivity shifter θ. We consider support the interval $[1, 2]$ as support for θ.

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

(a) Scale of production



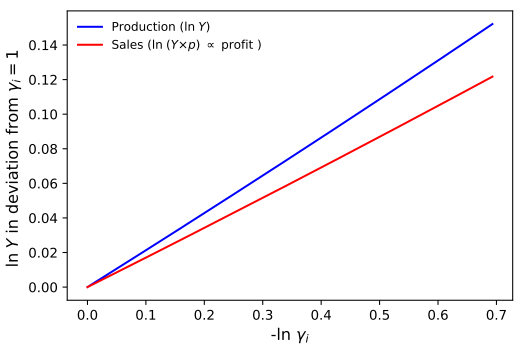
(b) Employment : in-house versus outsourced



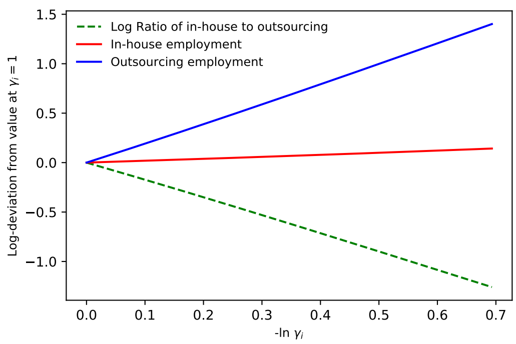
Notes: The figure provide comparative statics with respect to an increase in Hicks neutral productivity shifter θ. We consider support the interval $[1, 2]$ as support for θ.

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

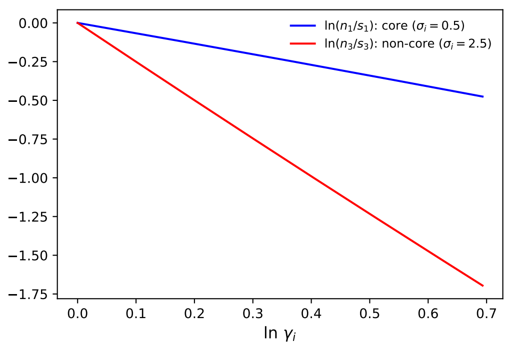
(a) Scale of production



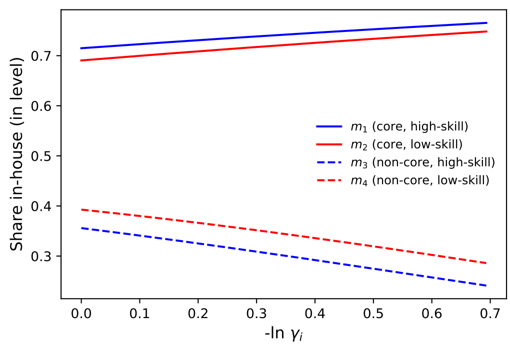
(b) Employment : in-house versus outsourced



(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 $\gamma : 1 \rightarrow 0.5$.