Location Effects or Sorting? Evidence from Firm Relocation*

Pauline Carry
Princeton University

Benny Kleinman
Stanford University

Elio Nimier-David
Cornell University

March 20, 2025

Abstract

Why are wages in cities like New York or Paris higher than in others? This paper leverages firm mobility to separate the role of "location effects" (e.g. local geography, infrastructure, and agglomeration) from the spatial sorting of workers and firms. Using comprehensive French administrative records and U.S. commercial data, we first document the prevalence of firm mobility and what it entails for a firm to relocate. We find that establishment mobility is widespread, occurring at an annual rate of about 4%. Establishments retain their main activity and structure as they move, but adjust their workforce and wages. Combining firm and worker mobility, we decompose wage disparities across French commuting zones: location effects account for only 2–5% of the gap, while differences in the composition of workers and establishments account for around 30% and 15%, respectively. The remaining half is accounted for by the co-location of high-wage workers and firms, especially in cities with high location effects. Revisiting the elasticity of local wages to population density, we find a significant coefficient of 0.007—two to three times lower than estimates not controlling for firm composition.

*We are grateful to Hadar Avivi, Michèle Belot, Marianne Bertrand, Zach Bleemer, Janet Currie, Ellora Derenoncourt, Joshua Gottlieb, John Grigsby, Rick Hornbeck, Anders Humlum, Erik Hurst, Simon Jäger, Philipp Kircher, Ilyana Kuziemko, Thibaut Lamadon, Jeremy Lise, Jack Mountjoy, Matt Notowidigdo, Ezra Oberfield, Roman Rivera, Richard Rogerson, Esteban Rossi-Hansberg, Karthik Sastry, Robert Shimer, Benjamin Schoefer, Chad Syverson, Owen Zidar and Eric Zwick for useful feedback. We would like to thank seminar participants at the University of Chicago, Cornell University, Princeton University, and HEC Paris for their comments. We thank Cedric Christensen, Silvia Pedersoli, and Noah Simon for outstanding research assistance.

1 Introduction

Why are wages in cities like New York or Paris higher than in others? Existing explanations fall into two broad categories. One view emphasizes "location effects," suggesting that these cities offer advantages—such as better infrastructure, agglomeration spillovers, or distinct labor market conditions—that boost productivity and wages. Under this perspective, a hypothetical relocation of a firm and its workers from a low-wage city to Paris would lead to higher wages in the firm. The alternative view attributes the wage premium to spatial sorting— i.e. differences in the local composition of firms and workers. In this view, moving a firm and its employees to Paris would not change their wages. Distinguishing between these explanations is crucial for policymakers, who allocate billions to attract firms through local business incentives and tax breaks (Bartik, 2020), assuming that bringing in "high-paying firms" can raise local prosperity. It is also essential for economists modeling cities, as it clarifies the relative importance of spatial sorting versus regional fundamentals and agglomeration effects.

In this paper, we leverage firm mobility across space to identify the importance of "location effects" separately from spatial sorting. To this end, we first establish a set of stylized facts about the understudied phenomenon of firm mobility. Using administrative data from France and commercial data from the United States, we examine how prevalent firm relocation is, what a relocation entails for a firm, and which firms move where. We also present a simple spatial equilibrium model with firm mobility that accounts for our findings. We then combine this new source of identification with worker mobility—both across firms and across space—to decompose local wage disparities in France, disentangling the contributions of firms, locations, and workers. We find that location effects account for only 2-5% of wage differences between commuting zones (CZs), while spatial sorting of workers and firms accounts for most regional wage disparities. Using our estimated location effects, we revisit the city-size wage premium, providing new estimates of the sensitivity of wages to local density when controlling for the composition of workers and firms.

Studying relocation decisions presents a key challenge: in many administrative data sources, an establishment's identifier changes when the establishment moves. We address this issue using unique data from France's *Register of Establishment Relocation* (1993–2021), which mandates reporting whenever an establishment changes its address.[1] In practice, and uniquely to these data, we observe on the same form the establishment's identity, its old location, and its new location, guaranteeing that we track the same establishment across space. We conduct a series of validation tests to ensure that relocations in these data align with the intuitive notion of a change in the establishment's

[1] In this context, a relocation is defined as "*the transfer of activities and equipment to a new location, with closure of the establishment at the former location and opening at the new location*". Entrepreneurs must show proof of closure and opening within three months to have a valid business license.

location of activity. To ensure our findings are not specific to France or to this particular definition of relocation, we replicate our main results using U.S. commercial data from Dun & Bradstreet, as well as data on the headquarters locations of U.S. publicly listed firms, tracking firms across space by name, age, and ownership structure. Throughout the paper, we measure the relocation of individual establishments and use the terms establishment relocation and firm relocation interchangeably. Most of the economy consists of single-establishment firms. For multi-establishment firms, we track the relocation of all establishment types in the French data, though in practice it is often the firm's headquarters that relocates.

We begin by presenting key facts about firm relocation decisions to both familiarize readers with the data and underscore its relevance to our wage decomposition analysis. First, we show that relocation is a common phenomenon, with approximately 4% of all establishments moving each year, spanning businesses of all sizes and sectors. Notably, a quarter of those relocations involve a change in commuting zone. Second, we explore the geography of relocation. Similar to worker migration, firm mobility is characterized by bidirectional flows between every two pair of locations, suggesting little role for shared location factors like market size or local taxation in driving destination choices. Also similar to worker flows, firm relocations are well approximated by a gravity equation, with distance playing a strong role in movement rates between locations. Third, we examine heterogeneity in firms' location decisions and find evidence of spatial sorting: conditional on the origin location and firm age, larger and higher-wage firms are more likely to choose larger, higher-wage destinations.

We next examine how establishments evolve following relocation. We find that relocation results in minimal changes to a firm's core activities or production structure. Using administrative data, we find no evidence of shifts in industry or legal status. We also find very little change in input shares, both for capital and for workforce occupational distribution. Moreover, relying on a survey of entrepreneurs, we document that establishments relocating have not experienced more problems, challenges or shocks than non-movers. Zooming into the workforce composition, we find that relocating establishments replace a large share of their former workforce with local employees in the new location. On average, half of former employees move with the establishment. The average wage of an establishment adjusts: moving to a commuting zone where wages are higher by 1% increases the average wage of the establishment by 0.15%. This small elasticity suggests a large establishment-component of wages. Moreover, as we show using our decomposition, the majority of the change is due to the hiring of different workers, rather than the location effect.

We develop a simple spatial equilibrium framework to rationalize the descriptive evidence on relocation decisions and to further motivate our wage decomposition exercise. Our framework incorporates both firm and worker mobility, providing a theoretical in-

terpretation of our estimated worker, firm, and location effects. In the model, firms with heterogeneous productivity select locations to maximize profits, subject to idiosyncratic location preferences, similar to the migration literature. Workers with heterogeneous abilities choose locations and match with firms in frictional local labor markets. The resulting equilibrium generates gravity-like patterns of firm relocation, predicts the sorting of more productive firms and workers into more productive locations, and yields a simple wage equation featuring distinct location, firm, and worker components.

In our main decomposition exercise, we exploit the widespread mobility of establishments and the persistence of their characteristics to examine the role of firms in generating spatial wage disparities, accounting for the composition of workers. We conduct a wage decomposition analysis inspired by Abowd et al. (1999) (hereafter AKM). Departing from the traditional wage decomposition literature, we disentangle the influence of locations from that of establishments. We estimate a three-way decomposition with worker, establishment and location fixed effects, leveraging both establishment mobility across locations and worker mobility across establishments and locations for identification. Worker mobility connects both establishments and locations and firm mobility connects locations. Identification relies on the assumption of random mobility of firms and workers, conditional on the model fixed effects. We present three tests supporting this assumption. First, we consider event studies for firm and worker wages as in Badinski et al. (2023); Card et al. (2013, 2016). We observe no significant pre- or post-movement wage dynamics for either establishments or workers, along with nearly symmetric wage adjustments when moving to higher-wage or lower-wage locations. Second, we present a placebo test enabled by our new setting: we track the wage evolution of establishments who relocate *within* the same commuting zone. In that case, wages remain flat in the year of the move, consistent with the establishment remaining in the same labor market. Third, the survey of entrepreneurs allows us to verify that establishments relocating did not experience shocks that would affect wages directly.

We use our estimated fixed effects to decompose the variance of average wages between commuting zones in France. We find that location effects per se account for only 2%–5% of the total variance in log wages across commuting zones. In contrast, the vast majority of spatial wage inequality arises from the sorting of firms and workers across space. Differences in worker composition across cities explain approximately 30% of the total variance, while differences in firm composition contribute an additional 15%. In the language of Song et al. (2019), these two components of spatial sorting reflect *spatial segregation*, where workers of similar types cluster together, as do firms. Moreover, about a third of the total variance stems from the *co-location* of high-type firms and high-type workers in the same areas. Finally, the sorting of high-type firms and workers into commuting zones with strong location effects accounts for an additional 15% of the total variance.

These findings suggest that spatial sorting and segregation play a far greater role in explaining regional disparities than local geography, physical infrastructure, or agglomeration effects. A key question, then, is whether standard evidence for agglomeration effects holds once we account for the composition of both workers and firms. To investigate this, we revisit the well-documented relationship between regional wages and population density. The elasticity of local wages with respect to population density is typically estimated at 0.06–0.08. Drawing on insights from previous literature (Glaeser and Gottlieb, 2009; Glaeser and Resseger, 2010; De la Roca and Puga, 2016; Duranton and Puga, 2023), we find that this elasticity declines substantially to 0.02–0.04 when controlling for worker composition. Once firm composition is also accounted for, the estimate drops further to a precise and significant 0.007. In sum, while the wage-density relationship weakens considerably when controlling for spatial sorting, it remains positive and significant, showing that local density is an important predictor of local wages beyond the role of firms.

We highlight two key implications of our results. First, location effects play a minor role in generating spatial wage disparities, whereas the spatial sorting of workers and firms accounts for the majority. These results suggest a relatively small role for agglomeration spillovers, and offer guidance on how to build and discipline spatial models of cities and regions. Second, they align with local policymakers' priors that "high-paying" firms are consequential for shaping regional income. Moreover, our findings indicate a strong role for local human capital—as captured by workers' fixed effects—in shaping firms' location decisions.

This paper relates to four strands of literature. First, we contribute to the literature examining the determinants of local wages. A significant part of this literature leverages worker mobility across locations to separate location-specific wage effects from worker characteristics (Glaeser and Maré, 2001; Combes et al., 2008; Enrico, 2011; De la Roca and Puga, 2016; Porcher et al., 2023; Moretti and Yi, 2024; Card et al., 2025), estimating AKM-like models. Relative to this literature, we disentangle the role of establishment sorting from location effects.[2] We also add to the extensive literature documenting higher wages in larger and denser cities (Glaeser and Gottlieb, 2009; Mion and Naticchioni, 2009; Combes et al., 2010; Glaeser and Resseger, 2010; De la Roca and Puga, 2016; Duranton and Puga, 2023). We emphasize how local worker and establishment composition explains most of the variation in wages observed across locations.[3] Nevertheless, we still find a significant association between our estimated location effects and measures of local size

[2]As a benchmark, we replicate the methods of Glaeser and Maré (2001); Combes et al. (2008); Card et al. (2025). Notably, we find estimates very close to what Card et al. (2025) obtain for the US in the French context: controlling for worker effects, other characteristics of the location account for 26% of spatial wage disparities in France (29% in the US).

[3]While our focus in on wages, our findings on the large role played by the local firm composition relate to Schoefer and Ziv (2022) who focus on the determinants of local productivity.

or density, which shows that location size affects wages beyond the role played by local establishments.

Second, we relate to the theoretical literature in spatial economics on firm sorting and its role in explaining differences in economic activity across space, including Suárez Serrato and Zidar (2016), Gaubert (2018), Lindenlaub et al. (2022), Lhuillier (2022), Hong (2023), and Oh (2023). These studies consider static firm location decisions and do not exploit information on relocation decisions. We contribute to these questions by quantifying the magnitude and significance of firm spatial sorting using a movers design, without relying on a fully-specified model. In addition, while our focus is mainly empirical, we provide a spatial equilibrium model with firm and worker mobility that aligns with our key findings. More generally, our results provide a useful set of empirical moments to discipline the importance of sorting relative to differences in location fundamentals and agglomeration forces in spatial equilibrium models.

Third, we contribute to the scarce literature on firm mobility. A few studies examine firm relocation decisions in specific contexts: Duranton and Puga (2001) focuses on business services exclusively and documents net migration of mature firms from diversified cities to specialized locations; Voget (2011) studies relocation of multinational headquarters as a response to taxation; Bergeaud and Ray (2020) investigates the role of real estate; Bryan and Guzman (2023) studies the movement of high-potential startups. Relative to these studies, we provide a comprehensive view of firm relocation decisions across the economy. To the best of our knowledge, we are the first to provide systematic evidence covering the prevalence of firm mobility in the economy (in France and the U.S.), the geography of relocation, and how firms change upon a move. We are the first to leverage this widespread phenomenon to quantify the importance of firms for spatial wage inequality.

Fourth, we relate to the literature on wage setting in multi-unit and multinational firms (Hjort et al., 2020; Setzler and Tintelnot, 2021; Hazell et al., 2024). An important theme in this literature is the importance of firm-level wage policies that can affect establishments in different locations. By studying individual establishments we account for the large variation in establishment wages within firms. Moreover, we account for spatial wage dispersion driven by both single-establishment firms–that constitute most of the economy and most of the workforce–and multi-establishment firms.

The remainder of the paper is structured as follows. In Section 2, we present the data we use for France and the United States and discuss the definition of relocation. Section 3 characterizes firm relocation and what it entails for an establishment to move. In Section 4, we present a spatial equilibrium model accounting for those results. We present the empirical model used to decompose spatial wage disparities in Section 5, and the results in Section 6. Section 7 concludes.

2 Context and data

This section presents the data, with an emphasis on measuring establishment relocation in France and the US. We begin by outlining the legal and administrative process associated with a change of address in France. Due to these legal requirements, we can comprehensively track establishment relocation in France. We then introduce the data used to proxy relocation rates for the US.

2.1 Process to relocate an establishment in France

The French statistical institute, INSEE, manages the registry of establishments in France and assigns both establishment and firm identifiers. This registry records the address, industry, opening date, and closure date of all establishments. INSEE defines an establishment relocation as "*the transfer of activities and equipment to a new location, with closure of the establishment at the former location and opening at the new location.*" In practice, an entrepreneur relocating an establishment must report a change of address to the local commercial court within three months of the move. The commercial court then updates the business license (Kbis) with the new address. This license is the sole document proving the existence of an operating establishment. An up-to-date license, with the correct address, is required for entrepreneurs to conduct their businesses (e.g., sign contracts or open bank accounts). To declare an address change, three documents are needed. First, the entrepreneur must provide proof of closure of the establishment at the former location. Second, they must present evidence of the establishment's reopening at the new location, which can take the form of a commercial lease, for example. Third, they must complete a one-page form, known as Form M2, stating both the former and new addresses. This form is depicted in Appendix Figure A.1. Relocation data, described in Section 2.2, are derived from this form.

Due to the formal reporting process required for a change of address, we believe there is minimal scope for misreporting relocations. First, an entrepreneur relocating an establishment has no incentive to declare it as a new establishment instead, as doing so would involve more bureaucratic procedures than simply reporting a change of address. Second, it is unlikely that a new establishment would be mistakenly classified as a relocation, given the requirement to provide proof of closure at one location and reopening at another within three months. In Section 3.3, we show that establishments retain the same activity and organizational structure after relocating, further supporting the notion that the same establishment is being observed in a new location.

2.2 French administrative and survey data

We combine two administrative data sources to track the mobility of establishments and workers over space: a registry of establishment relocation and linked employer-employee data. We complement these data with a survey of entrepreneurs to provide qualitative evidence on plant mobility.

Registry of Establishment Relocation. Measuring establishment relocation is often challenging. In many countries, establishment identifiers are location-specific, meaning that any change in location results in a new identifier. To address this issue, we leverage administrative data that track establishment relocations from the SIRENE register. These data originate from Form M2, which entrepreneurs complete to update the address on their business license (see Section 2.1). INSEE collects these forms to maintain an up-to-date registry of establishments.

For each year since 1993, the register provides a link between the establishment identifier at the destination and the initial location. We exploit this unique feature of the data to gain new insights into plant relocations. The register encompasses all types of establishments, including headquarters, research facilities, and production plants. It covers firms operating in the private sector, with the exception of those in agriculture and finance.

Linked Employer-Employee Data. We use comprehensive employment records at the job spell level (DADS Postes). These data are based on mandatory employer declarations for employees subject to French payroll taxes. They include information on employees' wages (gross and net) and the number of hours worked during the year. They also provide some characteristics of the employees (e.g., age, gender, place of residence, place of work) and of the companies and establishments for which they work (e.g., industry code and location at the municipality level). Importantly for our study, the data are reported at the establishment level, allowing us to track worker mobility between establishments and locations. The data are highly granular, as France is divided into approximately 35,000 municipalities ($\approx$ 40% of the total in Europe). We leverage this granularity to distinguish mobility within and across commuting zones. We consider the 304 commuting zones for metropolitan France as defined by INSEE in 2010.[4]

The raw data consist of a series of cross-sections that include all workers observed in year t along with information on the jobs they held in $t-1$. Using the algorithm developed by Babet et al. (2022), we link these cross-sections to construct a panel of workers covering the period 2002–2019. This enables us to track a large number of workers as they move between establishments and locations. For comparison, prior research has relied on a

[4]We discuss the use of commuting zones as our definition of local labor markets in Section 5.1.

restricted panel version covering approximately $1/25^{th}$ of the workforce (see, for example, Abowd et al. (1999)). We focus on employees aged 23 to 60 working in metropolitan France and retain only their highest-paying job during the year, determined by total gross earnings. For our main decomposition of interest, we use the hourly wage—obtained by dividing annual earnings by annual hours—converted to real terms for 2018.

For a subset of results, we also use firms' balance sheet data (Ficus-Fare), which allow us to compute the stock of capital. These data are reported at the firm level, and we restrict our analysis to single-establishment firms when utilizing them.

Survey of Entrepreneurs. The SINE survey provides detailed information on the characteristics of entrepreneurs, the development of the firm's activities, and its clientele. It also contains insights into potential challenges and shocks faced by entrepreneurs both before starting their business and during the initial years of operation (e.g., changes in local competition or difficulties in accessing credit). It consists of three waves: an initial interview and two follow-up questionnaires in years three and five. It is administered to a random sample of new entrepreneurs, covering approximately one-third of all firms created in the first semester of the survey year. In 2014, the survey included around 24,000 enterprises. We focus on companies that remained active at the time of the final interview and link this survey to our relocation register using a unique firm identifier. In section 3, we compare firms that relocated at least one establishment within their first five years of activity to those that did not relocate any activity.

2.3 Measures of establishment relocation in the United States

We provide additional evidence on firm relocation in the United States. We use establishment-level commercial data on the entire economy from Dun & Bradstreet, and firm-level data on headquarters of publicly-listed firms from firm filings to the US Securities and Exchange Commission (SEC) and Compustat data.

Dun & Bradstreet. The Dun & Bradstreet Historical Records provide comprehensive establishment-level data on both public and private companies dating back to 1969. The dataset includes information on establishment name, location, industry, creation date, employment, and firm linkages—specifically, headquarters identifiers for establishments within multi-unit firms, and parent identifiers for subsidiaries.

Unlike our data for France, relocation events are not explicitly reported. In principle, one can leverage the longitudinal nature of the data to identify cases where the same establishment identifier appears in different locations over time. However, establishments may change identifiers when relocating and apparent relocation events might instead reflect reporting errors by firms or significant changes in an establishment's iden-

tity and operations, such as mergers or acquisitions. Moreover, within multi-unit firms, distinguishing branch relocations from branch exits and new entries remains challenging.

Therefore, we adopt the following measurement approach. First, we focus exclusively on single-establishment firms or headquarters of multi-establishment firms, as these account for the vast majority of relocation events in the French data. We then define relocation events at the firm level, identifying firms based on their name, creation year, industry, and ownership structure. We restrict our analysis to firms observed for at least three consecutive years and require these identifying characteristics to remain constant throughout the entire period. Finally, we define a relocation event as a single change in the firm's location. Since street addresses are reported with considerable noise in these data, we define a change in zip code as a relocation.

Data on U.S. public corporations. We further complement our analysis using relocation events for publicly listed firms. Specifically, we obtain firms' annual 10-K reports from the SEC EDGAR (Electronic Data Gathering, Analysis, and Retrieval) system for the years 1993–2021 and extract each firm's headquarters location from its mailing address. Importantly, this mailing address captures the firm's administrative headquarters location, distinct from the incorporation address, which typically does not correspond to a site of physical operations. We then merge this information with firm-level balance sheet data from Compustat, restricting our sample to firms reporting positive employment to ensure meaningful economic activity. Firms are identified using their unique GVKEY code from Compustat, and we define relocation as a change in the address reported in consecutive annual 10-K filings.[5]

3 Characterization of firm relocation

While worker mobility—both between firms and across locations—has been widely studied in the literature, the use of firm relocation as a research focus is new. For this reason, we provide a detailed characterization of firm mobility in this section. While we discuss and test identification in section 5.2, we first describe what firm mobility implies in practice. We proceed in three steps. First, we assess the prevalence of firm mobility. Second, we examine the geographic patterns of relocations. Third, we investigate potential changes and adjustments in establishments after relocating. We document these patterns for France and, where data allow, for the US.

[5]While Compustat provides current addresses for firms, it does not retain historical addresses, making it insufficient for directly identifying relocations over time. Thus, address information from the annual 10-K filings is essential to identify firm relocations.

3.1 How prevalent is firm relocation?

We begin by documenting the prevalence of firm relocation. Figure 1 presents the time series of establishment relocations in France and the US between 1994 and 2023.[6] For each year, we compute the ratio of establishments that relocated to the total number of establishments operating in the economy. In France, the relocation rate remained relatively stable between 1994 and 2007, averaging around 4.0%. This rate declined to 3.5% during the Great Recession and has remained stable at that level since. Headquarters exhibit a slightly higher propensity to relocate, with rates between 4.0% and 4.5%. In the US, relocation rates are of similar magnitudes, with 2–5% of headquarters changing zip codes in both the Compustat sample and the D&B data. As expected, changes of address—which we can compute for the Compustat sample—occur slightly more frequently, at around 5–7%.

Establishment relocations occur across all industries and establishment size categories, though at varying rates. Figure B.1 illustrates the prevalence of relocation by industry. In France, the business services sector exhibits the highest relocation rates, averaging between 5.0% and 6.0% over the period (consistent with Duranton and Puga (2001) who focus on this industry).[7] In contrast, relocation rates in manufacturing, retail, and household services are approximately 3.0% per year. The hospitality and agrifood industries display the lowest relocation rates, though even these sectors experience an annual relocation rate of about 1.0%. Figure B.2 shows that establishments undergoing relocation employ approximately 2.5% of the French workforce.

3.2 Geography of firm relocation

We now examine the geographic patterns of firm relocation, addressing two key questions: (1) Where do establishments relocate? (2) Which establishments move where?

Relocation Distances. Panel (b) of Figure 1 shows that in both France and the US, approximately 20-25% of all establishment relocations involve a change in commuting zone (CZ).[8] To further investigate the role of distance, we estimate a gravity model:

$$log(\#\ \text{moves}_{ij}) = \eta_i + \mu_j + \delta\ log(\text{Distance}_{ij}) + \epsilon_{ij}, \tag{1}$$

[6] For the US, we compute several measures using our two datasets. With the D&B data, we only consider changes in zip codes (see Section 2.3), which likely underestimates relocations. In the US, relocations refer specifically to headquarters.

[7] This industry includes: telecommunications, real estate, IT, R&D, and services provided to companies such as consulting or marketing.

[8] Figure B.3 presents similar time series using different definitions of geographic boundaries. 65% of relocations involve changing cities.

where the dependent variable is the log of the number of establishments relocating from province i to province j. η_i and μ_j are origin and destination fixed effects, respectively. $Distance_{ij}$ is the number of kilometers between provinces i and j. Appendix Figure B.4 depicts the results for France and the US. We estimate negative elasticities of -0.96 for France, and -0.42 for the US, suggesting strong spatial frictions. In France, the average distance for an establishment relocation is 151 kilometers (median: 42 km). Appendix Figure B.1 compares establishment relocation distances with those of workers. For workers who relocate, the average distance is 178 kilometers (median: 69 km). Appendix Figure B.5 plots the distribution of relocation distances for establishments, highlighting the existence of relocations spanning several hundred kilometers.

Symmetry of flows. We investigate whether establishments systematically relocate to the same locations. For each pair of locations, we compute the following index:

$$\text{Index}^{\text{Directionality}} = \frac{|\#\text{ Outflows} - \#\text{ Inflows}|}{\#\text{ Outflows} + \#\text{ Inflows}}, \tag{2}$$

where # Outflows represents the number of establishments relocating from location i to j, and # Inflows denotes relocations from j to i. This index measures the asymmetry between gross and net flows, indicating the extent of unidirectional movement between two locations. An index value of zero implies perfectly symmetric bilateral flows. Figure 2 shows the distribution of $\text{Index}^{\text{Directionality}}$ for France and the US. The large mass of observations near zero suggests that gross flows are much larger than net flows—indicating that, for instance, there are roughly as many relocations from Paris to Marseille as there are from Marseille to Paris. The symmetric nature of these flows suggests that relocation decisions are not primarily driven by shared location-specific factors that would make certain destinations systematically more attractive.

Sorting associated with relocation. Having documented where establishments relocate, we now examine which types of establishments move to which destinations. Specifically, we compare the characteristics of establishments relocating from the same origin but to different destinations. To do so, we estimate the following specification:

$$x_{c(J)} = \eta_i + \gamma_1 x_J + \gamma_2 Age_J + u_J, \tag{3}$$

where x_J represents the characteristics of the establishment at the time of relocation (e.g., log size) and $x_{c(J)}$ denotes the characteristics of the destination commuting zone, measured before the move. We control for establishment age and commuting zone of origin, η_i. Figure 3 presents the results for establishment size in France and the US (Panel (a)) and for hourly wages in France (Panel (b)). First, we find that larger establishments relocate to larger commuting zones, with elasticities of 0.086 (standard error: 0.008) for France

and 0.057 (standard error: 0.014) for the US. Second, higher-paying establishments also relocate to commuting zones where wages are higher on average. Using the French data, we estimate an elasticity of 0.023 (standard error: 0.002). These correlations suggest that establishment relocation contributes to the sorting of higher-paying establishments into higher-paying locations, though at this stage we do not control for worker composition.

3.3 What happens to an establishment that relocates?

We now analyze the potential changes in an establishment following a relocation in France. We consider two types of adjustments. First, we examine structural changes related to the establishment's activity and organization. Second, we analyze labor adjustments in response to operating in a new labor market.

3.3.1 Activity and organization

We combine administrative data with a survey of entrepreneurs to assess whether relocating establishments implement structural changes beyond their geographic move. This analysis allows us to evaluate the persistence of establishment characteristics through relocation.

First, using administrative data, we construct several proxies for the type of activity performed by an establishment. Table 1 presents the results by mover status. Following a relocation, 94% of establishments remain in the same industry, 99% retain their headquarter status, and 96% continue under the same legal category (e.g., simplified corporation or limited liability company). These figures suggest that relocation does not involve significant changes in business activity or organizational structure. Entrepreneurs must report updated information when completing the change-of-address form, making it more likely that changes are recorded for relocating establishments. To proxy the identity of the top manager, we use the highest-paid employee.[9] The last column of Table 1 indicates that the top manager remains the same in 67% of relocating establishments, compared to 77% of non-relocating establishments. Table B.2 reports a difference-in-differences analysis, controlling for potential differences between movers and non-movers, as well as year fixed effects. The results are similar in magnitude for all the variables, confirming that only a small fraction of establishments undergo significant changes after relocation.

Second, we use linked employer-employee and balance-sheet data to track the evolution of key input shares after relocation. Figure 4 plots the correlation between input shares in year t and year $t+1$, separately for establishments that relocated between t and $t+1$ and those that did not. We examine the share of low-skill workers among employees (Panel (a)) and the capital-to-labor ratio (Panel (b)). For movers, we estimate slopes of

[9]This approach may fail to correctly identify the top manager if that person is not an employee of the establishment.

0.924 and 0.922, respectively, indicating that input shares remain nearly unchanged after relocation.

Our findings from administrative data indicate that relocating establishments do not significantly alter their operations or the type of goods or services they produce. To complement these results, we analyze data from the SINE survey, which asks entrepreneurs about changes implemented since firm creation and challenges encountered (see Section 2.2). We match this survey with administrative relocation data to compare responses from entrepreneurs who relocated their establishments with those who did not. Figure 5 presents responses to the questions: "Did you experience any major change since the creation of the firm?" and "When you set up your company, did you encounter any difficulties?". The responses are strikingly similar for movers and non-movers. Among both groups, 57% of entrepreneurs report no major changes since firm creation, and 39% report no difficulties in the first two years. Movers and non-movers also exhibit similar probabilities of having changed the types of services offered (15%), the location of their clients (9%), the number of clients (31%), and the types of clients (11%). Beyond confirming that establishments do not implement more changes when they relocate, these results suggest that the decision to relocate is not driven by external shocks. For instance, movers and non-movers are equally likely to report problems related to labor force constraints (7%) or competition (27% for movers and 29% for non-movers) during the first two years. Additional results from the survey, presented in Appendix Figure B.6, further support the absence of specific shocks as primary drivers of relocation decisions.

3.3.2 Labor adjustments

Although establishments do not appear to undergo structural changes after relocation, we now examine whether they adjust their labor input. Specifically, we investigate whether they begin employing workers from the new location and whether the wages paid by the establishment respond to local wage levels.

First, we find that half of the workers employed in the establishment before relocation continue working there in the year following the move. Appendix Figure B.7 presents the distribution of the share of workers retained after relocation (Panel (a)) and the workforce composition based on workers' place of residence (Panel (b)). Prior to relocation, establishments already employ workers from the destination commuting zone, representing 20–25% of the workforce. Immediately after relocation, this share increases to 50%, as workers from the commuting zone of origin are replaced by those from the new location. This indicates that a significant portion of an establishment's workforce is replaced by local workers following a move.

Given that both location and workforce composition change with relocation, we now examine how wages evolve within establishments. Specifically, we test whether estab-

lishments relocating to higher-wage areas increase their wages. To systematically assess this, we compare wage trajectories based on the wage gap between the destination and origin commuting zones. We estimate the following regression, adapting the specification of Badinski et al. (2023) for the healthcare sector to our context:

$$log(wage_{Jt}) = \phi_J + \delta_t + \theta_{r(J,t)}\Delta_J + \beta x_{Jt} + \nu_{Jt}, \tag{4}$$

where the outcome variable is the log of the average hourly wage in establishment J in year t. ϕ_J and δ_t are establishment and year fixed effects, respectively. Δ_J represents the difference in average log hourly wages between the destination and origin commuting zones, measured in the year of relocation. x_{Jt} consists of indicators for years relative to establishment J's relocation, with $r(J,t)$ denoting the relative year of the move. The main parameters of interest, $\theta_{r(J,t)}$, correspond to the interactions between relative year indicators and the wage gap between the destination and origin commuting zones. This captures how an establishment's average wage evolves in the years before and after relocation, relative to the wage difference between its new and original location.

We estimate two versions of Equation (4). In the first, we include both movers and non-movers, setting the relative year and Δ_J to zero for non-movers. In the second, we focus exclusively on relocating establishments and omit the relative year fixed effects (x_{Jt}). In both cases, we use a balanced panel of establishments to ensure that composition effects do not bias our estimates.

As a benchmark, we also conduct this event study for workers relocating across space. In this case, the dependent variable is the log hourly wage of the worker, and Δ_J still represents the wage gap between destination and origin commuting zones.[10]

Figure 6 presents the results for establishment relocations (Panel (a)) and worker relocations (Panel (b)). In both cases, wages remain stable in the four years preceding relocation, increase precisely in the year of the move, and remain relatively stable afterward (with some modest post-move dynamics for workers). For establishments, relocating to a commuting zone with wages 1% higher is associated with a wage increase of 0.15%, whereas the elasticity for worker relocations is 0.25. The fact that this coefficient differs from one highlights the significant role of establishments in wage setting. Meanwhile, the coefficient being different from zero indicates that the location also affects wages—either through the effect of the local labor market or through the composition of workers employed after relocation. In Section 5, we further disentangle the contributions of the local workforce and the location itself.

[10] Appendix Figure B.8 reports the results for workers changing employers, where Δ_J corresponds to the wage gap between the new and former firms.

3.4 Summary

Firm mobility is prevalent and can be systematically tracked in our data. The persistence of establishment characteristics after relocation confirms that the same establishment is observed before and after a move. Furthermore, relocation decisions are not driven by specific constraints or problems faced by entrepreneurs.

Our results support the use of establishment mobility as a setting to study spatial wage disparities. We are able to observe the same establishment—with the same activity and input composition—operating in different locations and assess how wages vary across these locations. When combined with worker mobility data, this approach allows us to disentangle the contributions of workers, establishments, and locations to wage disparities.

4 A model of firm relocation

In this section, we develop a simple model of firm location decisions, designed to serve two primary purposes. First, it offers a stylized framework that can rationalize the key patterns in firms' relocation decisions, as documented in Section 3. Second, it provides an example of a theoretical setting that aligns with the assumptions underlying our wage decomposition exercise in Section 5. It also offers a possible interpretation of our empirical estimates of worker, location, and firm effects presented in Section 6. To this end, we develop a model with worker and firm heterogeneity and spatial sorting.

4.1 General Setting

We consider a static model with N discrete locations. The economy consists of a unit mass of firm owners and a measure L of workers. Each agent is endowed with an initial location $o \in \{1...N\}$. Workers born in location o can choose to relocate to a new location $n \in \{1...N\}$. Similarly, firm owners who initially establish their businesses in location o can opt to operate from a different location n. All agents consume freely traded goods.

4.2 Workers

Each worker is endowed with an innate ability a, representing the efficient units of labor they supply to firms. This ability is drawn at birth from a cumulative distribution function (CDF) $F_A(.)$. Let $w_n(a)$ denote the expected wage for a worker of type a if they choose to live in location n. Upon birth in location o, each worker selects the location n that maximizes:

$$\max_{n \in \{1...N\}} w_n(a) - \kappa^w_{on} + \epsilon^w_n,$$

where κ_{on}^{w} represents the bilateral migration friction between locations o and n, and ϵ_{n}^{w} is an idiosyncratic preference shock for location n. We assume that these shocks are drawn from a Type-1 extreme value distribution with shape parameter ξ^{w}.

4.3 Firm owners

Each firm owner is endowed with productivity level z, drawn upon entry from a CDF $F_{Z}(.)$. Let $\pi_{n}(z)$ denote the expected profits of a type-z firm when operating from location n (see Appendix C.4 for the derivation). A firm owner who starts their business in location o chooses the location n that maximizes:

$$\max_{n\in\{1...N\}} \pi_{n}(z) - \kappa_{on}^{f} + \epsilon_{n}^{f},$$

where κ_{on}^{f} is a bilateral relocation friction between locations o and n, and ϵ_{n}^{f} is an idiosyncratic preference shock for location n. These shocks are also drawn from a Type-1 extreme value distribution, with shape parameter ξ^{f}.[11]

4.4 Production

The output of a firm is the sum of outputs across all of its workers. We assume that each worker's output is super-modular in their individual ability and the firm's productivity. Specifically, the output of type-a worker in an n-based firm of type z is given by

$$y_{n}(z,a,x) = \Phi_{n} z a x,$$

where Φ_{n} captures a productivity component shared by all n-based firms, and x is an i.i.d match-specific productivity shock, drawn from a CDF $F_{X}(.)$ with mean 1. We allow Φ_{n} to potentially depend on the number of local workers and firms, thus capturing agglomeration forces in addition to exogenous regional fundamentals. For instance, regional productivity may follow the standard functional form from the agglomeration literature, $\Phi_{n} = \bar{\Phi}_{n} L_{n}^{\alpha}$.[12]

4.5 Labor markets

The local labor market is frictional and characterized by random search. A firm posts v vacancies in its location and incurs a vacancy-posting cost $H(v)$. Each posted vacancy is matched with a random worker at a market-specific matching rate q_{n}. Under standard

[11] In our model, a relocation results from a potential trade-off between profits and idiosyncratic preferences. We do not make assumptions on which of the two is the main driver. The prevalence of bilateral gross relocation flows in the data aligns with such interpretation.

[12] In this formulation, regional productivity consists of an exogenous term $\bar{\Phi}_{n}$, and increases with local employment density (L_{n}), by a magnitude governed by elasticity α.

assumptions about the matching function, this rate is given by $q_n = \vartheta_n^{-\eta}$, where ϑ_n denotes local market tightness—the ratio of total vacancies posted by firms selecting market n to the total number of workers choosing market n—and η represents the elasticity of the matching function with respect to the number of workers choosing market n. Accordingly, the share of job-seekers matched with vacancies is $p_n = \vartheta_n^{1-\eta}$.

When a firm meets a job-seeker, it can decide whether or not to hire and train the worker. The cost of hiring and training, c, is randomly drawn from a distribution with CDF $F_C(.)$. When making this decision, the firm observes the worker's type a, but not the idiosyncratic match-specific productivity shock x, which becomes known only at the production stage. After choosing to hire and train the worker, output is produced, and the firm pays a constant share $1-\beta$ of output as a wage to the worker. This wage-setting arrangement can represent, for instance, a standard Nash bargaining protocol, as both parties' outside options after the sunk cost c has been paid are zero.[13] Under these assumptions, the probability that a firm of type z hires a worker of type a, conditional on a match, is given by $\mathcal{P}_n(z,a) = F_C(\beta y_n(z,a))$, where $y_n(z,a)$ is the expectation of output, integrating over the match-specific shock x.[14]

4.6 Model implications

We now summarize the key implications from the above setting, leaving most derivations for Appendix C. We highlight that this setting aligns with the key facts about firm relocation flows from Section 3, and with the main assumptions that we will impose in the wage-decomposition exercise in Section 5.

Proposition 1. ***Characteristics of firm relocation and wages in the model.*** *We denote by $s_{on}^f(z)$ the share of type-z firms that start in location o and move to location n, and by $M_{on}(z)$ their mass.*

(a) ***Gravity structure for firm relocation rates:*** *the share of type-z firms that start in location o and choose location n admits a log-linear structure that depends on a destination fixed effect, an origin fixed effect, and bilateral relocation frictions:*

$$\log s_{on}^f(z) = \underbrace{\frac{1}{\xi^f}\pi_n(z)}_{\text{Destination FE}} \underbrace{- \log \sum_{n'=1}^{N} \exp\left(\frac{1}{\xi^f}\left(\pi_{n'}(z) - \kappa_{on'}^f\right)\right)}_{\text{Origin FE}} - \underbrace{\frac{\kappa_{on}^f}{\xi^f}}_{\text{Bilateral frictions}}.$$

(b) ***The model features gross bilateral relocation flows:*** *if κ_{on}^f and κ_{no}^f are finite, then $s_{on}^f(z) > 0$ and $s_{no}^f(z) > 0$ for all o, n, z. In addition, the ratio of net-flows*

[13] We abstract from an explicit modeling of workers' outside options.

[14] The firm decides whether to pay c knowing it will get a share β of the job surplus afterwards, as c is sunk by the time of bargaining.

to gross-flows between locations o and n for a particular firm type z, $\frac{|M_{on}(z)-M_{no}(z)|}{M_{on}(z)+M_{no}(z)}$, converges to one when the dispersion of the idiosyncratic firm shocks, ξ^f, converges to zero.

(c) ***Firm spatial sorting:*** *suppose that the inverse of the hiring and training cost, c, is distributed Pareto, and that the cost of posting vacancies is a power function: $H(v) = \frac{v^{1+\delta}}{1+\delta}$. Then $s^f_{on}(z)$ is log-supermodular in the firm's productivity z and the location productivity Φ_n. Conditional on the origin location o, higher productivity firms select locations with higher productivity (Φ_n), higher human capital ($\mathcal{A}_n$) and higher vacancy matching rate (q_n).*

(d) ***Worker spatial sorting:*** *suppose that the inverse of the hiring and training cost, c, is distributed Pareto. Then the probability that a worker of type a from location o chooses location n, $s^w_{on}(a)$, is log-supermodular in the worker's ability a and the location productivity Φ_n. Conditional on the origin location o, higher ability workers select locations with higher productivity (Φ_n), higher average firm productivity, and higher worker matching rate (p_n).*

(e) ***Within-location sorting:*** *furthermore, suppose that the inverse of the hiring and training cost, c, is distributed according to the Generalized Pareto Distribution, with location μ, scale σ, and shape $\zeta \neq 0$. Then, if $\mu > \frac{\sigma}{\zeta}$, the probability that a random match is turned into an employment relationship is log-supermodular in the firm productivity and the worker's ability: $\frac{\partial^2 \log \mathcal{P}_n(z,a)}{\partial z \partial a} \geq 0$. In this case, within each location n, higher productivity firms employ, on average, higher ability workers.*

(f) ***Log-linear wage structure:*** *consider a worker of type a that is employed by a firm of type z in location n. The wage paid to this worker is given by:*

$$(1-\beta)\,\Phi_n z a x.$$

The predictions of our model align closely with observed patterns of worker and plant mobility in France and in the US, as outlined in Section 3. Specifically, propositions (a), (b), and (c) are consistent with the gravity structure depicted in Figure B.4, the large bilateral gross flows in Figure 2, and the sorting of higher fixed effect firms into locations with higher average worker effect and somewhat larger location effects (see Figure D.3). Finally, proposition (e) aligns with the assortative matching documented in the AKM literature (see for example Card et al. (2013); Song et al. (2019)).

While our empirical analyses does not require the parametric assumptions made to obtain the wage equation, having this model in mind is useful as a suggestion of interpretation of our fixed effects. It also shows that this log-linear wage equation is consistent with a framework featuring positive sorting (between and within locations).

5 Empirical model and identification

In this section, we introduce the empirical model we use to estimate the contributions of workers, establishments and locations to spatial wage disparities and discuss identification.

5.1 Statistical model

We leverage both establishment mobility across space and worker mobility (between establishments and between locations) to estimate the contributions of worker heterogeneity, establishments, and locations to spatial wage disparities. Our main analysis is based on the following model:

$$Y_{ict} = \alpha_i + \phi_{J(i,t)} + \psi_c + \mathrm{X'}_{it}.\beta + \epsilon_{ict}, \tag{5}$$

where Y_{ict} is the (log) gross hourly wage of employee i, working in plant $J(i,t)$, located in commuting zone c, observed at time t. The term α_i represents a worker fixed effect, capturing time-invariant characteristics that influence worker i's compensation (e.g., ability).

The establishment-specific component of earnings, $\phi_{J(i,t)}$, reflects differences in productivity and pay policy between establishments. We assume that the establishment component of wages, ϕ, remains unchanged after the move. This implies, for instance, that establishments do not use relocation as an opportunity to alter their type of activity in a way that would affect wages. Our assumption is supported by the findings in Section 3.3, which show that establishments maintain the same activity type, legal status, headquarter status, and input composition. Moreover, only a change in ϕ which is systematically correlated with the location effect would bias our results. This would lead to over-estimating the contribution of locations effects, relative to that of establishments. As our estimates indicate a small role of location effects, this is unlikely.

We consider a specification in which each establishment has its own fixed effect, as well as an alternative specification where establishments are clustered. Such clustering aims to correct for the limited mobility of workers between establishments (Bonhomme et al., 2019). Establishments are grouped into 10 clusters based on percentiles of the wage distribution using a K-means algorithm.

The location fixed effect, ψ_c, captures compensation differences between locations while controlling for worker and establishment characteristics. A positive ψ_c implies that similar workers and establishments experience higher wages in this location compared to the average location. As discussed in Section 4, this effect accounts for differences in infrastructure, agglomeration forces, and local competition. In our main analysis, we define location at the level of the commuting zone. Commuting zones provide a

relevant definition of local labor markets, as they are determined based on commuting flows between place of residence and place of work.[15] In France, commuting zones are delineated by the Statistical Institute (INSEE) and the Ministry of Labor (DARES), following Eurostat guidelines. One key criterion in their definition is that at least 60% of workers in a given zone must both live and work there.

Finally, the model includes time-varying controls, X_{it}, which in our main specification consist of a polynomial of workers' age and year fixed effects.

In Section 6.2, we present evidence supporting our specification choices and we provide a sensitivity analysis.

5.2 Identification and assumptions

Identification: Our strategy relies on observing the same worker employed in different locations and establishments, as well as the same establishment operating in different locations. The model is identified using (i) worker mobility between establishments and commuting zones and (ii) establishment mobility between commuting zones. This two-sided movers design is illustrated in Figure 7. A standard two-way fixed effects model does not allow for the separation of establishment and location effects. When a worker moves between two commuting zones, the effects of both the new establishment and the new location are jointly identified (ϕ and ψ change simultaneously). To address this, we exploit establishment mobility between locations as a second source of identification. This additional variation allows us to compare the same worker in the same establishment across different locations.

More precisely, two features of the data enable us to separately identify establishment from location effects. First, we use observed wage changes for workers who relocate with their establishment. For instance, in Figure 7, Firm B relocates from location α to location β while retaining some of its employees. In this case, the only change for those workers is the location. However, this first source of identification may be affected by wage rigidity for incumbents, as workers may not experience a wage decrease even if their establishment relocates to a lower-paying location.[16] To address this potential concern, we leverage a second source of identification that does not rely on incumbent workers moving with their establishment. Instead, we exploit worker mobility that connects establishments within both the origin and destination locations. In Figure 7, some workers transition to or from Firm B when it is in location α, while others connect Firm B to different employers after it relocates to location β. The location effect is identified as

[15] The definition of commuting zones has been criticized as an inadequate representation of local labor markets in the U.S (Bartik, 2024). In contrast, commuting zones in France are defined using a different methodology with stricter commuting flow requirements, making them a more precise reflection of local labor markets.

[16] The concern that wage rigidity may bias our estimates is mitigated by the legal requirement that all employment contracts must be renegotiated for relocations exceeding 10 kilometers.

long as Firm B is connected to the rest of the sample through mobility in both locations. In our main estimation, we combine these two sources of identification. As a robustness check, we exclude workers who move with their establishments to rule out potential biases due to wage rigidities. This alternative specification produces nearly identical results. We now discuss the main identification assumptions.

Additive separability: Following a large body of literature, we assume the additive separability of log wages, meaning that wages can be expressed as a sum of fixed effects. In Section 4, we obtain such an equation in a search and matching model with worker and firm mobility. Empirically, the high adjusted R^2 of 0.88 suggests that this specification provides a good approximation of the data.

Conditional exogenous mobility: Conditional on fixed effects (i.e., workers' ability, plants' productivity, and local conditions), mobility must be as good as random. This assumption allows for different types of establishments to move to different types of locations, or for different types of workers to sort into different establishments and locations. However, it requires that mobility is not driven by time-varying unobservables (ϵ_{ict}). In practice, this assumption excludes two scenarios: (i) mobility induced by unobserved shocks that directly affect wages, and (ii) mobility driven by match effects—i.e., establishments (or workers) relocating to a particular location (or establishment) because they expect to benefit more than average from that location. We implement three main tests to assess the validity of the conditional exogenous mobility assumption.

First, we use event studies to analyze the evolution of wages before, during, and after a move. Following the approach of Badinski et al. (2023), Figure 6 and Appendix Figure B.8 show estimates of Equation (4) and plot the coefficient associated with moving to a higher-paying establishment and location respectively. For both worker and establishment relocation, as well as job changes, there is a distinct wage change precisely in the year of the move, with no pre-trends and minimal post-move dynamics. The results remain consistent when we focus only on movers or when we also include non-movers. These event studies support the absence of wage-related shocks driving mobility, ruling out concerns such as Ashenfelter dips. We further refine this analysis by considering event studies separately by estimated fixed effects quartiles (following Card et al. (2013, 2016)). Figure 8 presents the evolution of log hourly wages (residualized for age and year fixed effects) for three types of mobility: (a) workers changing establishments within the same location, (b) workers changing both establishment and location, and (c) establishments changing location. Wages are plotted separately by quartiles of origin and destination fixed effects, estimated from Equation (5). For workers moving across both jobs and locations, quartiles are based on the sum of establishment and location fixed effects. For establishments changing location, quartiles are based on the sum of location fixed effects

and the average worker fixed effects in that location. These results yield three key observations. First, wages were flat in the years before the move across all quartile-to-quartile transitions, and they are flat again post-move, supporting the assumption that observed wage changes are a direct consequence of relocation rather than unobserved shocks. Second, when moves involve a change in quartile of fixed effects, wages adjust immediately, while they remain stable for moves within the same quartile. Moreover, the magnitude of wage changes is close to symmetric: establishments moving from Q1 to Q4 experience wage increases of a similar magnitude as wage decreases for establishments moving from Q4 to Q1. This symmetry supports the absence of match effects driving relocation or job changes.[17] Third, while wage changes for workers switching establishments are similar regardless of whether they move within or across commuting zones, the magnitude of wage changes for establishments relocating is smaller, suggesting greater heterogeneity in establishment effects than in location effects.

Second, our framework allows for a novel placebo test: we examine wage evolution for establishments that relocate *within* the same commuting zone. According to our statistical model, these establishments should not experience wage changes as they retain the same location effect and continue to access the same labor market. On the contrary, wages would change if relocation is primarily driven by unobserved shocks or if the process itself impacts wages. Panel (c) of Figure 8 confirms that average establishment wages remain remarkably stable at the time of the move, reinforcing our interpretation that location and worker composition drive observed wage changes for those relocating across commuting zones.

Third, we leverage the survey of entrepreneurs to assess whether relocating establishments systematically face specific challenges or shocks. In Section 3, we present evidence that movers do not undergo more changes than non-movers. Panel (b) of Figure 5 shows that establishments that relocated were no more likely to experience operational difficulties in their first two years than those that did not relocate. The reported challenges—including workforce-related issues, production costs, and financial constraints—are strikingly similar across movers and non-movers, with approximately 40% of both groups reporting no difficulties. Appendix Figure B.6, panel (a), further shows that both groups report comparable changes in local competition, with 60% indicating no change. Panel (b) confirms that movers and non-movers faced similar obstacles when opening their establishments. These results suggest that establishment relocations are not primarily driven by shocks or binding constraints.

[17] We also confirm the expected ordering consistent with prior research on the AKM framework: individuals moving from high to low quartiles were initially paid less than those moving to relatively higher quartiles.

Timing. We do not model the fact that the effect of changing firm or location might take time to fully materialize, for instance because of learning (De la Roca and Puga, 2016). Our estimates stem from the average wage change observed in the years after a job change or a relocation, as we include all subsequent observations. However, our evidence suggests that dynamic effects are mild. We have estimated dynamic event studies based on Equation (4). Figure 6 shows that for both establishments relocating and workers moving to new locations, most wage effects materialize immediately in the year following the move, with limited subsequent dynamics. Specifically, for relocating workers, the wage coefficient rises from 0.23 immediately after the move to 0.27 four years later, indicating some dynamic effects, but those are modest compared to the overall effect. Similarly, Appendix Figure B.8 shows that for workers changing firms—both within and across commuting zones—the effect is largely immediate.

5.3 Sample and connected set

We estimate Equation (5) on the universe of companies operating in the private sector over the period 2002–2016, focusing on establishments that employ at least one full-time worker during the year. Table 2 presents summary statistics for establishments and workers by mover status. On average, establishments that relocate have 11 employees, compared to 12 for non-movers. Their share of high-skill workers is slightly larger, and industry composition remains similar across both groups.

To ensure identification, we restrict our analysis to the largest connected set of workers, establishments, and locations, excluding singletons. This connected set comprises more than 161 million worker-year observations, covering approximately 1.8 million unique establishments, 22.7 million unique workers, and 304 commuting zones. The relatively small number of commuting zones, combined with our large sample size and the high number of moves between CZs, ensures strong connectivity. Appendix Table D.1 reports the distribution of the number of unique workers and establishments per CZ (averaging 98,920 and 5,937, respectively), as well as the number of moves for workers and establishments (averaging 28,750 and 287, respectively).

Existing studies on wage setting emphasize that the limited mobility of workers between firms can bias estimates of the variance (Abowd et al., 2004; Andrews et al., 2008; Kline et al., 2020). However, this issue is less concerning in our case for two reasons. First, we observe a large number of unique observations per CZ and a substantial volume of worker and establishment moves. Second, we decompose differences across only 304 commuting zones. As explained in Section 5.4, we collapse estimated fixed effects at the CZ level before computing the variance, which helps mitigate noise.

To further verify that our results are not affected by limited mobility bias, we implement two additional strategies. First, following Bonhomme, Lamadon and Manresa

(2019), we cluster establishments into 10 groups using a k-means algorithm based on percentiles of the within-establishment wage distribution. Second, we apply a variance correction method using a split-sample approach (Babet et al., 2022). Specifically, we estimate Equation (5) on two separate worker samples and compute variance adjustments. Details of this method are provided in Appendix D.2. The results are strikingly similar across the baseline model, the clustered version, and the split sample approach, confirming that limited mobility does not pose a significant concern.

5.4 Variance Decomposition Framework

Our objective is to decompose disparities in average hourly wages across commuting zones. This section outlines the decomposition method used to this purpose.

First, we compute the average of log hourly wages, the fixed effects estimated from Equation (5), the time-varying controls, and the residuals at the commuting zone level. Each average is weighted by the number of worker-year observations. Our interest lies in assessing the contribution of these components to the variation in wages across commuting zones. We rely on the following variance decomposition:

$$
\begin{aligned}
Var(\overline{log(Y_c)}) \;=\; & \underbrace{Var(\overline{\alpha_c})}_{\text{Workers contribution}} + \underbrace{Var(\overline{\phi_{J(c)}})}_{\text{Plants contribution}} + \underbrace{Var(\psi_c)}_{\text{Area contribution}} \\
& + \underbrace{2.cov(\overline{\alpha_c}, \overline{\phi_{J(c)}}) + 2.cov(\psi_c, \overline{\alpha_c}) + 2.cov(\psi_c, \overline{\phi_{J(c)}})}_{\text{“Sorting” of workers and firms}} \\
& + \underbrace{Var(\overline{X_c}) + 2.cov(\overline{X_c}, \overline{\alpha_c}) + 2.cov(\overline{X_c}, \overline{\phi_{J(c)}}) + 2.cov(\overline{X_c}, \psi_c)}_{\text{Demographic controls}} + \underbrace{Var(\overline{\epsilon_c})}_{=0}
\end{aligned}
\tag{6}
$$

The variance of average wages across locations can thus be decomposed into the variance of the average fixed effects between commuting zones and their respective covariances. We interpret the variance components as capturing (a) differences in intrinsic location productivity (ψ_c), and (b) differences in the local composition of workers and establishments ($\overline{\alpha_c}, \overline{\phi_{J(c)}}$). The covariance terms provide insight into the role of worker and establishment sorting between locations. A positive $cov(\psi_c, \overline{\phi_c})$ (respectively, $cov(\psi_c, \overline{\alpha_c})$) suggests that productive establishments tend to locate in high-productivity areas (respectively, high-ability workers sort into high-productivity locations). Finally, $cov(\overline{\alpha_c}), \overline{\phi_c})$ informs us on the degree of co-location of productive establishment and high ability workers, irrespective of the productivity of the location. In the following analy-

sis, we present the contribution of each component as a fraction of the total variance in average wages across commuting zones, $Var(\overline{log(Y_c)})$.

6 Results

6.1 Contribution of worker, plant and location effects to spatial wage disparities

We present the main results of the variance decomposition in Table 3. Column (1) reports the results from the baseline model in which each establishment has a fixed effect, while column (2) shows a version where establishments are clustered to account for the limited mobility of workers across establishments. The table provides the standard deviation of average log hourly wages between commuting zones, along with the standard deviations of worker, plant, and location effects, as well as time-varying controls and covariance terms (multiplied by two) following Equation (6).

First, nearly 50% of the variance in wages across commuting zones can be attributed to heterogeneity in workers and establishments. Worker effects ($\overline{\alpha_c}$) explain approximately 30% of the variance between CZs (31% in the clustered version), while establishment effects ($\overline{\phi_{J(c)}}$) account for around 17% (13% in the clustered version). In contrast, location effects *per se* contribute only 4.5% of the variance (2% in the clustered version). These results indicate that differences in location effects play a minimal role in explaining spatial wage disparities, with most variation stemming from worker and establishment heterogeneity. In Appendix Figure D.2, we plot average wages and estimated fixed effects on the map of France. The spatial distribution of worker and establishment fixed effects closely mirrors the distribution of gross hourly wages. However, location effects present a different pattern, highlighting geographically concentrated pockets of high productivity, particularly in the Paris region and the northern and northeastern quadrants. Conversely, regions such as Brittany (northwest) and the southwestern quadrant exhibit lower location effects.

Second, the sorting of workers, establishments, and locations accounts for the remaining half of the variance. The co-location of more productive workers and higher-paying establishments—independent of location—explains 33% of spatial wage disparities (40% in the clustered version). Additionally, we find that more productive workers tend to locate in higher location-effect areas (explaining 12% of the variance, 8% in the clustered version) and higher-paying establishments sorting into higher location-effect areas (5% of the variance, 7% in the clustered version). Appendix Figure D.1 illustrates the correlation between worker and plant effects (panel (a)), worker and area effects (panel (b)), and plant and area effects (panel (c)). The substantial role of sorting implies that, even though heterogeneity in location effects explains only a small share of spatial wage dis-

parities, the sorting of workers and firms across these heterogeneous locations amplifies the overall impact of those differences. Furthermore, the relative contributions of these components remain stable between the first and second halves of the study period (see Appendix Figure D.4).

In Section 3, we document that, on average, larger and higher-paying establishments relocate to larger and higher-paying commuting zones. To further disentangle the drivers of this sorting through relocations, we rely on the estimated fixed effects. Appendix Figure D.3 shows the correlation between the fixed effects of relocating establishments and the difference between destination and origin in (i) location fixed effects (panel (a)), (ii) average fixed effects of the local workforce (panel (b)), and (iii) average fixed effects of local establishments (panel (c)). These results indicate that sorting through establishment relocation primarily occurs through the local workforce and local establishments.

6.2 Robustness and heterogeneity analysis

We consider alternative versions of our model and sample. The resulting decompositions are presented in Table 4.

First, we include additional time-varying controls in our specification. In columns (2) and (3), we add a polynomial of the establishment's age to account for the fact that firm pay-setting practices can evolve over time (Brown and Medoff, 2003; Babina et al., 2019).[18] This addition does not affect our results.

Second, in columns (4) and (5), we exclude workers who remain with their establishment when it relocates and re-estimate our model. As discussed in Section 5.1, our parameters of interest remain identified through worker mobility between establishments and establishment mobility between locations. This analysis serves to ensure that our findings are not influenced by potential wage rigidity for incumbent workers or a wage premium compensating for increased commuting time. When using only the wages of new hires after establishment relocation, we find similar results, with location effects explaining 8% of wage disparities in the baseline version and 3% in the clustered version.

Third, we examine alternative clustering methods for establishments and different numbers of clusters. Column (6) reports the results when clustering establishments using their mean and standard deviation instead of percentiles. In column (7), we increase the number of clusters from ten to 100. Column (8) combines both modifications. While adding more clusters slightly increases the contribution of establishments and reduces that of locations, the overall magnitude of the different components remains largely unchanged.

Fourth, while the baseline model estimates all parameters jointly, we now adopt a two-step procedure. In the first step, we use a standard AKM framework, estimating worker

[18] For instance, young firms may back-load workers' compensation to mitigate financial constraints (Michelacci and Quadrini, 2005).

and establishment × location effects. In the second step, we decompose the establishment × location effect by leveraging establishment mobility across locations.[19] This approach produces results nearly identical to those of the baseline specification, shown in column (9).

Fifth, fixed effects are assumed to be time-invariant in our model, with data from 2002 to 2016 pooled together for estimation. This approach enhances the connectedness of the dataset. Consequently, our estimated location fixed effects can be interpreted as averages over the 2002-2016 period. As a robustness check, we also estimated our model separately for the first and second halves of the period. The contributions of the different components to spatial wage disparities remain almost identical across both periods (see Appendix Figure D.4).

Heterogeneity Analysis by Industry. We examine the potential heterogeneity of our results across major industries: manufacturing, lower-skill services (e.g., retail, accommodation, services to households), and higher-skill services (e.g., consulting, marketing, accounting). This allows us to assess whether agglomeration forces play a more significant role in certain sectors. Appendix Table D.2 presents the results of separate estimations by industry, where establishments are clustered. First, we observe that wage disparities between locations are larger in more skill-intensive industries. Second, while the contribution of location effects varies across industries, it remains consistently below 5%. Notably, location effects account for a slightly larger share of the variance in lower-skilled industries. Sorting patterns also exhibit heterogeneity. In manufacturing, there is stronger co-location of high-ability workers and high-paying establishments than in other industries, but little evidence of productive establishments and workers sorting into high-paying locations. In services, by contrast, more productive workers and higher-paying establishments are more likely to locate in high-fixed-effect areas.

6.3 Comparison with Alternative Approaches

In this section, we compare our results to alternative models used in the literature. The comparisons are reported in Table 5, where column (1) presents our main estimates (baseline and clustered versions).

First, the seminal approach of Glaeser and Maré (2001); Combes et al. (2008) relies on a two-way fixed effects model that decomposes spatial wage disparities into worker heterogeneity and other location-specific factors. Identification in this framework is based on worker mobility across locations. We replicate this approach using our main sample as a benchmark.[20] Column (2) reports the resulting decomposition.

[19] In the first step, we estimate: $Y_{ict} = \alpha_i + \phi_{J(i,t),c} + \text{X'}_{it}\beta + \epsilon_{ict}$. In the second step, we decompose: $\phi_{J,c} = \phi_J + \psi_c + \eta$.

[20] The model we estimate is $Y_{ict} = \alpha_i + \psi_c + \text{X}_{it}\beta + \epsilon_{ict}$.

The main limitation of this approach is that it does not account for the productivity of the establishments where individuals work. As highlighted by Card et al. (2025), mobility between locations is often accompanied by changes in the hierarchy of the establishments where individuals are employed. This change in hierarchy leads to an underestimation of the actual contribution of locations to spatial wage disparities. To address this potential bias, Card et al. (2025) estimate a standard AKM model with worker and establishment fixed effects, identified through worker mobility. They then aggregate establishment fixed effects at the location level. This strategy captures the average wage change for workers moving from the average establishment in their origin location to the average establishment in their destination. However, since establishment and location effects are perfectly collinear in this framework, this method does not allow for the separate identification of establishment and pure location effects, nor does it permit the study of establishment sorting across space.

In column (3), we report the estimates from Card et al. (2025) for the US. In column (4), we apply the same method to our sample for France. The results for France and the US are strikingly similar: location differences explain close to 30% of wage disparities in both countries. However, the interpretation of the location effect in this approach differs from that in our main results, as it includes both the local composition of establishments and the sorting of establishments across locations.

A final benchmark we consider is an AKM model with establishment and location fixed effects, omitting worker heterogeneity:

$$Y_{Jct} = \phi_J + \psi_c + \mathrm{X}_{Jt}\beta + \epsilon_{Jct}, \tag{7}$$

where Y_{Jct} is the (log) average hourly wage paid by establishment J, observed in location c at time t. Identification in this model relies on establishment mobility across locations. This approach mirrors that of Glaeser and Maré (2001); Combes et al. (2008), but instead of relying on worker mobility, it uses establishment mobility. It allows us to separately estimate the contributions of establishments and locations, as well as their sorting patterns. However, it abstracts entirely from worker heterogeneity, both within and across establishments and locations. The results, presented in column (5), indicate that establishments account for 38% of the variance, locations for 26%, and sorting between the two for 40%. However, both establishment and location effects in this model are influenced by worker composition.

Despite these methodological differences, columns (2) to (5) yield similar estimates for the contribution of location effects, although their interpretations vary across specifications. Our model, which explicitly accounts for workers, establishments, and locations (column (1)), demonstrates that much of the location effect in existing approaches is, in fact, driven by differences in establishment characteristics across space and the co-location

of more productive establishments and workers in more productive areas.

6.4 Revisiting the location size premium

We now examine the relationship between city size and wages. The literature has well documented that larger locations tend to offer higher wages, even after controlling for worker characteristics (Glaeser and Gottlieb, 2009; Glaeser and Resseger, 2010; De la Roca and Puga, 2016; Duranton and Puga, 2023). A key question is the extent to which this size effect arises due to the local composition of establishments—e.g., a worker moving to a larger city may also transition to a more productive employer. The location effects estimated from Equation (5) allow us to assess the relationship between city size and wages while accounting for the spatial sorting of both workers and establishments.

First, the elasticity of the average raw local wage to local density is equal to 0.067 (standard error: 0.007—see Appendix Table D.3). We then decompose this elasticity between the respective elasticities of the location effect, local worker effects and local establishment effect to the location density.

The elasticity of the estimated location effect from Equation (5) is equal to 0.007 (standard error: 0.0006), and the correlation is depicted on panel (a) of Figure 9.[21] These results indicate that city size positively affects wages, even after controlling for the type of establishments where workers are employed in cities of different sizes. However, our estimate is two to three times smaller than the estimate we would obtain without controlling for establishments effects. Table D.3 shows that the majority of the relationship between local wages and density is driven by local worker effects and local plant effects—with respective elasticities of 0.038 and 0.022.[22] Panel (b) of Figure 9 further compares our estimated elasticity to previous findings from the literature and the estimate we would obtain without controlling for establishments. Using the approach of Card et al. (2025), we obtain an elasticity to density of 0.026, while a two-way fixed effects model similar to Glaeser and Maré (2001) yields an elasticity of 0.021. These elasticities have distinct interpretations: the elasticity from our main specification captures the effect of location size beyond the role of local workers and establishments, while estimates from the literature reflect the broader impact of moving to a larger city for a given worker, which may also involve transitioning to a more productive establishment.

[21] Panel A of Table Appendix Table D.3 demonstrates that this estimate remains robust across different definitions of size.

[22] Appendix Figure D.5 confirms this result and depicts the distribution of worker and establishment fixed effects, separately for the largest cities and the rest of France.

7 Conclusion

Similar to worker migration, firm relocation is a crucial aspect of the economy and a key mechanism through which economic activity is distributed across space. In this paper, we combine data from France and the U.S. to provide a detailed characterization of firm relocation decisions. Using linked employer-employee data for France, together with information on firm relocations, we examine a central question in labor and spatial economics: to what extent do spatial wage differences across cities arise from "location effects"—such as local geography, infrastructure, and agglomeration—versus the sorting of workers and firms across locations? Our findings indicate that location effects account for only a small fraction of spatial wage disparities, with the majority of the city size premium driven by the local composition of workers and firms.

These findings have several implications for economists and policymakers. First, our results suggest that agglomeration spillovers play a relatively modest role in explaining wage differences, compared the distribution of workers and firms across space. Our estimates provide useful moments for calibrating and disciplining spatial equilibrium models. Second, the significant role of the local composition of establishments in driving spatial wage disparities aligns with the perspective of local policymakers, who view the location decisions of high-paying firms as key determinants of local earnings. Third, the observed relocation patterns suggest that idiosyncratic factors are at play, as firm relocation flows are bidirectional. Fourth, our results emphasize the importance of local human capital in shaping firms' location choices. Finally, from a national policy standpoint, our findings highlight concerns related to inter-locality competition, particularly regarding the use of potentially counterproductive "beggar-thy-neighbor" policies.

This paper leaves several avenues open for future research. First, firm relocation information could be leveraged to evaluate and compare policies aimed at attracting firms, helping to inform the design of optimal local and national business incentives. Second, further research could seek to disentangle the drivers of location effects, particularly assessing the extent to which they reflect agglomeration spillovers rather than local geography and infrastructure.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

_ , _ , Paul Lengermann, and Sebastien Perez-Duarte, "Are Good Workers Employed by Good Firms? A Simple Test of Positive Assortative Matching Models," Econometric Society 2004 North American Winter Meetings 385, Econometric Society August 2004.

Andrews, M. J., L. Gill, Thorsten Schank, and Richard Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society Series A*, 2008, *171* (3), 673–697.

Babet, Damien, Olivier Godechot, and Marco G Palladino, "In the Land of AKM: Explaining the Dynamics of Wage Inequality in France," October 2022. working paper or preprint.

Babina, Tania, Wenting Ma, Christian Moser, Paige Ouimet, and Rebecca Zarutskie, "Pay, Employment, and Dynamics of Young Firms," Working Paper 19-25, Kenan Institute of Private Enterprise Research July 2019.

Badinski, Ivan, Amy Finkelstein, Matthew Gentzkow, and Peter Hull, "Geographic Variation in Healthcare Utilization: The Role of Physicians," Working Paper 31749, National Bureau of Economic Research October 2023.

Bartik, Timothy J, "Using Place-Based Jobs Policies to Help Distressed Communities," *Journal of Economic Perspectives*, 2020, *34* (3), 99–127.

Bartik, Timothy J., "Local Labor Markets Should Be Redefined: New Definitions Based on Estimated Demand-Shock Spillovers," Upjohn Working Papers 24-407, W.E. Upjohn Institute for Employment Research October 2024.

Bergeaud, Antonin and Simon Ray, "Adjustment Costs and Factor Demand: New Evidence from Firms' Real Estate," *The Economic Journal*, 07 2020, *131* (633), 70–100.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Brown, Charles and James L. Medoff, "Firm Age and Wages," *Journal of Labor Economics*, July 2003, *21* (3), 677–698.

Bryan, Kevin A. and Jorge Guzman, "Entrepreneurial Migration," *The Review of Economics and Statistics*, 10 2023, pp. 1–45.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantigying the Impact of Firms On the Relative Pay of Women," *The Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

_ , Jesse Rothstein, and Moises Yi, "Location, Location, Location," *American Economic Journal: Applied Economics*, January 2025, *17* (1), 297–336.

_ , Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality*," *The Quarterly Journal of Economics*, 05 2013, *128* (3), 967–1015.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial Wage Disparities: Sorting Matters!," *Journal of Urban Economics*, 2008, *63* (2), 723–742.

_ , _ , _ , and Sébastien Roux, "Estimating Agglomeration Economies with History, Geology, and Worker Effects," *in Glaeser, E. L. (ed.) Agglomeration Economics, (Chicago, IL: Chicago University)*, 2010, (2), 15–65.

Duranton, Gilles and Diego Puga, "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," *American Economic Review*, December 2001, *91* (5), 1454–1477.

_ and _ , "Urban growth and its aggregate implications," *Econometrica*, 2023, *91* (6), 2219–2259.

Enrico, Moretti, "Local Labor Markets," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1237–1313.

Gaubert, Cecile, "Firm Sorting and Agglomeration," *American Economic Review*, November 2018, *108* (11), 3117–53.

Glaeser, Edward L. and David C. Maré, "Cities and Skills," *Journal of Labor Economics*, 2001, *19* (2), 316–342.

_ and Joshua D. Gottlieb, "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, December 2009, *47* (4), 983–1028.

_ and Matthew G. Resseger, "The Complementarity Between Cities and Skills*," *Journal of Regional Science*, 2010, *50* (1), 221–244.

Hazell, Jonathon, Christina Patterson, Heather Sarsons, and Bledi Taska, "National Wage Setting," NBER Working Papers 30623 2024.

Hjort, Jonas, Xuan Li, and Heather Sarsons, "Across-Country Wage Compression in Multinationals," NBER Working Papers 26788 2020.

Hong, Guangbin, "Two-Sided Sorting of Workers and Firms: Implications for Spatial Inequality and Welfare," 2023.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-Out Estimation of Variance Components," *Econometrica*, September 2020, *88* (5), 1859–1898.

la Roca, Jorge De and Diego Puga, "Learning by Working in Big Cities," *The Review of Economic Studies*, 07 2016, *84* (1), 106–142.

Lhuillier, Hugo, "The Local Root of Wage Inequality," 2022.

Lindenlaub, Ilse, Ryungha Oh, and Michael Peters, "Firm Sorting and Spatial Inequality," NBER Working Paper 30637 2022.

Michelacci, Claudio and Vincenzo Quadrini, "Borrowing from Employees: Wage Dynamics with Financial Constraints," *Journal of the European Economic Association*, 2005, *3* (2/3), 360–369.

Mion, Giordano and Paolo Naticchioni, "The Spatial Sorting and Matching of Skills and Firms," *Canadian Journal of Economics*, 2009, *42* (1), 28–55.

Moretti, Enrico and Moises Yi, "Size Matters: Matching Externalities and the Advantages of Large Labor Markets," Technical Report 2024.

Oh, Ryungha, "Spatial Sorting of Workers and Firms," 2023.

Porcher, Charly, Hannah Rubinton, and Clara Santamaría, "JUE insight: The Role of Establishment Size in the City-size Earnings Premium," *Journal of Urban Economics*, 2023, *136*, 103556.

Schoefer, Benjamin and Oren Ziv, "Productivity, Place, and Plants," *The Review of Economics and Statistics*, 11 2022, pp. 1–46.

Serrato, Juan Carlos Suárez and Owen Zidar, "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms," *American Economic Review*, September 2016, *106* (9), 2582–2624.

Setzler, Bradley and Felix Tintelnot, "The Effects of Foreign Multinationals on Workers and Firms in the United States," *The Quarterly Journal of Economics*, 2021, *136* (3), 1943–1991.

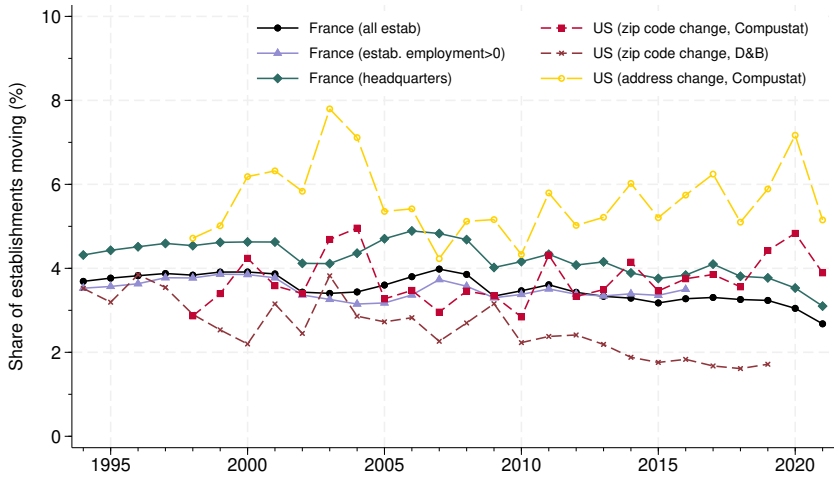
Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming Up Inequality," *The Quarterly Journal of Economics*, 2019, *134* (1), 1–50.

Voget, Johannes, "Relocation of Headquarters and International Taxation," *Journal of Public Economics*, 2011, *95* (9-10), 1067–1081.

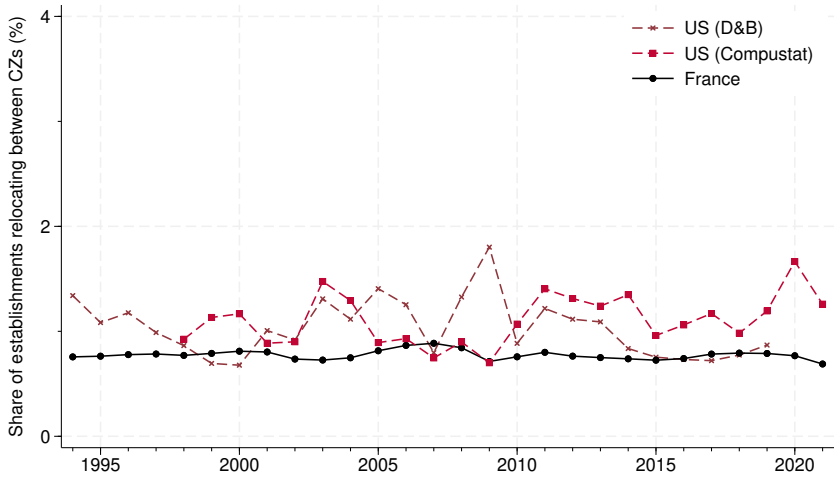
Figures

Figure 1: Share of Establishments Relocating in France and in the US

(a) All moves



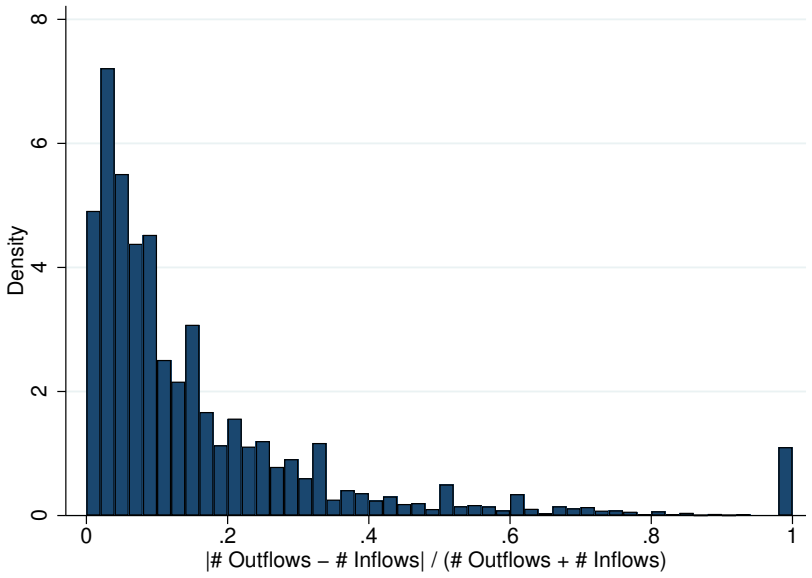
(b) Between-CZ moves



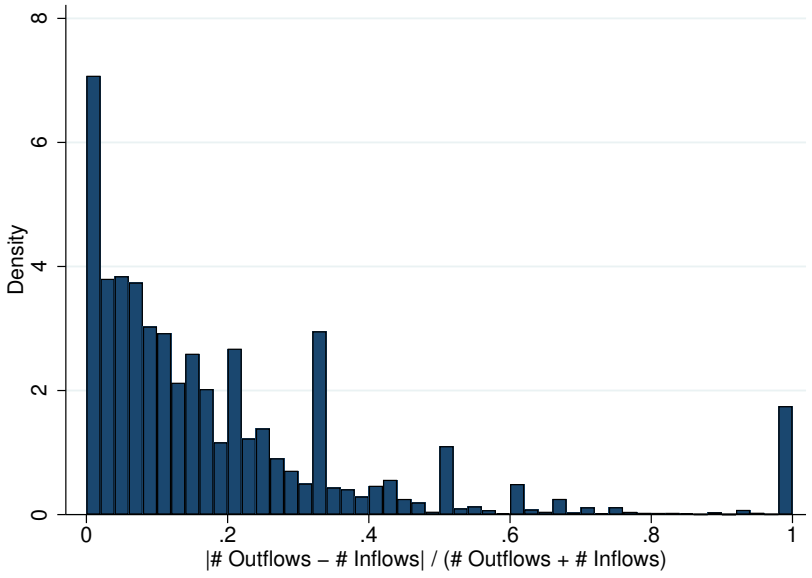
Notes: Figure 1 plots, for each year between 1994 and 2023, the share of establishments relocating as a fraction of the total number of establishments for France (solid lines) and for the United States (dashed lines). The solid dark line with circular markers includes all types of establishments. The solid blue line with diamonds restricts the sample to headquarters. The solid lavender line with triangles considers only establishments with positive employment, as measured in the linked employer-employee data. Turning to the US, the dark-red dashed line with crosses includes all headquarters of firms observed in the Dun and Bradstreet data, while the lighter red dashed line with squares focuses on relocation of headquarters from Compustat firms. Panel (a) includes all relocations while Panel (b) focuses on those between two different commuting zones. Data: Linked employer-employee data, relocation register, register of establishments, Dun and Bradstreet data, Compustat. Go back to main text

Figure 2: Ratio of Net and Gross Relocation Flows for Every Pair of Locations

(a) France



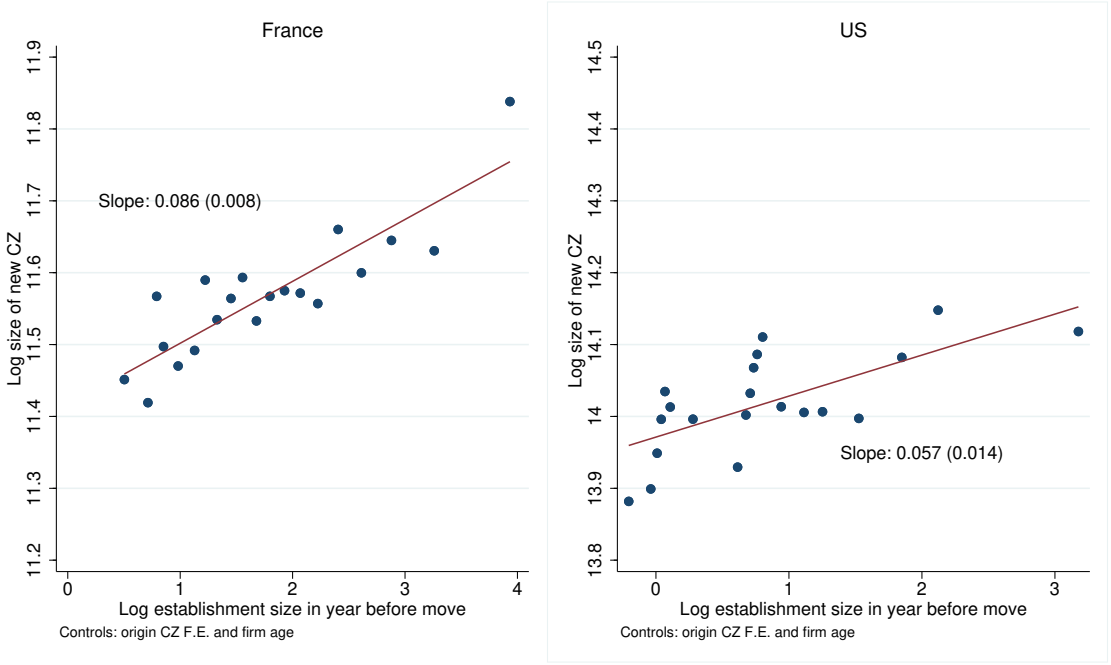
(b) US



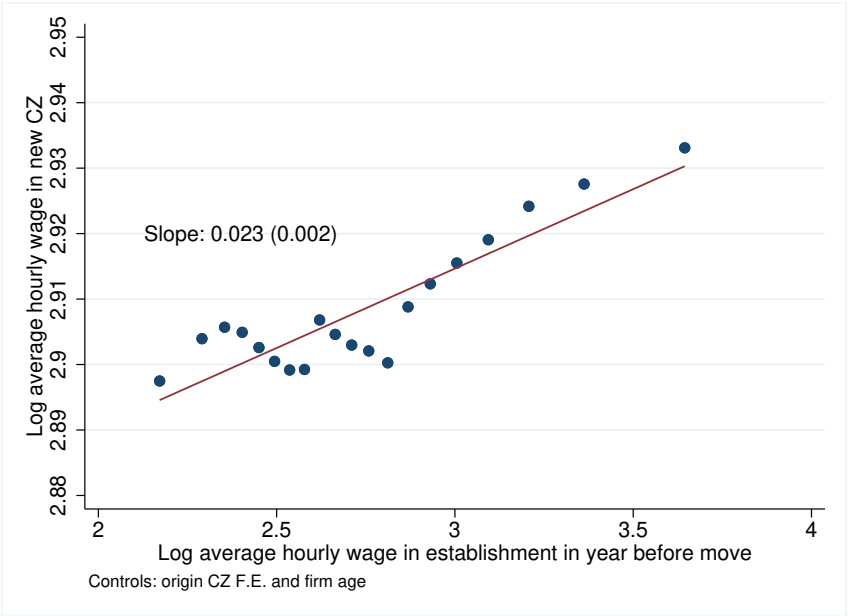
Notes: Figure 2 plots the distribution of the ratio between net and gross flows, for every pair of Provinces (France) or commuting zones (US). This index takes values between 0 and 1. A value of 1 implies unidirectional flows: establishments are either going from location A to B, or from location B to A, but not both. A value of 0 means that entries and exits perfectly offset each other. Every pair is weighted by the total number of moves (outflows + inflows). The index is computed over subperiods of seven years. Panel (a) shows the distribution for France over 1994-2021 and panel (b) for the US over 1994-2019. Data: relocation register (France) and Dun and Bradstreet (US). Go back to main text

Figure 3: Correlation Between Establishment Size and Pay and Destination Size and Pay

(a) Establishment Size and Destination Size (France and US)



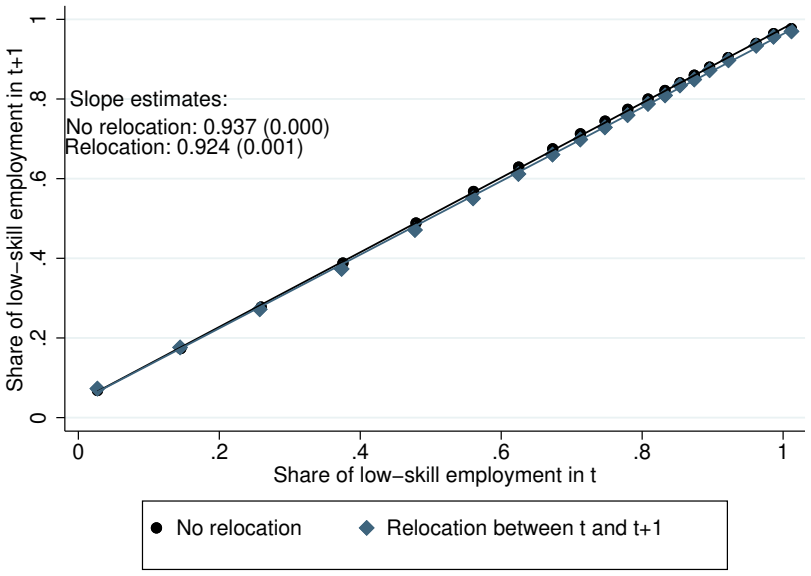
(b) Establishment Wage and Destination Wage (France)



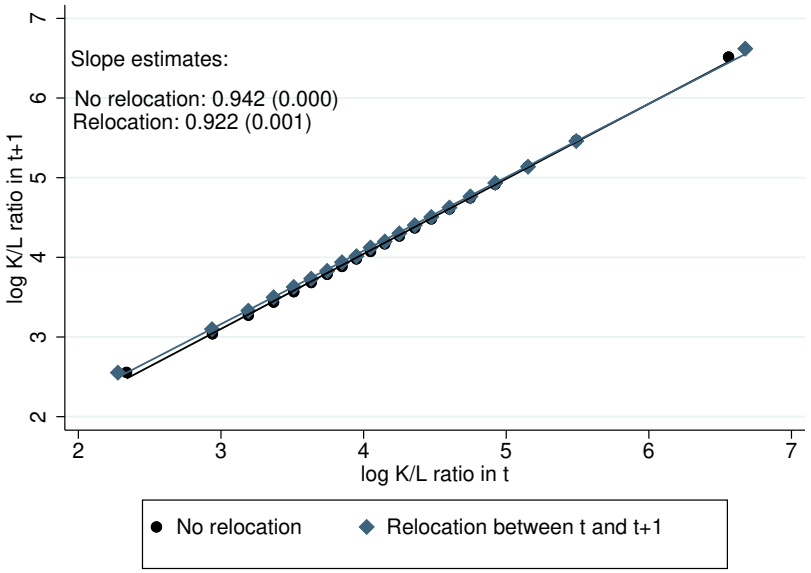
Notes: Figure 3 shows the correlation between characteristics of relocating establishments and characteristics of destination CZs. The fitted lines and slopes report estimates of γ_1 in Equation (3), for $x =$ log size, log average hourly wage. The regression controls for establishment age at the time of the relocation and commuting zone of origin fixed effects. Robust standard errors are reported in parentheses. Panel (a) focuses on log size, while Panel (b) depicts estimates for the log hourly wage. Data: linked employer-employee data and relocation register (France) and Dun and Bradsreet (US). Go back to main text

Figure 4: Evolution of Input Composition for Movers and Non-Movers

(a) Skill Composition



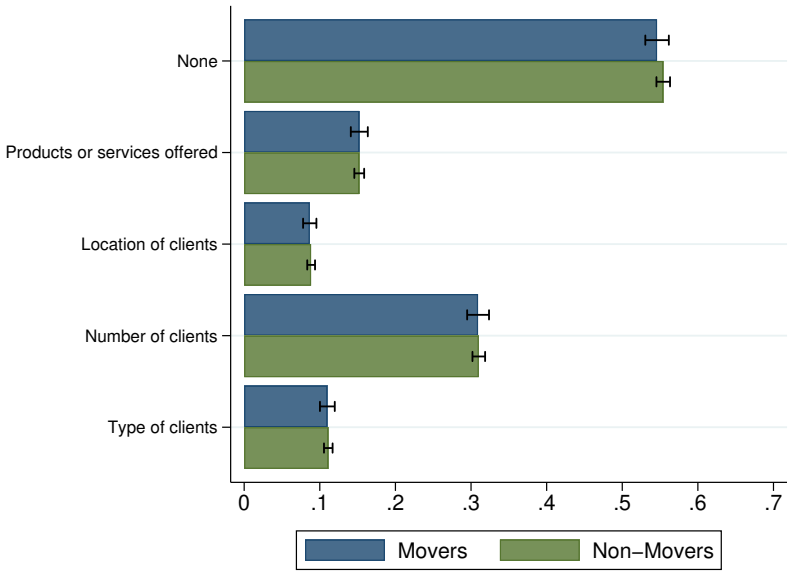
(b) Capital-to-Labor Ratio



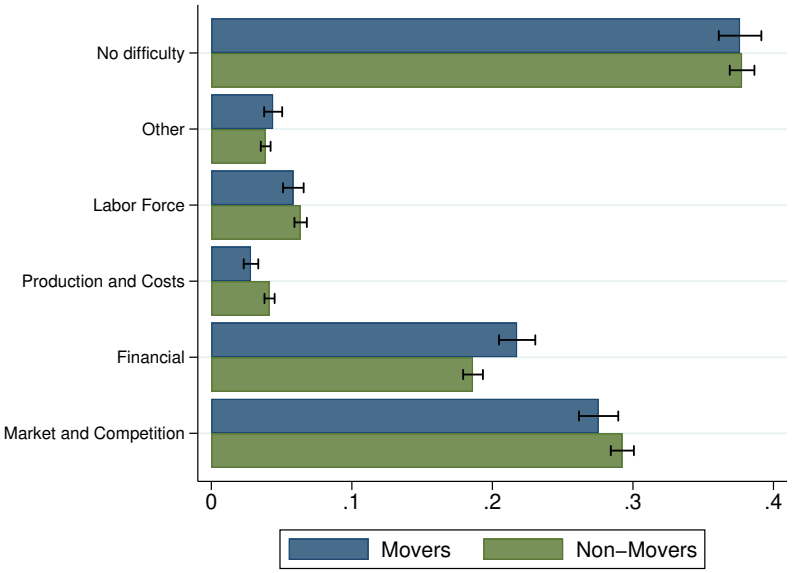
Notes: Figure 4 shows the correlation between an establishment characteristic in a given year (x-axis) and the same characteristic the subsequent year (y-axis), for two continuous variable. In panel (a), the variable under consideration is the share of low-skill workers employed in the establishment, defined as blue-collars and low-skill white collars. In panel (b), the characteristic is the logarithm of the capital to labor ratio. The two panels are binscatter graphs, where the circles are for establishments that do not relocate to a different location between year t and year t+1, while the diamonds are for establishments that experience a relocation. Estimates of the slopes are reported, with robust standard errors in parentheses. Data: DADS and FICUS-FARE. Go back to main text

Figure 5: Changes Faced According to Entrepreneurs

(a) Major Changes Experienced Since Creation



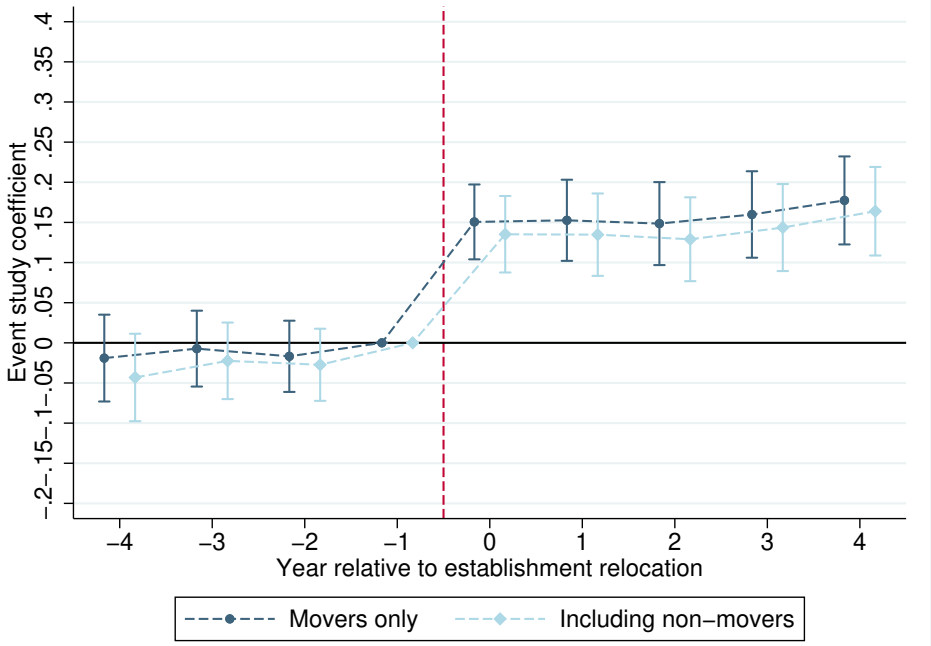
(b) Problems in the First Two Years



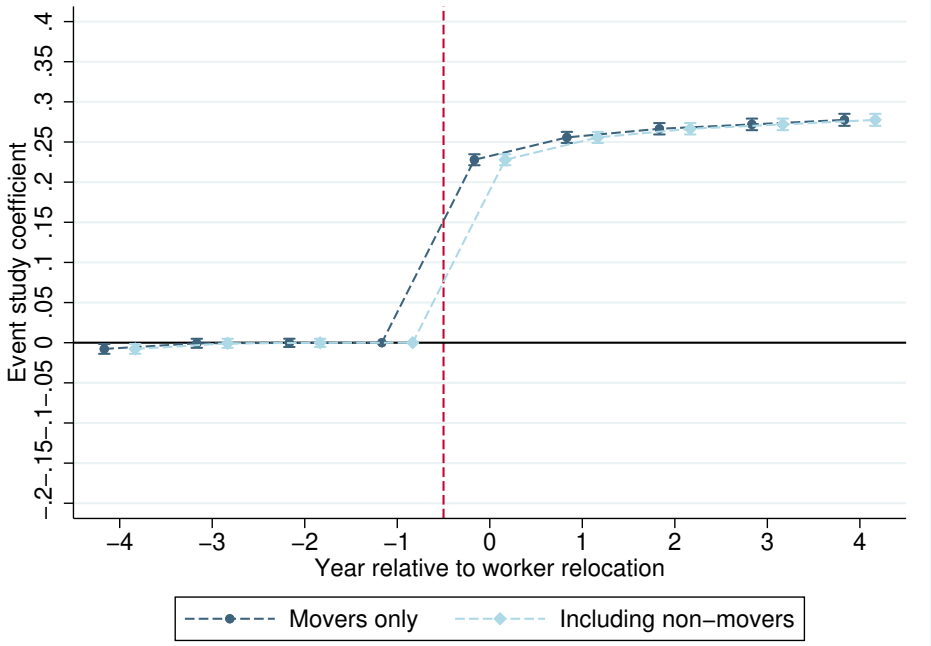
Notes: Figure 5 compares answers from entrepreneurs who relocated at least one establishment during their first five years of economic activity, with answers of entrepreneurs who did not. Panel (a) shows answers to the question "**Did you experience any major change since the creation of the firm?**". Entrepreneurs can report experiencing multiple changes. Panel (b) reports answers to the question "**When you set up your company, did you encounter any difficulties?**". The dark lines correspond to the 95% confidence interval computed using robust standard errors. Appendix Figure B.6 shows additional questions on changes encountered. Data: SINE survey of entrepreneurs (2014) and relocation register. Go back to main text

Figure 6: Wage Changes for Establishments and Workers

(a) Establishment Relocating

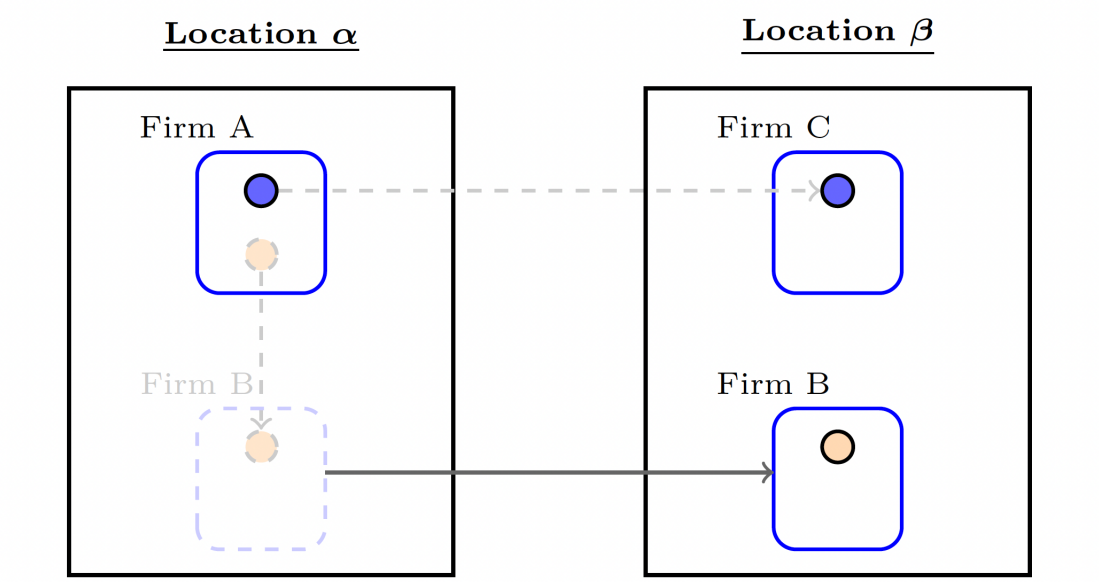


(b) Worker Relocating (and Changing Firm)



Notes: Figure 6 plots estimates of θ in Equation (4) for four types of changes. Panel (a) considers establishments relocating to a different commuting zones. Panels (c)-(d) are worker-level panels focusing on workers changing commuting zone, changing firm, and changing firm within the same commuting zone, respectively. Year 0 is the first year of the change (relocation of change in employer). Each panel reports estimates on two different samples: the one of establishments or worker ever making a change, and the whole sample. Standard errors are clustered at the establishment level. Data: DADS and relocation register. Go back to main text

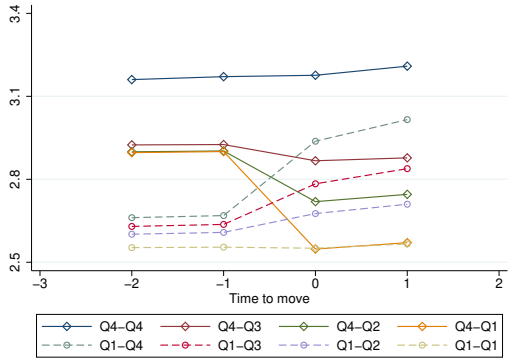
Figure 7: Identification with the Two Sided Movers Design



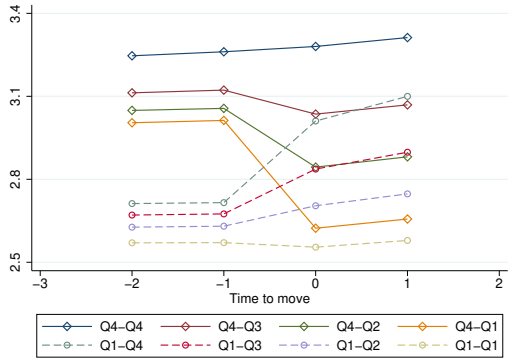
Notes: Figure 7 illustrates the two sided movers design used to estimate equation 5. Rectangles with straight angles represent locations α and β. Rounded rectangles correspond to establishments A, B and C. Finally, the arrows represent the mobility of agents (workers and establishments) between establishments and locations. Go back to main text

Figure 8: Change in Wages by Quartile of Estimated Fixed Effects

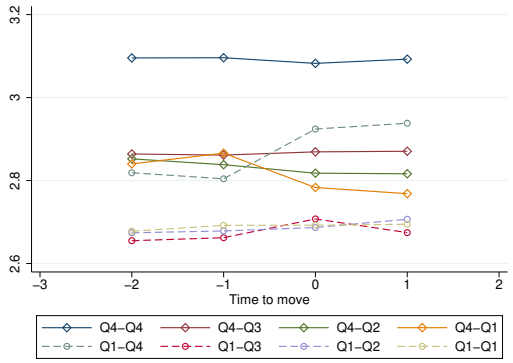
(a) Workers Changing Establishment Within Location



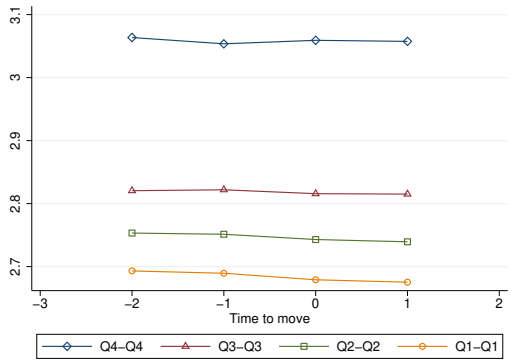
(b) Workers Changing Establishment and Location



(c) Establishments Changing Location



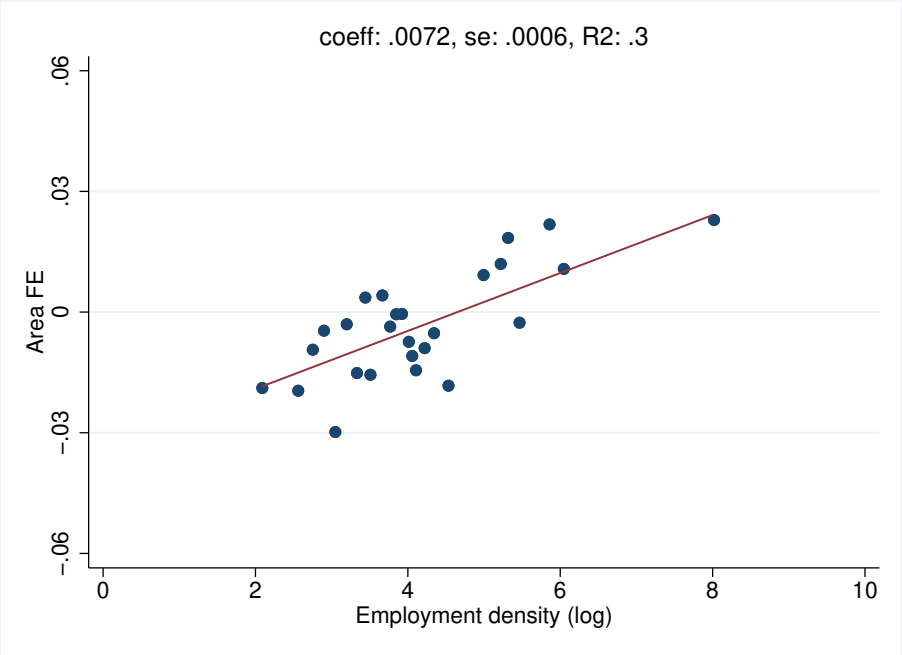
(d) Placebo: Establishments Moving Within Location



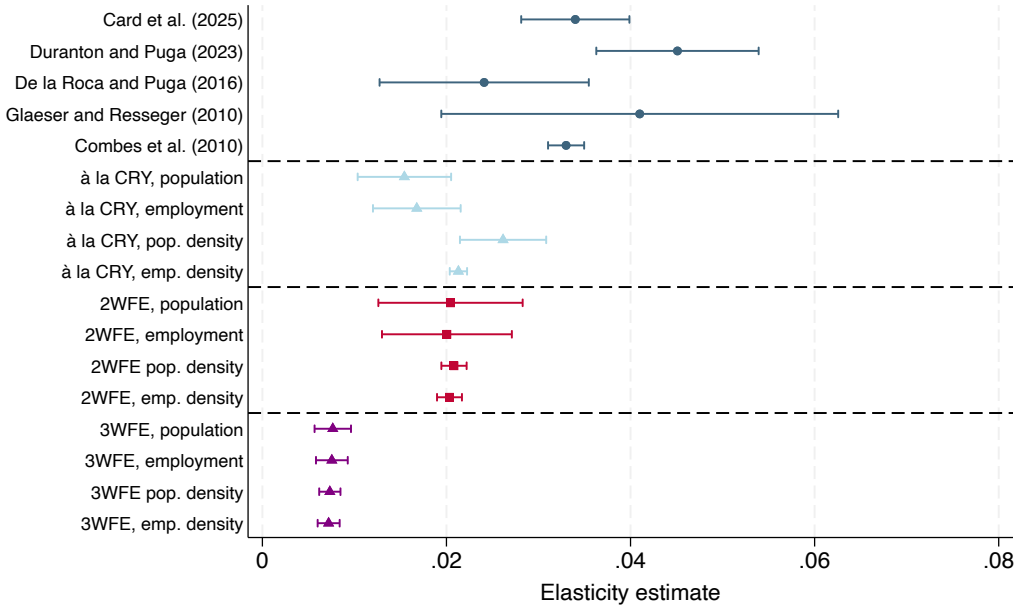
Notes: Figure 8 plots the evolution of the average (log) wages earned by workers and paid by establishments around a relocation, separately for different quartiles of origin and destination fixed effects. Panel (a) focuses on workers moving between high and low paying establishments, within the same commuting zones. Establishments are divided into four quartiles of fixed effects based on estimation of equation 5. Panel (b) reproduces the same figure for workers changing establishment between commuting zones. The quartiles are computed using the sum of establishments and location fixed effects. Panel (c) turns to the average wage paid by establishments following a relocation between commuting zones. The quartiles reflect the sum of area and average worker fixed effects. Panel (d) focuses on establishments relocating within the same commuting zone, hence within the same quartile of fixed effects. Time to move equal to 0 corresponds to the first full year after the move. We focus on a balance sample of workers and establishments observed two years around the move and exclude the transition year. Data: DADS and relocation register (2002-2019). Go back to main text

Figure 9: Location-Size Premium

(a) Estimated Location Effect and Size



(b) Comparison with Different Estimates of the Size Premium



Notes: Figure 9 panel (a) plots the correlation between the location fixed effects estimated from Equation (5) and the log of employment per square kilometer of the commuting zone, observed in the 2015 census. The coefficient, standard error and R^2 come from the estimation of the corresponding linear regression with robust standard errors. Panel (b) compares the resulting slope to alternative measures of the impact of size, for several measures of the location size. The first four estimates are taken from the literature. The "CRY" and "2WFE" result from our own computations, using the method in Card et al. (2025), and a standard two-way fixed effects model. The bottom set of estimates are obtained from our main specification in Equation (5), for different measures of size. Go back to main text

Tables

Table 1: Establishments Preserving their Activity and Organization

	Same Industry	Preserve HQ Status	Same Legal Category	Same Top Manager
Relocation	0.944	0.989	0.957	0.670
No Relocation	0.990	0.998	0.977	0.766

Notes: Table 1 shows the share of establishment keeping identical the characteristic indicated in the first row, between year t and year t+1. Those shares are presented separately for establishments experiencing a relocation between two consecutive years, and for those that do not. For instance, 94.4% of establishment relocating keep the same industry code the year after their relocation, and 99.0% of establishment that do not relocate keep the same code between two consecutive years. Go back to main text

Table 2: Worker-Level and Establishment-Level Summary Statistics

	Workers			Establishments	
	Stayers	Movers (All)	Movers (CZ)	Stayers	Movers (CZ)
Demographics					
Share Female	0.39	0.37	0.32	0.42	0.33
Worker Age	40.93	38.45	37.32	39.35	39.73
Gross hourly Wage					
Mean	18.18	18.75	19.11	16.13	18.83
S.D.	10.05	10.32	10.64	6.64	8.60
Establishment Size					
Mean	346.00	283.41	243.71	11.99	10.85
S.D.	1265.22	1096.77	906.47	60.39	37.88
Occupations					
Blue Collar	0.40	0.34	0.32	0.31	0.32
Clerks	0.28	0.28	0.25	0.38	0.24
Supervisors	0.18	0.20	0.22	0.17	0.21
Managers	0.13	0.18	0.20	0.11	0.19
Executives	0.01	0.01	0.01	0.03	0.04
Industries					
Manufacturing	0.32	0.23	0.21	0.12	0.13
Retail	0.22	0.24	0.24	0.32	0.26
Services	0.35	0.43	0.45	0.42	0.42
Construction	0.11	0.10	0.10	0.14	0.19
Number of Units					
Total Units	23,465,100	11,988,646	5,754,340	2,408,220	56,210
Unit-Year Obs.	79,196,845	98,740,414	47,183,902	14,344,972	471,040
% Changing Firms					
1 Time	.	0.45	0.30	.	.
2 Times	.	0.26	0.28	.	.
3+ Times	.	0.28	0.40	.	.
% Changing Location					
1 Time	.	0.27	0.56	.	0.91
2 Times	.	0.13	0.27	.	0.08
3+ Times	.	0.08	0.16	.	0.01

Notes: Table 2 reports descriptive statistics for workers and establishments. Column 1 includes non-movers (2002–2016), while column 2 covers all movers (across commuting zones and/or establishments). Column 3 shows only workers switching commuting zones. Columns 4 and 5 show establishments that remain or relocate across commuting zones, respectively. Means are reported with standard deviations where applicable. Establishment-level variables (e.g., size, industry) are worker-weighted in columns 1–3. Worker-level variables (e.g., gender, occupation) are firm-averaged and then unweighted across firms in columns 4–5. Go back to main text.

Table 3: Decomposition of Wage Disparities into Worker, Establishment, Location Effects and Sorting

	(1)	(2)	(3)
	Baseline	Clustered	Variance corrected
Standard deviation of log wages	0.113	0.113	0.110
Number of worker-year observations	161,262,266	161,383,668	161,043,566
Number of workers	22,664,335	22,684,178	22,632,016
Number of plants	1,765,559	1,867,831	1,730,018
Number of locations	304	304	304
Std. dev. of worker effect	0.062	0.063	0.061
Std. dev. of plant effects	0.047	0.041	0.046
Std. dev. of location effects	0.024	0.018	0.022
Std. dev. of Xb	0.003	0.002	
Correlation of worker/plant effects	0.725	0.971	0.688
Correlation of plant/location effects	0.275	0.637	0.266
Correlation of worker/location effects	0.514	0.452	0.482
Share of variance of log wages due to:			
Person effects	0.302	0.311	0.311
Plant effects	0.172	0.134	0.175
Location effects	0.045	0.024	0.041
Covariance of worker and plant effects	0.330	0.397	0.347
Covariance of plant and location effects	0.048	0.073	0.052
Covariance of worker and location effects	0.120	0.078	0.124
Xb and other covariances	-0.017	-0.017	-0.019
Residual	0.000	0.000	0.000
RMSE of model	0.16	0.16	.
Adjusted R2	0.88	0.88	.

Notes: Table 3 provides the results of the variance decomposition of the average log gross hourly wage between commuting zones. The formula is detailed in Equation (6). Column (1) reports the results of the plug-in estimator (see Equation (5)), while column (2) provides results of the model with clustered establishment fixed effects. Following Bonhomme et al. (2019) we use a K-means algorithm to cluster establishments into ten groups based on percentiles of their distribution of gross hourly wage. We then estimate one fixed effect per group of establishments. Column (3) reports the decomposition using the split-sample correction method. Go back to main text

Table 4: Robustness: Alternative Samples and Specifications for the Decomposition

	(1) Baseline	(2) Controlling for Plant Age	(3) Controlling for Plant Age (Clustered)	(4) Without Stayers	(5) Without Stayers (Clustered)	(6) Alternative Clustering	(7) Clustered (100)	(8) Alternative Clustering (100)	(9) Two Steps Estimation
Standard deviation of log wages	0.113	0.116	0.116	0.105	0.105	0.113	0.113	0.113	0.113
Number of worker-year observations	161,262,266	152,774,762	152,581,685	146,022,117	146,154,670	161,342,616	161,383,668	161,342,616	161,238,989
Number of workers	22,664,335	22,090,472	21,856,018	21,539,489	21,562,547	22,678,788	22,684,178	22,678,788	22,661,243
Number of plants	1,765,559	1,846,417	1,848,224	1,741,129	1,851,464	1,831,817	1,867,831	1,831,817	1,763,518
Number of locations	304	304	304	304	304	304	304	304	304
Std. dev. of worker effect	0.062	0.063	0.065	0.057	0.058	0.063	0.062	0.062	0.062
Std. dev. of plant effects	0.047	0.047	0.043	0.049	0.039	0.042	0.044	0.045	0.046
Std. dev. of location effects	0.024	0.023	0.018	0.029	0.017	0.017	0.016	0.016	0.025
Std. dev. of Xb	0.003	0.004	0.003	0.003	0.003	0.002	0.002	0.002	0.003
Correlation of worker/plant effects	0.725	0.7435	0.975	0.422	0.967	0.972	0.968	0.969	0.709
Correlation of plant/location effects	0.275	0.336	0.656	-0.0341	0.63	0.623	0.567	0.544	0.305
Correlation of worker/location effects	0.514	0.529	0.485	0.635	0.428	0.435	0.372	0.345	0.564
Share of variance of log wages due to:									
Person effects	0.302	0.295	0.310	0.293	0.305	0.313	0.302	0.300	0.301
Plant effects	0.172	0.164	0.135	0.214	0.138	0.136	0.155	0.160	0.164
Location effects	0.045	0.039	0.023	0.079	0.028	0.023	0.020	0.019	0.048
Covariance of worker and plant effects	0.330	0.327	0.399	0.319	0.397	0.401	0.419	0.424	0.315
Covariance of plant and location effects	0.048	0.054	0.074	-0.009	0.078	0.070	0.063	0.061	0.054
Covariance of worker and location effects	0.120	0.113	0.083	0.128	0.078	0.074	0.058	0.053	0.135
Xb and other covariances	-0.017	0.009	-0.024	-0.024	-0.024	-0.017	-0.017	-0.017	-0.017
Residual	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Number of clusters	.	10	10	.	10	10	100	100	.
RMSE of model	0.164	0.164	0.164	0.164	0.164	0.165	0.164	0.164	.
Adjusted R^2	0.88	0.88	0.88	0.85	0.85	0.85	0.85	0.85	.

Notes: Table 4 compares the results of our main specifications, displayed in column (1), to alternative specifications. Column (2) expands the set of time varying controls in equation 5, and includes a polynomial of order three of the age of the establishment. Column (3) is the clustered version of this specification (10 clusters of establishments). Columns (4) and (5) re-estimate the main model, respectively the clustered model, by excluding workers who continue to work in the same establishment after its relocation. The model is estimated only through establishment relocation and workers mobility between establishments, not within the same establishment moving between locations. Columns (6), (7) and (8) investigate alternative clustering approaches. First, column (6) uses mean and standard deviation of wages as clustering variables, instead of percentiles of the wage distribution. Second, column (7), replicates the standard clustering approach with 100 clusters instead of 10. Column (8) combines approaches (6) and (7) with 100 clusters based on mean and standard deviation of wages within establishments. Finally, column (9) provides a decomposition of between CZ variance of wages using a two steps estimation approach (see footnote 19). Data: linked employer-employee data and register of firm relocation (2002-2019). Go back to main text

Table 5: Comparison with Alternative Approaches

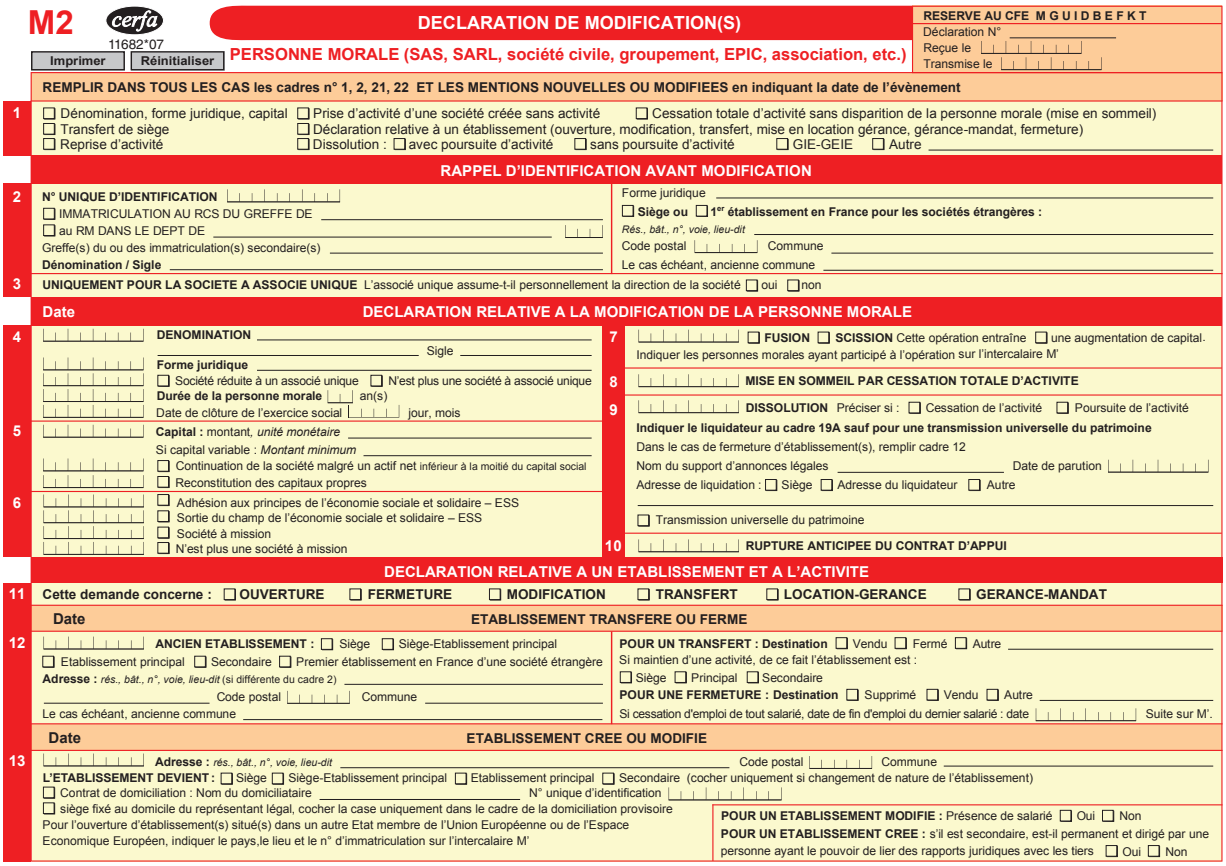
	(1)	(2)	(3)	(4)	(5)
	Baseline	TWFE Worker mobility	CRY (2025)	à la CRY (2025)	TWFE Plant mobility
	FRA	FRA	USA	FRA	FRA
Standard deviation of log wages	0.113	0.114	0.147	0.114	0.091
Number of observations	161,262,266	161,083,676	2,523M quarters	161,083,680	20,613,261
Number of workers	22,664,335	22,594,519	112M	22,594,519	.
Number of plants	1,765,559	.		1,967,420	2,665,472
Number of locations	304	304	741	304	304
Std. dev. of worker effect	0.062	0.0792	0.081	0.0627	.
Std. dev. of plant effects	0.047	.	.	.	0.0562
Std. dev. of location effects	0.024	0.0426	0.079	0.0582	0.0465
Std. dev. of Xb	0.003	0.0026	0.006	0.0025	0.0038
Correlation of worker/plant effects	0.725	.	.	.	.
Correlation of plant/location effects	0.275	.	.	.	0.63
Correlation of worker/location effects	0.514	0.75	0.64	0.80	.
Share of variance of log wages due to:					
Person effects	0.302	0.485	0.303	0.304	.
Plant effects	0.172	.	.	.	0.382
City effects	0.045	0.14	0.293	0.262	0.261
Covariance of worker and plant effects	0.330	.	.	.	.
Covariance of plant and city effects	0.048	.	.	.	0.401
Covariance of worker and city effects	0.120	0.393	0.382	0.45	.
Xb and other covariances	-0.017	-0.018	0.022	-0.017	-0.043
Residual	0.000	0	0	0	0
Number of clusters	.	.	.	.	.
RMSE of model	0.16	0.17	.	0.16	0.18
Adjusted R2	0.88	0.84	.	0.86	0.75

Notes: Table 5 provides a comparison of our main estimates based on equation 5 (column (1)) and alternative approaches. Column (2) displays the results of a two-way fixed effect model with workers and locations. The model is estimated using worker mobility between commuting zones. Column (3) reports the results from the Card et al. (2025) paper based on LEHD data in the US. Column (4) replicates the approach from Card et al. (2025) on our French data. We estimate a standard AKM (1999) model using worker mobility between establishments. We then aggregate the establishments fixed effects at the commuting zone level, that we interpret as a location fixed effect. Column (5) displays the results of a two-way fixed effect model with establishment and location effects (see Equation (7)). Data: linked employer-employee data and register of firm relocation (2002-2019). Go back to main text

Online Appendix

A Context

Figure A.1: 1-Page "M2" Form to Report Establishment Address Change in France



M2 cerfa 11682*07

Imprimer | Réinitialiser

DECLARATION DE MODIFICATION(S)

PERSONNE MORALE (SAS, SARL, société civile, groupement, EPIC, association, etc.)

RESERVE AU CFE M G U I D B E F K T
Déclaration N° ____
Reçue le ____
Transmise le ____

REMPLIR DANS TOUS LES CAS les cadres n° 1, 2, 21, 22 ET LES MENTIONS NOUVELLES OU MODIFIEES en indiquant la date de l'évènement

1 ☐ Dénomination, forme juridique, capital ☐ Prise d'activité d'une société créée sans activité ☐ Cessation totale d'activité sans disparition de la personne morale (mise en sommeil)
☐ Transfert de siège ☐ Déclaration relative à un établissement (ouverture, modification, transfert, mise en location gérance, gérance-mandat, fermeture)
☐ Reprise d'activité ☐ Dissolution : ☐ avec poursuite d'activité ☐ sans poursuite d'activité ☐ GIE-GEIE ☐ Autre ____

RAPPEL D'IDENTIFICATION AVANT MODIFICATION

2 **N° UNIQUE D'IDENTIFICATION** ____
☐ IMMATRICULATION AU RCS DU GREFFE DE ____
☐ au RM DANS LE DEPT DE ____ ____
Greffe(s) du ou des immatriculation(s) secondaire(s) ____
Dénomination / Sigle ____
Forme juridique ____
☐ **Siège ou** ☐ **1er établissement en France pour les sociétés étrangères :**
Rés., bât., n°, voie, lieu-dit ____
Code postal ____ Commune ____
Le cas échéant, ancienne commune ____

3 **UNIQUEMENT POUR LA SOCIETE A ASSOCIE UNIQUE** L'associé unique assume-t-il personnellement la direction de la société ☐ oui ☐ non

Date **DECLARATION RELATIVE A LA MODIFICATION DE LA PERSONNE MORALE**

4 ____ **DENOMINATION** ____ Sigle ____
____ **Forme juridique** ____
____ ☐ Société réduite à un associé unique ☐ N'est plus une société à associé unique
____ **Durée de la personne morale** ____ an(s)
____ Date de clôture de l'exercice social ____ jour, mois

5 ____ **Capital :** montant, *unité monétaire* ____
Si capital variable : *Montant minimum* ____
____ ☐ Continuation de la société malgré un actif net inférieur à la moitié du capital social
____ ☐ Reconstitution des capitaux propres

6 ____ ☐ Adhésion aux principes de l'économie sociale et solidaire – ESS
____ ☐ Sortie du champ de l'économie sociale et solidaire – ESS
____ ☐ Société à mission
____ ☐ N'est plus une société à mission

7 ____ ☐ **FUSION** ☐ **SCISSION** Cette opération entraîne ☐ une augmentation de capital.
Indiquer les personnes morales ayant participé à l'opération sur l'intercalaire M'

8 ____ **MISE EN SOMMEIL PAR CESSATION TOTALE D'ACTIVITE**

9 ____ **DISSOLUTION** Préciser si : ☐ Cessation de l'activité ☐ Poursuite de l'activité
Indiquer le liquidateur au cadre 19A sauf pour une transmission universelle du patrimoine
Dans le cas de fermeture d'établissement(s), remplir cadre 12
Nom du support d'annonces légales ____ Date de parution ____
Adresse de liquidation : ☐ Siège ☐ Adresse du liquidateur ☐ Autre
____
☐ Transmission universelle du patrimoine

10 ____ **RUPTURE ANTICIPEE DU CONTRAT D'APPUI**

DECLARATION RELATIVE A UN ETABLISSEMENT ET A L'ACTIVITE

11 **Cette demande concerne :** ☐ **OUVERTURE** ☐ **FERMETURE** ☐ **MODIFICATION** ☐ **TRANSFERT** ☐ **LOCATION-GERANCE** ☐ **GERANCE-MANDAT**

Date **ETABLISSEMENT TRANSFERE OU FERME**

12 ____ **ANCIEN ETABLISSEMENT :** ☐ Siège ☐ Siège-Etablissement principal
☐ Etablissement principal ☐ Secondaire ☐ Premier établissement en France d'une société étrangère
Adresse : *rés., bât., n°, voie, lieu-dit* (si différente du cadre 2) ____
____ Code postal ____ Commune ____
Le cas échéant, ancienne commune ____
POUR UN TRANSFERT : Destination ☐ Vendu ☐ Fermé ☐ Autre ____
Si maintien d'une activité, de ce fait l'établissement est :
☐ Siège ☐ Principal ☐ Secondaire
POUR UNE FERMETURE : Destination ☐ Supprimé ☐ Vendu ☐ Autre ____
Si cessation d'emploi de tout salarié, date de fin d'emploi du dernier salarié : date ____ Suite sur M'.

Date **ETABLISSEMENT CREE OU MODIFIE**

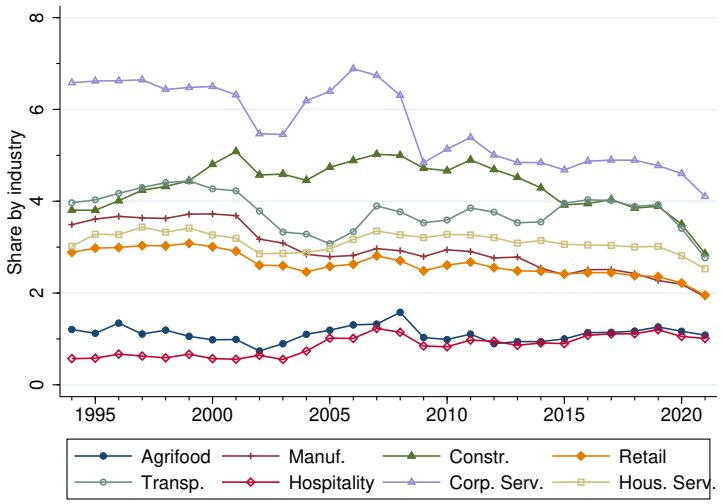
13 ____ **Adresse :** *rés., bât., n°, voie, lieu-dit* ____ Code postal ____ Commune ____
L'ETABLISSEMENT DEVIENT : ☐ Siège ☐ Siège-Etablissement principal ☐ Etablissement principal ☐ Secondaire (cocher uniquement si changement de nature de l'établissement)
☐ Contrat de domiciliation : Nom du domiciliataire ____ N° unique d'identification ____
☐ siège fixé au domicile du représentant légal, cocher la case uniquement dans le cadre de la domiciliation provisoire
Pour l'ouverture d'établissement(s) situé(s) dans un autre Etat membre de l'Union Européenne ou de l'Espace Economique Européen, indiquer le pays,le lieu et le n° d'immatriculation sur l'intercalaire M'
POUR UN ETABLISSEMENT MODIFIE : Présence de salarié ☐ Oui ☐ Non
POUR UN ETABLISSEMENT CREE : s'il est secondaire, est-il permanent et dirigé par une personne ayant le pouvoir de lier des rapports juridiques avec les tiers ☐ Oui ☐ Non

Notes: Figure A.1 presents the one-page form "M2", which entrepreneurs are required to complete when relocating their establishment. The form mandates the completion of three key sections: boxes (1), (2), and (13). Box (1) specifies the type of change being reported, where entrepreneurs must check the option "Déclaration relative à un établissement" (Declaration concerning an establishment) to indicate a relocation. Box (2) details the initial characteristics of the establishment, while Box (13) provides the new address. Entrepreneurs must submit this form as part of the relocation process with the local commercial court within three months of the relocation. Additionally, they are required to include a proof of closure of the former establishment, and evidence of the current or future opening of the new one. Go back to main text

B Additional results on establishment relocation

B.1 Prevalence of establishment relocation

Figure B.1: Establishment Relocation by Industry

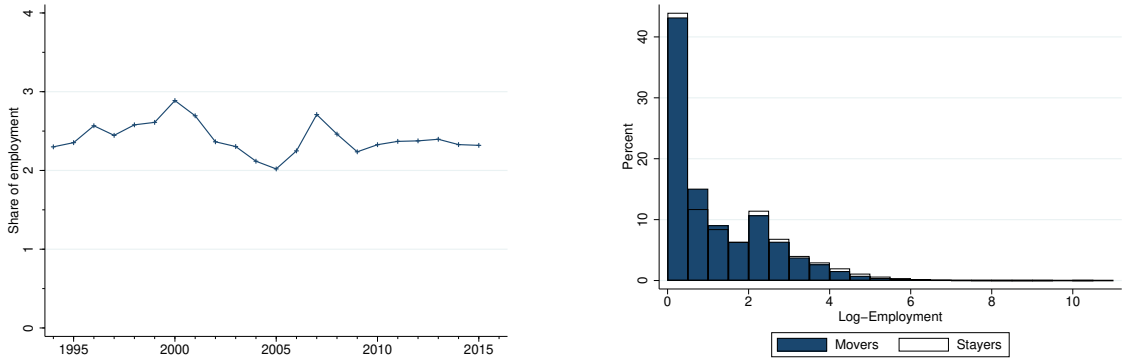


Notes: Figure B.1 plots the share of establishments moving every year by major industry. Data: relocation register. Go back to main text

Figure B.2: Size of establishments relocating in France

(a) Establishment Relocation Weighted by Employment Share

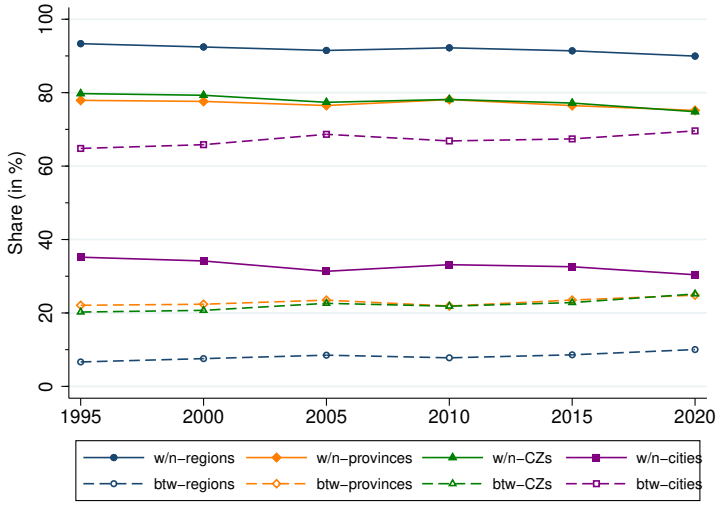
(b) Size distribution



Notes: Figure B.2 provides evidence on the size of establishments relocating. Panel (a) reports the share of establishments moving weighted by their pre-move size, measured in the linked employer-employee data. Panel (b) depicts the establishment (log) size distribution for those that relocate and those that do not. Data: relocation register and linked employer- employee data. Go back to main text

B.2 Distances of relocations

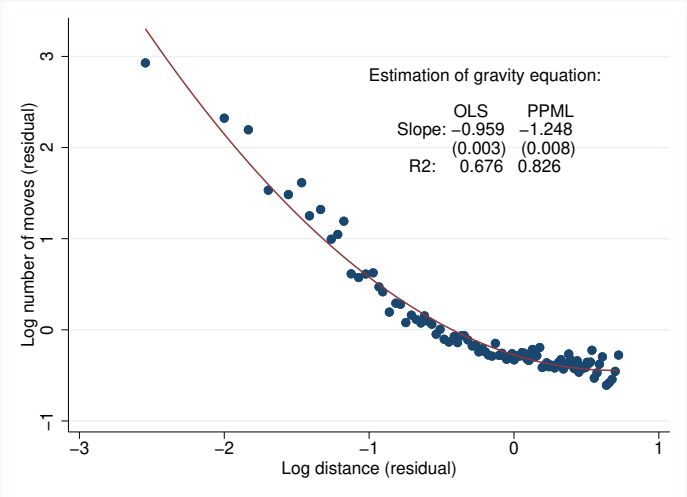
Figure B.3: Establishment Relocation by Geographical Unit



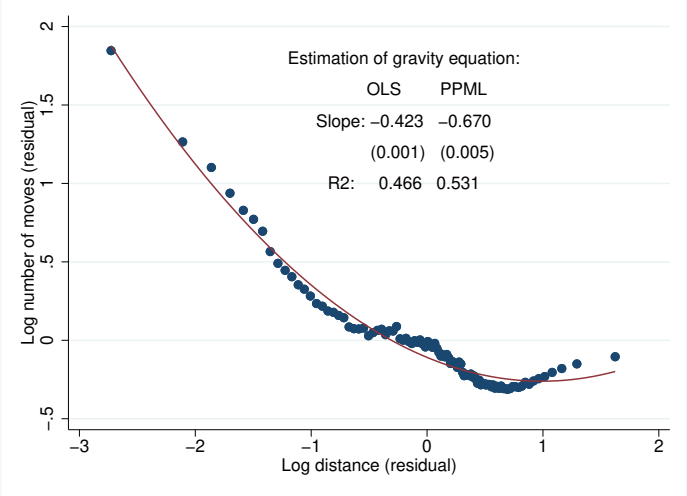
Notes: Figure B.3 reports the share of moves that took place within vs. between locations, for four types of administrative boundaries. There are 22 regions in France, 96 provinces, 304 commuting zones and about 35,000 municipalities. Data: relocation register. Go back to main text

Figure B.4: Gravity patterns

(a) France



(b) US



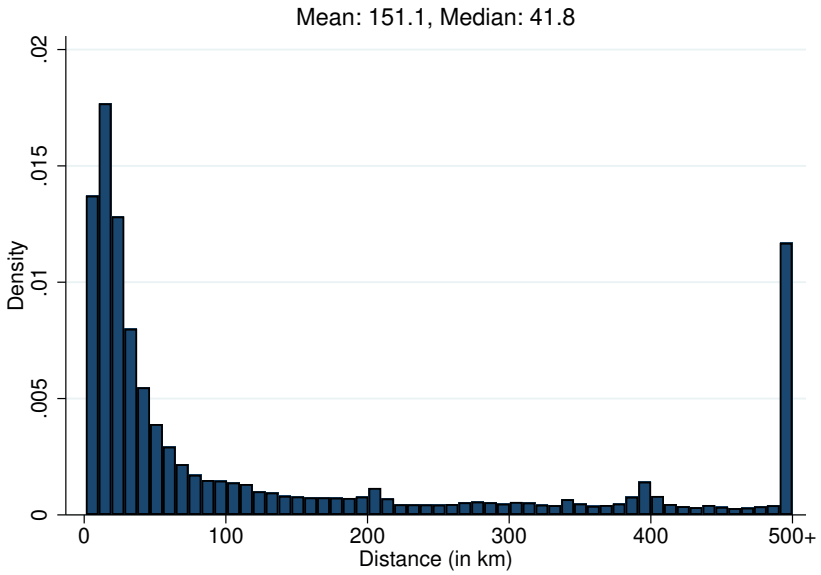
Notes: Figure B.4 is a binscatter that plots the relationship between the (log) number of establishments relocating between every pairs of location and the (log) distance between these two locations. Both quantities are first residualized for origin by year and destination by year fixed effects to account for systematic differences in market access. The corresponding specification is: $log(\#moves_{ijt}) = log(distance_{ij}) + \alpha_{it} + \gamma_{jt} + \epsilon_{ijt}$. The table reports the coefficients of a log-log regression model (column 1) and a pseudo-poison model (column 2). Panel (a) is for France and focuses on moves between every two pairs of provinces (there are 96 of them) over 1996-2021. Panel (b) is for the US and considers pairs of commuting zones over 1994-2019. Data: relocation register (France) and Dun and Bradstreet (US). Go back to main text

Table B.1: Distribution of Distance: Between CZ Moves in France

	P5	P10	P25	P50	Mean	P75	P90	P95
Establishment relocation (in km)	6	9	17	42	151	210	500	639
Worker relocation (in km)	20	21	30	69	178	277	517	636
Minimal distance between CZs (km)	18	21	26	31	32	39	45	49

Notes: Table B.1 reports the distribution of the distance traveled by establishments and workers moving between commuting zones. The third line corresponds to the distribution of the shortest distance that needs to be traveled to change commuting zone in France, for each CZ. The distance is measured in kilometers as the shortest path between the centroids of the two locations along the surface of a mathematical model of the earth. Go back to main text

Figure B.5: Distribution of Distance: Between CZ Moves in France



Notes: Figure B.5 depicts the distribution of relocation distances for establishment moving between different commuting zones. Half of relocations involve a move further than 41.8 kilometers. Relocations exceeding 500 kilometers are grouped inot a single category. Data: relocation register. Go back to main text

B.2.1 Post-relocation evolution of activity

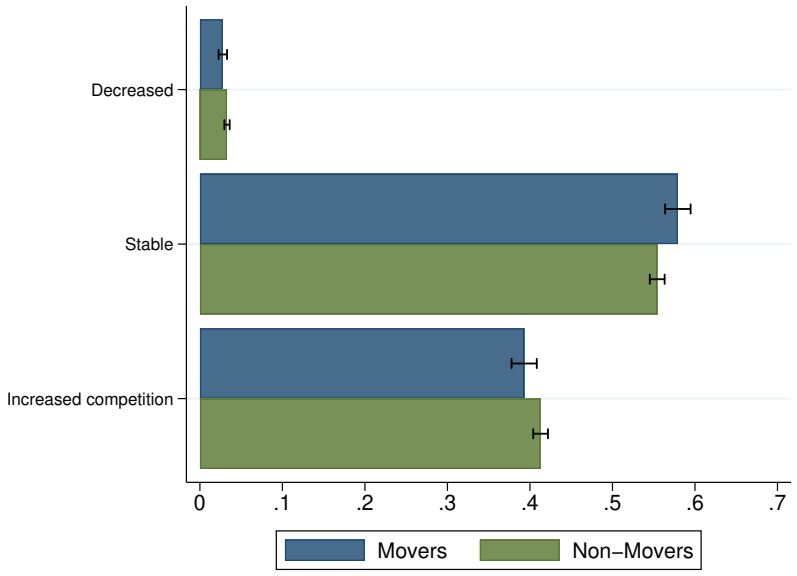
Table B.2: Difference-in-Differences: Establishments Preserving their Activity and Organization

	Same Industry	Preserve HQ Status	Same Legal Category	Same Top Manager
Mover x Year Post Move	-0.044	-0.007	-0.020	-0.098
	(0.000)	(0.000)	(0.000)	(0.001)
Mover	-0.002	-0.003	0.025	0.003
	(0.000)	(0.000)	(0.000)	(0.001)
Year Post Move	-0.003	0.000	-0.005	-0.027
	(0.000)	(0.000)	(0.000)	(0.001)
Year F.E.	Yes	Yes	Yes	Yes

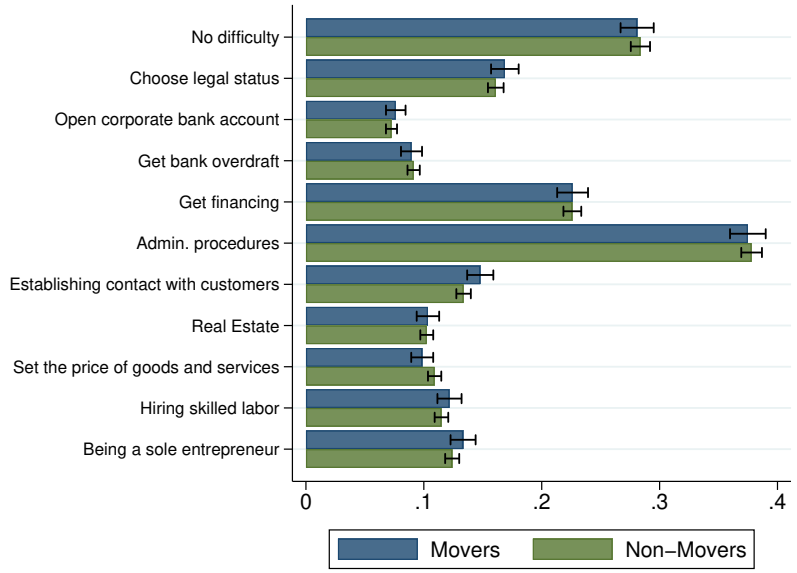
Notes: Table B.2 reports the estimates of a difference-in-difference specification in which the outcome listed in the first row if the outcome variable. Each outcome is a dummy variable equal to one if the establishment keeps a given characteristic identical from one year to the other. The specification is the following: $y_{it} = \delta_t + \beta_1 \text{Mover˙i x Year Post Move}_t + \beta_2 \text{Mover}_i + u_{it}$. For an establishment relocating, the variable Year Post Move$_t$ equals one only the first year post-relocation. The standard errors are clustered at the establishment level and are reported in parentheses. For example, an establishment relocating is 4.4 percentage points less likely to keep an identical industry code in the year following the relocation. Go back to main text

Figure B.6: Changes Faced by Entrepreneurs Relocating their Establishment

(a) Changes in Local Competition



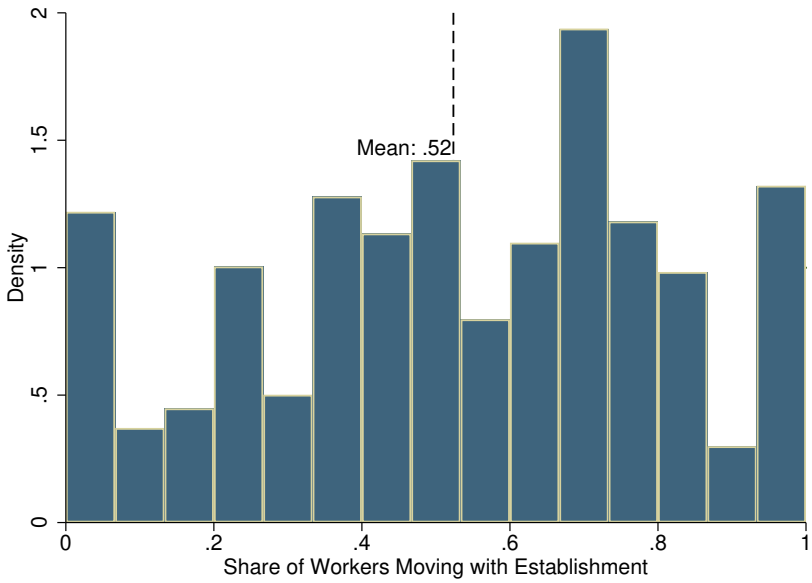
(b) Problems During Creation



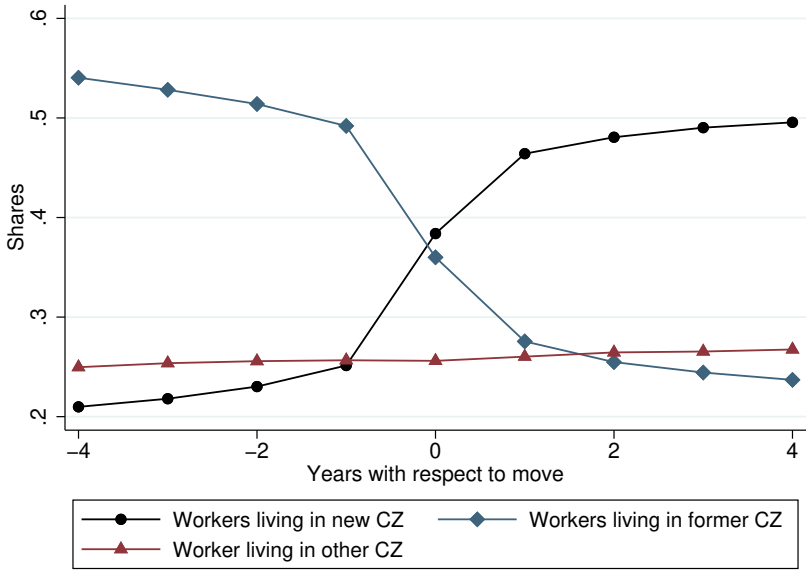
Notes: Figure B.6 compares answers from entrepreneurs who relocated at least one establishment during their first five years of economic activity, with answers of entrepreneurs who did not. There are two questions of interest (on top of the two questions already depicted in Figure 5). Panel (a): "Since the creation of your firm, would you say that the level of competition: (i) increased, (ii) is stable, (iii) decreased?". Panel (b): "Since the creation of your company, what has been THE MAIN OBSTACLE to its development?". Entrepreneurs can choose multiple responses. The solid dark lines correspond to the 95% confidence interval computed using robust standard errors. Data: SINE survey of entrepreneurs (2014) and relocation register. Go back to main text

Figure B.7: Composition of Workforce in Establishment Around Relocation

(a) Share of Workers Staying in Establishment After Relocation



(b) Workforce Living Location



Notes: Panel (a) of Figure B.7 displays the distribution of the share of workers moving with their establishment. It is the ratio between the number of workers in the establishment in both $t-1$ and $t+1$ and employment in $t-1$, where t is the relocation year. Panel (b) depicts the average share of workers residing in the destination commuting zone, the commuting zone of origin, and another commuting zone. Averages are computed for each year before and after relocation. Year 0 corresponds to the year of the move. Data: DADS and relocation register. Go back to main text

Figure B.8: Wage Changes for Workers Changing Firm

(a) Worker changing firm, any location



(b) Worker changing firm, within location



Notes: Figure B.8 plots estimates of θ in Equation (4) for four types of changes. Panel (a) considers establishments relocating to a different commuting zones. Panels (c)-(d) are worker-level panels focusing on workers changing commuting zone, changing firm, and changing firm within the same commuting zone, respectively. Year 0 is the first year of the change (relocation of change in employer). Each panel reports estimates on two different samples: the one of establishments or worker ever making a change, and the whole sample. Standard errors are clustered at the establishment level. Data: DADS and relocation register. Go back to main text

C Model appendix

C.1 Firm Location Decisions and Firm flows

Recall the firm's location decision problem:

$$\max_{n\in\{1...N\}} \pi_n(a) - \kappa^f_{on} + \epsilon^f_n.$$

Since ϵ^f_n is drawn from a Type-1 extreme value distribution with shape parameter ξ^f, the probability that a type-z firm chooses location n conditional on starting in location o is

$$s^f_{on}(z) = \frac{\exp\left(\frac{1}{\xi^f}\left(\pi_n(z) - \kappa^f_{on}\right)\right)}{\sum_{n'=1}^N \exp\left(\frac{1}{\xi^f}\left(\pi_{n'}(z) - \kappa^f_{on'}\right)\right)}.$$

Taking logs, we obtain a gravity structure:

$$\log s^f_{on}(z) = \underbrace{\frac{1}{\xi^f}\pi_n(z)}_{\text{Destination FE}} \underbrace{- \log \sum_{n'=1}^N \exp\left(\frac{1}{\xi^f}\left(\pi_{n'}(z) - \kappa^f_{on'}\right)\right)}_{\text{Origin FE}} - \underbrace{\frac{\kappa^f_{on}}{\xi^f}}_{\text{Bilateral frictions}}.$$

We assume that the mass of firms that start in location o is given by $\bar{M}_o$. We can then write the mass of firms that locate in location n, M_n, as

$$M_n = \sum_{o=1}^N \bar{M}_o \int_0^\infty s^f_{on}(z')\,\mathrm{d}F_Z(z').$$

The distribution of firm types in location n is given by the CDF

$$F_{Z,n}(z) = \sum_{o=1}^N \frac{\bar{M}_o}{M_n} \int_0^z s^f_{on}(z')\,\mathrm{d}F_Z(z').$$

C.2 Worker Location Decisions and Worker flows

Recall the worker's location decision problem

$$\max_{n\in\{1...N\}} w_n(z) - \kappa^w_{on} + \epsilon^w_n.$$

Since ϵ^w_n is drawn from a Type-1 extreme value distribution with shape parameter ξ^w, the probability that a type-a worker chooses location n conditional on starting in location o is

$$s^w_{on}(a) = \frac{\exp\left(\frac{1}{\xi^w}\left(w_n(a) - \kappa^w_{on}\right)\right)}{\sum_{n'=1}^N \exp\left(\frac{1}{\xi^w}\left(w_{n'}(a) - \kappa^w_{on'}\right)\right)}.$$

We assume that the mass of workers that start in location o is given by $\bar{L}_o$. We can then write the mass of workers that locate in location n, L_n, as

$$L_n = \sum_{o=1}^{N} \bar{L}_o \int_0^{\infty} s_{on}^w (a') \, \mathrm{d}F_A (a') .$$

The distribution of worker types in location n is given by the CDF

$$F_{A,n} (a) = \sum_{o=1}^{N} \frac{\bar{L}_o}{L_n} \int_0^{a} s_{on}^w (a') \, \mathrm{d}_A (a') .$$

C.3 Search and Matching

The mass of job-seekers in market n is L_n. The mass of vacancies posted in market n is $V_n = M_n \int_z v_n (z) \, \mathrm{d}F_{Z,n} (z)$, where $v_n (z)$ is the chosen number of vacancies by firms of type z in location n (to be derived later). Denote the market tightness in market n by $\vartheta_n = \frac{V_n}{L_n}$. We assume that the matching function takes the form $L_n^{\eta} V_n^{1-\eta}$. Then the probability that a job-seeker is matched to a vacancy is $p_n = \vartheta_n^{1-\eta}$, and the probability that a vacancy is matched to a job-seeker is $q_n = \vartheta_n^{-\eta}$.

C.4 Firm Expected Profits

The expected profits to the firm from operating in market n are given by

$$\pi_n (z) = \max_v \left\{ v q_n \int_a \mathcal{P}_n (z, a) \left(\beta y_n (z, a) - \mathbb{E} \left[c | c < \beta y_n (z, a) \right] \right) \mathrm{d}F_{A,n} (a) - H (v) \right\},$$

where $F_{A,n} (a)$ is the distribution of worker types in location n, derived from the location decisions of workers. For each posted vacancy, the probability to fill it is $q_n \int_a \mathcal{P}_n (z, a) \, \mathrm{d}F_{A,n} (a)$. For each worker of type a that the firm hires, it gets a constant share of output, $\beta y_n (z, a)$, minus the expected hiring and training cost, $\mathbb{E} \left[c | c < \beta y_n (z, a) \right]$.

A particular simple case is when c^{-1} is distributed Pareto with shape parameter θ, and the vacancy-posting cost is given by $H (v) = \frac{1}{\delta+1} v^{\delta+1}$. In this case, $\pi_n (z)$ equals to

$$\pi_n (z) = \pi_0 q_n^{\left(\frac{1}{\delta} + 1 \right)} \left(\Phi_n \mathcal{A}_n z \right)^{(\theta+1)\left(\frac{1}{\delta} + 1 \right)},$$

where π_0 is a constant subsuming various parameters, and $\mathcal{A}_n \equiv \left(\int_a a^{1+\theta} \mathrm{d}F_{A,n} (a) \right)^{\frac{1}{1+\theta}}$ is a measure of location-n human capital, captured by a power-mean of workers ability in location n, and weighted by the distribution of ability in n. In this case, firm expected profits are a power function of the firm's productivity z, the location productivity Φ_n, the endogenous local stock of human capital $\mathcal{A}_n$, and the matching rate q_n.

C.5 Worker Expected Wages

The expected wage of a worker of type a from choosing location n is

$$w_n (a) = p_n (1 - \beta) \int_z \mathcal{P}_n (z, a) \, y_n (z, a) \, \mathrm{d}F_{Z,n} (z) .$$

C.6 Pareto Distribution

Suppose that c^{-1} is distributed Pareto with shape parameter θ and scale 1, and the vacancy-posting cost is given by $H(v) = \frac{1}{\delta+1}v^{\delta+1}$. In this case,

$$\begin{aligned}\mathcal{P}_n(z,a) &= F_C(\beta y_n(z,a)), \\ &= \Pr(c < \beta y_n(z,a)), \\ &= 1 - \Pr\left(\frac{1}{\beta y_n(z,a)} > \frac{1}{c}\right), \\ &= (\beta y_n(z,a))^{\theta}.\end{aligned}$$

Profits, $\pi_n(z)$, are given by

$$\pi_n(z) = \pi_0 q_n^{\left(\frac{1}{\delta}+1\right)} \left(\Phi_n \mathcal{A}_n z\right)^{(\theta+1)\left(\frac{1}{\delta}+1\right)},$$

where π_0 is a constant subsuming various parameters, and $\mathcal{A}_n \equiv \left(\int_a a^{1+\theta} \mathrm{d}F_{A,n}(a)\right)^{\frac{1}{1+\theta}}$ is a measure of location-n human capital, captured by a power-mean of workers ability in location n, and weighted by the distribution of ability in n. In this case, firm expected profits are a power function of the firm's productivity z, the location productivity Φ_n, the endogenous local stock of human capital $\mathcal{A}_n$, and the matching rate q_n.

The firm's expected wages are given by

$$(1-\beta)\frac{\int_a \mathcal{P}_n(z,a)\, y_n(z,a)\, dF_{A,n}(a)}{\int_a \mathcal{P}_n(z,a)\, dF_{A,n}(a)} = (1-\beta)\Phi_n z \frac{\int_a a^{\theta+1} dF_{A,n}(a)}{\int_a a^{\theta} dF_{A,n}(a)}.$$

Also, $w_n(a)$ equals to

$$w_n(a) = p_n(1-\beta)\beta^{\theta}\left(\Phi_n \mathcal{Z}_n a\right)^{(\theta+1)},$$

where $\mathcal{Z}_n$ is a measure of location-n firm productivity, captured by a power-mean of firms productivity in location n, and weighted by the distribution of productivity:

$$\mathcal{Z}_n \equiv \left(\int_z z^{1+\theta} dF_{Z,n}(z)\right)^{\frac{1}{1+\theta}}.$$

In this case, we have

$$s_{on}^f(z) = \frac{\exp\left(\frac{1}{\xi^f}\left(\pi_0 q_n^{\left(\frac{1}{\delta}+1\right)}\left(\Phi_n \mathcal{A}_n z\right)^{(\theta+1)\left(\frac{1}{\delta}+1\right)} - \kappa_{on}^f\right)\right)}{\sum_{n'=1}^{N}\exp\left(\frac{1}{\xi^f}\left(\pi_0 q_{n'}^{\left(\frac{1}{\delta}+1\right)}\left(\Phi_{n'} \mathcal{A}_{n'} z\right)^{(\theta+1)\left(\frac{1}{\delta}+1\right)} - \kappa_{on'}^f\right)\right)},$$

and

$$s_{on}^w(a) = \frac{\exp\left(\frac{1}{\xi^w}\left((1-\beta)\beta^{\theta} p_n\left(\Phi_n \mathcal{Z}_n a\right)^{(\theta+1)} - \kappa_{on}^w\right)\right)}{\sum_{n'=1}^{N}\exp\left(\frac{1}{\xi^w}\left((1-\beta)\beta^{\theta} p_{n'}\left(\Phi_{n'} \mathcal{Z}_{n'} a\right)^{(\theta+1)} - \kappa_{on'}^w\right)\right)}.$$

C.7 Generalized Pareto Distribution

Now suppose that c^{-1} is distributed according to the Generalized Pareto Distribution, with location μ, scale σ, and shape $\zeta \neq 0$:

$$\Pr\left(\frac{1}{c} < h\right) = 1 - \left(1 + \zeta\frac{h-\mu}{\sigma}\right)^{-1/\zeta}, \quad \text{for } 1 + \zeta\frac{h-\mu}{\sigma} > 0.$$

In this case,

$$\mathcal{P}_n(z,a) = \left(1 + \frac{\zeta}{\sigma}\left(\frac{1}{\beta y_n(z,a)} - \mu\right)\right)^{-1/\zeta}.$$

Then, if $\mu > \frac{\sigma}{\zeta}$,

$$\frac{\partial^2 \log \mathcal{P}_n(z,a)}{\partial z \partial a} \geq 0.$$

I.e., the probability that a random match turns into an employment relationship is log-supermodular in a and z.

D Decomposition of spatial wage disparities

D.1 Sample

Table D.1: Number of Observations and Movers by Commuting Zone

	Mean	P10	P25	P50	P75	P90
Panel A: Unique Observations per CZ						
Establishments	5,937	802	1,250	2,504	5,334	11,212
Workers	98,920	12,851	21,242	39,091	94,022	198,215
Panel B: Number of Moves per CZ						
Establishments	287	29	49	94	202	418
Workers	28,750	3,903	6,008	10,826	24,517	52,628

Notes: Table D.1 reports the distribution of the number of unique observations per CZ (Panel A) and number of movers by CZ (Panel B) in our estimation sample. On average, there are 5,937 unique establishments and 98,920 unique workers (respectively 2,504 establishments and 39,091 workers for the median) by CZ. On average, a commuting zone experiences 287 establishment relocations and 28,750 worker moves (respectively 94 and 10,826 for the median). Data: relocation register and linked employer-employee data. Go back to main text

D.2 Variance correction with split sample

Our model of interest involves a high degree of dimensionality, as it incorporates three sets of fixed effects and comprehensive individual-level data spanning a long period. For this reason, we implement a variance correction method that is computationally efficient. The purpose of correcting the variance is to ensure that our decomposition is not biased by the limited mobility of workers between establishments. However, this concern is mitigated in our setting, as we aggregate fixed effects at the CZ level before computing the variance.

The split sample strategy builds on Kline et al. (2020) and Babet et al. (2022) and requires only two splits of the data and hence two estimations, instead of as many splits

as there are observations in Kline et al. (2020). Intuitively, we obtain two independent, unbiased sets of estimates from two separate estimations. The covariance between these independent estimates provides information about the underlying variance or covariance.

In practice, we implement this method in our three-way fixed effects model as follows:

1. We randomly split our 2002-2016 sample based on worker identifiers, ensuring that each worker appears only in either sample A or sample B.

2. We estimate Equation (5) separately on sample A and obtain estimates $(\hat{\alpha}^A, \hat{\phi}^A, \hat{\psi}^A, \hat{\beta}^A)$, and on sample B and obtain estimates $(\hat{\alpha}^B, \hat{\phi}^B, \hat{\psi}^B, \hat{\beta}^B)$.

3. In sample A, we predict worker fixed effects using the estimates obtained from sample B and average them across all observations for each worker:

$$\hat{\alpha}_i^{A(B)} = \frac{1}{n_i} \sum Y_{ict} - \hat{\phi}^B_{J(i,t)} - \hat{\psi}^B - X'_{it}.\hat{\beta}^B \tag{A.1}$$

4. We combine averages at the CZ level, $(\bar{\hat{\alpha}}^k, \bar{\hat{\phi}}^k, \hat{\psi}^k, \bar{\hat{\beta}}^k)$, for k =A, B.

5. We use information from both samples to estimate the variance and covariance terms as follows:

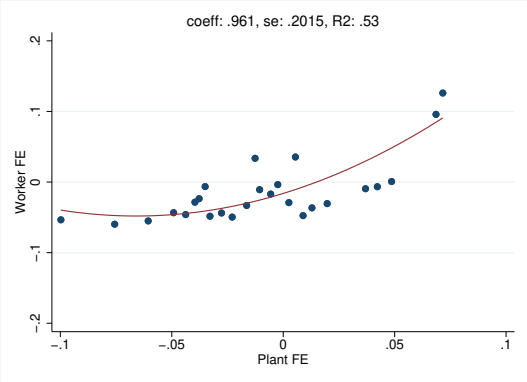
$$\begin{aligned} Var(\overline{\alpha_c}) &= Cov(\bar{\hat{\alpha}}^{A(B)}, \bar{\hat{\alpha}}^A) \\ Var(\overline{\phi_{J(c)}}) &= Cov(\bar{\hat{\phi}}^A, \bar{\hat{\phi}}^B) \\ Var(\psi_c) &= Cov(\hat{\psi}^A, \hat{\psi}^B). \end{aligned}$$

   The covariance between different sets of fixed effects is computed by taking each set from a different estimation. For example, the covariance between worker and plant fixed effects is calculated as: $Cov(\psi_c, \overline{\phi_{J(c)}}) = Cov(\hat{\psi}^A, \bar{\hat{\phi}}^B)$.

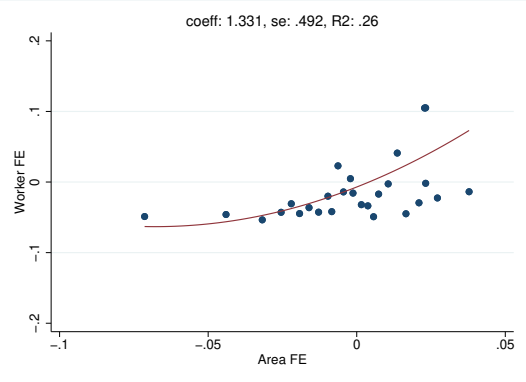
D.3 Model estimates

Figure D.1: Correlation Between Estimated Fixed Effects

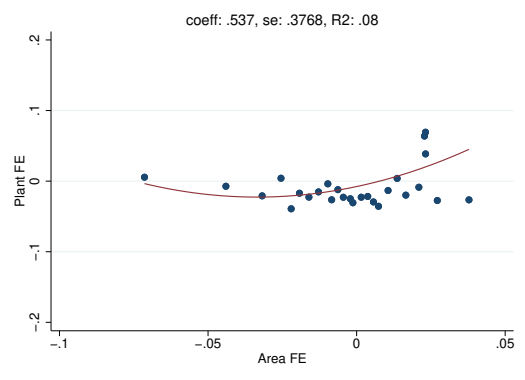
(a) Worker and Establishment



(b) Worker and Location



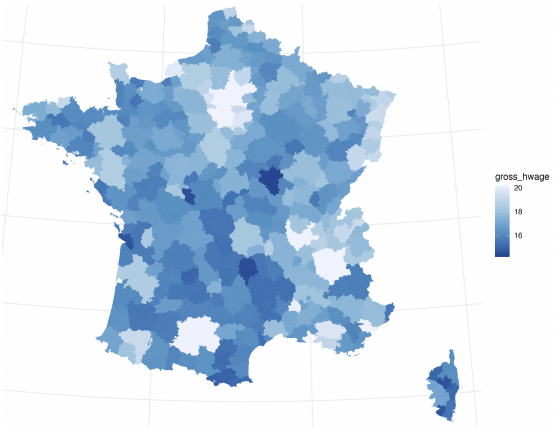
(c) Establishment and Location



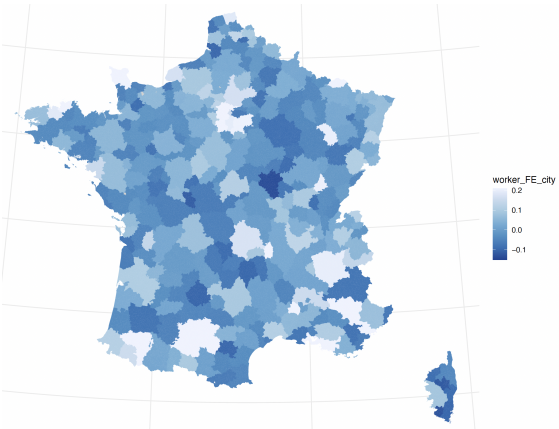
Notes: Figure D.1 shows the result of a binscatter that displays the correlation between average worker, establishment and area fixed effects, between commuting zones. The fixed effects result from the estimation of equation 5. The coefficients, standard errors and R^2 are derived from a linear regression using the underlying micro data at the commuting zone level.Data: DADS and relocation register (2002-2019).
Go back to main text

Figure D.2: Spatial Disparities in Wages and Estimated Fixed Effects

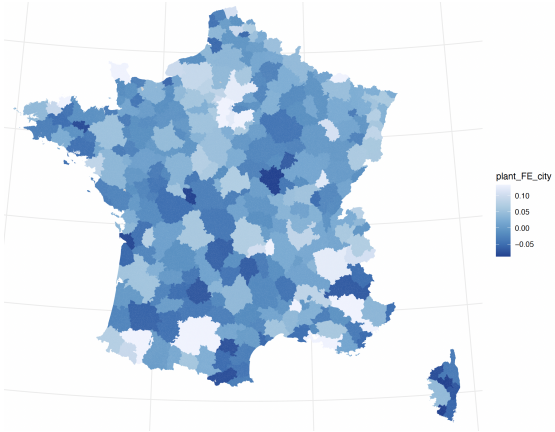
(a) Gross Hourly Wage



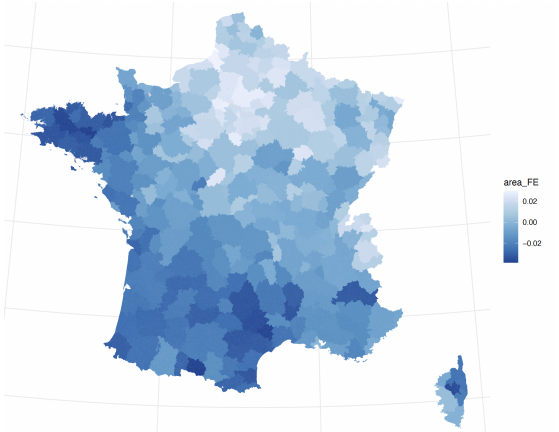
(b) Average Worker Fixed Effects



(c) Average Establishment Fixed Effects



(d) Location Fixed Effects



Notes: Figure D.2 plots the distribution of wages between commuting zones. Panel (a) displays the average gross hourly wage per CZ. Panel (b), (c), (d), plots respectively the average worker, establishment and location fixed effects per CZ. The fixed effects are estimated based on equation 5. Go back to main text

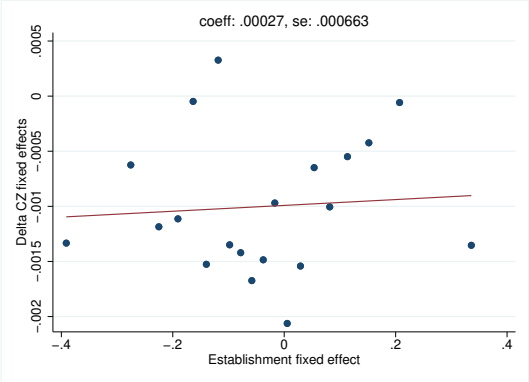
Table D.2: Three-Way Decomposition by Industry

	(1)	(2)	(3)	(4)
	All	Manufacturing	Lower-Skill Services	Higher-Skill Services
Standard deviation of log wages	0.113	0.145	0.099	0.158
Number of worker-year observations	161,383,668	44,576,542	67,237,473	30,615,442
Number of workers	22,699,490	6,140,108	10,806,411	5,479,442
Number of plants	1,867,831	226,763	1,030,656	364,242
Number of locations	304	304	304	304
Std. dev. of worker effect	0.063	0.099	0.046	0.098
Std. dev. of plant effects	0.041	0.049	0.038	0.05
Std. dev. of location effects	0.018	0.019	0.02	0.022
Std. dev. of Xb	0.002	0.003	0.002	0.006
Correlation of worker/plant effects	0.971	0.971	0.965	0.99
Correlation of plant/location effects	0.637	-0.024	0.818	0.684
Correlation of worker/location effects	0.452	-0.233	0.657	0.588
Share of variance of log wages due to:				
Person effects	0.311	0.467	0.216	0.391
Plant effects	0.134	0.112	0.149	0.101
Location effects	0.024	0.017	0.042	0.019
Covariance of worker and plant effects	0.397	0.445	0.347	0.391
Covariance of plant and Location effects	0.073	-0.002	0.129	0.06
Covariance of worker and Location effects	0.078	-0.042	0.125	0.101
Xb and other covariances	-0.017	0.001	-0.008	-0.056
Residual	0	0	0	0
Number of clusters	10	10	10	10
RMSE of model	0.164	0.157	0.156	0.166
Adjusted R^2	0.88	0.86	0.84	0.88

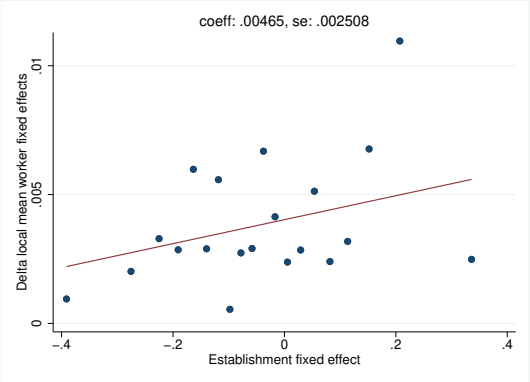
Notes: Table D.2 reproduces the variance decomposition exercise from equation 6, separately by major industry. Lower-skill services include retail, accommodation, food services and other services to households. Higher skill services include services to firms and real estate services. Go back to main text

Figure D.3: Correlation between Fixed Effects of Establishment Moving and Change in Fixed Effects of:

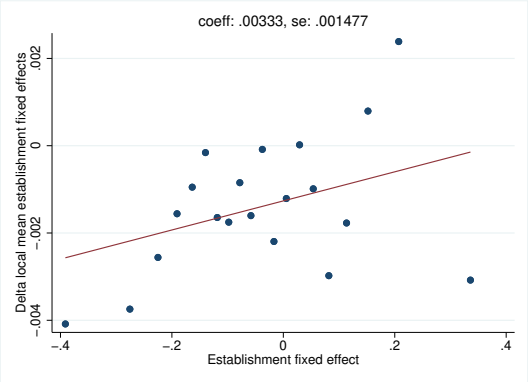
(a) Location



(b) Local Workforce

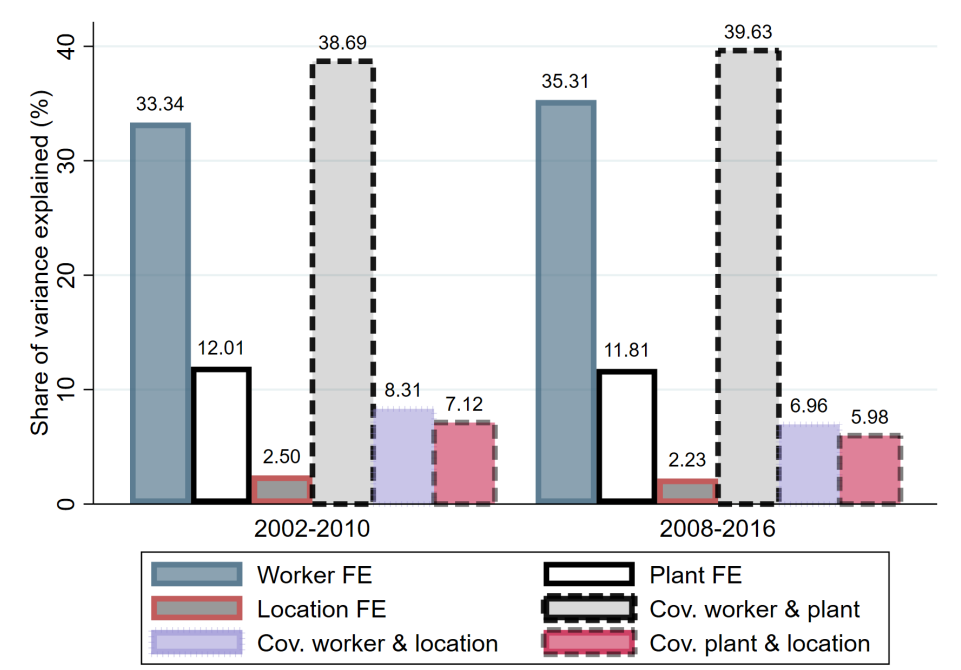


(c) Local Establishments



Notes: Figure D.3 depicts the correlation between the estimated fixed effect of the establishment relocation (x-axis) and the change in average fixed effects between destination and initial commuting zones (y-axis). The change considered is the difference between location fixed effects (panel (a)), between average local worker fixed effects (panel (b)) and average local establishment fixed-effect (panel (c)). The average local fixed effects are measures the year after the relocation for the initial commuting zone, and the year before relocation for the destination, to avoid contamination from the establishment itself. Go back to main text

Figure D.4: Decomposition of Spatial Disparities over Two Periods



Notes: Figure D.4 presents the results of two separate decompositions, resulting from the estimation of Equation (5) over 2 periods. The first period is 2001-2010 and the second one is 2008-2016. Each bars depicts the share of the variance of mean wages between commuting zones explained by a given component. For each period, the first three bars are the variance components (for workers, plants and locations). The subsequent three bars are the covariances (multiplied by two). We omit the components related to the controls (age and year). For example, over 2002-2010, heterogeneity in workers between commuting zones explains 33.34% of the variance, and it explains 35.31% over 2008-2016. Go back to main text

D.3.1 Additional evidence on the size premium

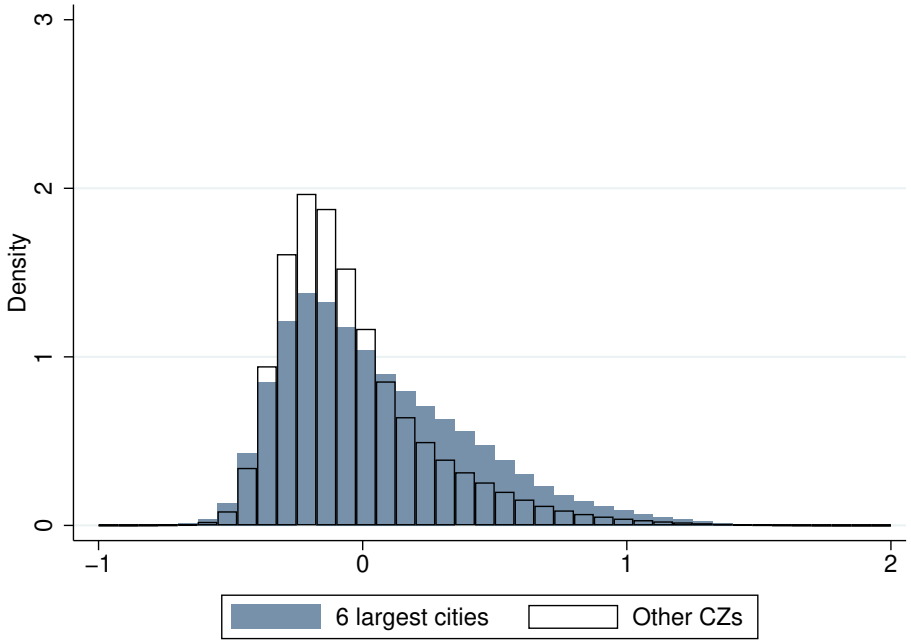
Table D.3: Elasticity of Location Effects With Respect to Density

	(1) Population (log)	(2) Employment (log)	(3) Population density (log)	(4) Employment density (log)	(5) College share	(6) College density (log)
Panel A: Correlation between local wages and size						
Average Wage	0.06709***	0.06553***	0.05823***	0.05707***	0.00958***	0.05180***
	(0.00751)	(0.00650)	(0.00235)	(0.00209)	(0.00076)	(0.00184)
Observations	304	304	304	304	304	304
Adj R^2	0.66	0.69	0.76	0.78	0.74	0.79
Panel B: Location effect (from 3WFE model) and size						
Area Fixed Effect	0.00765***	0.00755***	0.00733***	0.00720***	0.00105***	0.00635***
	(0.00101)	(0.00088)	(0.00059)	(0.00061)	(0.00013)	(0.00050)
Observations	304	304	304	304	304	304
Adj R^2	0.20	0.22	0.29	0.30	0.21	0.28
Panel C: Local worker effects (from 3WFE model) and size						
Worker Fixed Effect	0.03838***	0.03734***	0.03057***	0.02999***	0.00569***	0.02788***
	(0.00325)	(0.00278)	(0.00173)	(0.00149)	(0.00023)	(0.00113)
Observations	304	304	304	304	304	304
Adj R^2	0.72	0.74	0.69	0.71	0.87	0.76
Panel D: Local plant effects (from 3WFE model) and size						
Plant Fixed Effect	0.02163***	0.02120***	0.02095***	0.02049***	0.00293***	0.01811***
	(0.00393)	(0.00353)	(0.00094)	(0.00086)	(0.00052)	(0.00082)
Observations	304	304	304	304	304	304
Adj R^2	0.39	0.41	0.56	0.57	0.39	0.55

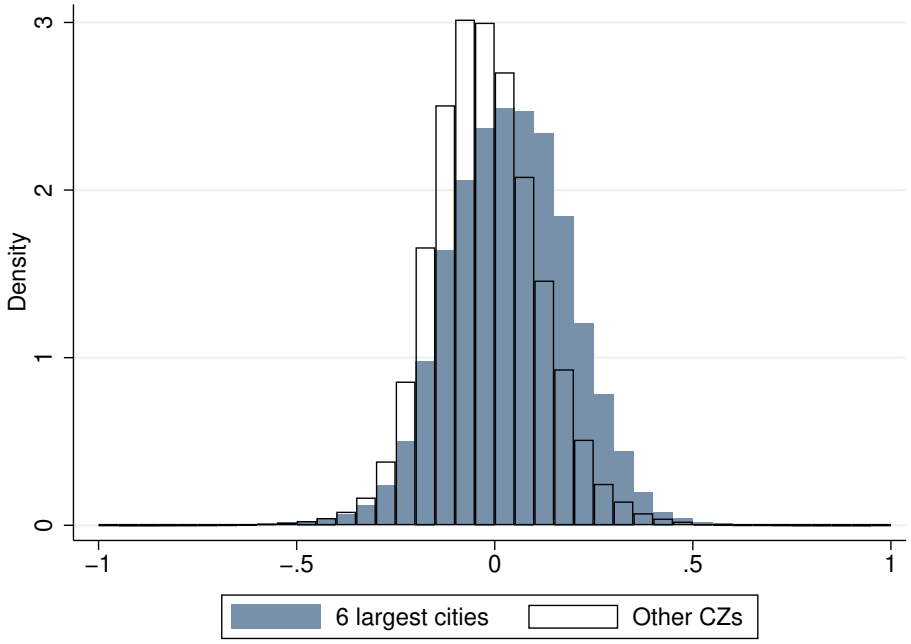
Notes: Table D.3 reports the results of a linear regression of the wage and estimated fixed effects on different measures of location density: (1) log of population, (2) log of employment, (3) log of population per meter square, (4) log of employment per square kilometer, (5) share of college graduates, (6) log of college graduates per square kilometer, observed in the 2015 census. The outcome of the regression is the average local wage (Panel A), the location fixed effect (Panel B), the average local worker fixed effects (Panel C) and the average local plant fixed effects (Panel D). All fixed effects have been estimated in Equation (5). The standard errors are robust to heteroskedasticity. Go back to main text

Figure D.5: Distribution of Worker and Establishment Fixed Effects by Location Size

(a) Local Worker Fixed Effects



(b) Local Establishment Fixed Effects



Notes: Figure D.5 plots estimated worker (panel (a)) and establishment (panel (b)) fixed effects, separately for the six largest French commuting zones, and all other commuting zones. Data: DADS and relocation register. Go back to main text